

RESEARCH ARTICLE OPEN ACCESS

Merging Weather Surveillance Radar Precipitation Estimates From Different Sources: A Quality-Index Approach

David R. L. Dufton¹  | Tamora D. James²  | Mark Whitling³ | Ryan R. Neely III^{1,4} 

¹National Centre for Atmospheric Science, University of Leeds, Leeds, UK | ²Centre for Environmental Modelling and Computation, University of Leeds, Leeds, UK | ³Environment Agency, York, UK | ⁴Institute for Climate and Atmospheric Science, University of Leeds, Leeds, UK

Correspondence: David R. L. Dufton (david.dufton@ncas.ac.uk)

Received: 26 November 2024 | **Revised:** 10 April 2025 | **Accepted:** 9 June 2025

Funding: This work was supported by Flood and Coastal Risk Management Joint R&D Programme sponsored by the Environment Agency, Natural Resources Wales, Defra and the Welsh Government (FRS21237). David Dufton's time on the project was also funded by the National Environment Research Council (NERC) project HydroJULES (NE/X019063/1) and NCAS's core NERC funding. The views expressed are of the authors and do not reflect the position of the funders.

Keywords: precipitation estimation | quality | radar data compositing | weather radar

ABSTRACT

Weather surveillance radar (WSR) provide distributed quantitative precipitation estimates (QPEs) of great value to the modelling, understanding and management of many hydro-meteorological processes. To obtain these observations over regional or larger scale domains it is necessary to composite data from multiple WSRs. These composites are often produced operationally by national or international meteorological agencies yet valuable data from ad-hoc sources such as research groups and local-level WSR operators are not included in these products. This study presents a methodology for incorporating data from a research radar deployment (the National Centre for Atmospheric Science mobile X-band weather radar, NXPOL-1) into a national scale composite (the UK Met Office British Isles gridded composite) using a quality-index. Firstly a quality-index is developed for NXPOL-1 using an intuitive, multi-factor approach. The quality-index is then cross-referenced with the existing quality-index for the national composite, to allow production of a dynamically merged two source WSR QPE. The method developed is then evaluated using surface precipitation measurements from an extensive rain gauge network. Merging QPE from the two sources using a quality-index improves the accuracy of WSR QPE when compared to either individual data source, showing it is possible to combine ad-hoc WSR data with national products dynamically such that precipitation estimation is improved. Improving local QPE using additional radar deployments will benefit flood forecasting accuracy and local incident response, particularly when that data is used to enhance existing coverage.

1 | Introduction

Quantitative precipitation estimates (QPEs) from weather surveillance radars (WSRs) are crucial observations for nowcasting, flood-risk modelling, water resources management and climatological assessment (Ravuri et al. 2021; Saltikoff, Friedrich, et al. 2019; Fabry 2018; He et al. 2011; Cole and Moore 2009).

These observations are a critical component of flood incident response with better nowcasts improving the effective lead-time of flood forecasts and distributed observations allowing a more targeted incident response (Kox et al. 2018; Werner and Cranston 2009). WSRs are particularly suited to these applications as they provide high-resolution (sub-hourly, kilometre scale) remote observations over wide geographic areas. To cover

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 Crown copyright and The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society. This article is published with the permission of the Controller of HMSO and the King's Printer for Scotland.

even larger areas, observations from multiple WSRs are often combined into a single dataset ranging from regional to even continental scales (Jurczyk et al. 2019). These datasets transform the original polar coordinate WSR moment data into a processed gridded surface precipitation estimate, which is far more accessible to the wider scientific community, government agencies and the general public (Villarini and Krajewski 2010).

The accuracy of these gridded datasets depends not only on the underlying QPE processing employed in transforming radar moments into precipitation intensity, on which there is an extensive body of growing research (Germann et al. 2022; Ryzhkov et al. 2022; Fabry 2018), but also on the procedures used for gridding and compositing the WSR observations. The use of a quality indicator during gridding and compositing, calculated based on the underlying WSR processing and radar parameters, has been shown to improve accuracy in comparison to using simple approaches such as the data from the nearest radar, the maximum intensity observed by any radar or even distance weighted averaging of all overlapping radars (Barbieri et al. 2022; Jurczyk et al. 2019; Zhang et al. 2011).

As the majority of WSR observations used for compositing are made by national meteorological agencies operating national networks consisting of several radars with common hardware and standardised processing chains, a single processing, quality and compositing scheme is often employed. For example, the UK Met Office (UKMO) composites 18 C-band weather radars (15 operated by the UKMO) to cover the British Isles using the Radarnet system for QPE generation and compositing (Harrison et al. 2012; Golding 1998).

In contrast, the Pan-European radar composite generated by the Operational Programme for the Exchange of Weather Radar Information (OPERA) takes in data from 164 radar sites across 25 countries with significant heterogeneity in terms of weather radar hardware, signal processing techniques, and scan strategy. However, a common processing methodology is still applied at the OPERA Radar Data Center (Odyssey) to allow the generation of Europe-wide radar composite products (Saltikoff, Haase, et al. 2019; Huuskonen et al. 2014).

Another source of WSR data is the research community which, in contrast, tends to deploy either a single radar or a small, local network of 2–3 radars (Pejic et al. 2022; Neely III et al. 2018; Junyent et al. 2010). These radars often operate with shorter ranges (up to 150 km) and are deployed for local monitoring purposes or targeted atmospheric science studies. Similarly, local water resource agencies and regional agencies have also begun to deploy smaller WSRs (Hosseini et al. 2023). These systems all have the potential to act as ‘gap-fillers’ in operational networks and could provide local enhancements in QPE accuracy when combined with wider operational WSR composites if subject to effective data processing and compositing procedures.

While the process of combining radars into a multi-radar composite is an ongoing area of research, most studies still focus on combining radar observations when using a common processing methodology. As the processing used in operational networks can often be closed source, with external data requirements (including forecast model data) it is not often possible for

researchers to recreate the same level of processing, especially in real time. In this work, we demonstrate a method for merging data from a research radar deployment into a wider, pre-existing composite generated by a national meteorological service where the processing methodologies are distinct. Taking such an approach has the potential to improve resolution and coverage in a local area while retaining the wider details of a national composite for case-study analysis, while the method developed also has the potential to be used for real-time data pipelines in the future.

In Section 2 the observations used in the work are introduced, including data from the National Centre for Atmospheric Science (NCAS) X-band research radar (NXPol-1), the operational Cartesian composite produced by the UK Met Office and surface observations used for evaluation of the presented methodology. Section 3 then describes a new method of calculating a quality index for the research radar and Section 4 details how that index has been used in this work to generate a new, Cartesian QPE product. The Cartesian QPE product is then evaluated in Section 5 using visual analysis and surface observations. Section 6 discusses improvements to the methodology and how the methodology is more widely applicable beyond this example, before concluding with Section 7.

2 | Data and Background

The following section describes the data used in this study, both from research and operational weather radars and surface observations. It also summarises the research background and operational motivation for the study, which contribute to the direction taken here. This study originated as a result of the Radar Applications in Northern England (RAiN-E) project, funded by the Environment Agency (England, UK). The Environment Agency is a non-departmental government body that has the responsibility for managing the risk of flooding from main rivers, reservoirs, estuaries, and the sea in England. It co-funds the UK weather radar network and jointly operates the Flood Forecasting Centre (UK) with the Met Office. The study uses observations collected during the RAiN-E project (Section 2.1) alongside data from the UK weather radar network (Section 2.2) and from UK hydrometric telemetry networks (Section 2.3).

2.1 | NXPol-1 and the Radar Applications in Northern England (RAiN-E) Project

During the Radar Applications in Northern England campaign, NCAS's first X-band, dual polarisation weather radar, NXPol-1, was deployed from October 2018 to December 2020 at the United Utilities reservoir near Sandwith in Cumbria (54.517°N, 3.615°W).

NXPol-1 is a dual-polarisation mobile X-band research radar manufactured by Leonardo Germany GmbH (Neely III et al. 2018). During the RAiN-E campaign, NXPol-1 collected multi-elevation (ranging from 0.5° to 20°) volume scans roughly every 6 min, with a fixed set of parameters, shown in Table 1, used at each elevation. Dual-polarisation moments, including reflectivity in both horizontal and vertical

TABLE 1 | Radar parameters for the NXPOL-1 (left column) radar scans during RAIN-E and those typically used by the UKMO C-band radars when operating in long-pulse (LP, middle column) and short-pulse (SP, right column) mode.

	NXPOL-1	C-band LP	C-band SP
Frequency	9.4 GHz	5.6 GHz	5.6 GHz
Half-power beam width	1.0°	1.0°	1.0°
Azimuth spacing	1.0°	1.0°	1.0°
Gate spacing	150 m	600 m	600 m
Maximum range	150 km	255 km	112 km
Pulse width	1.0 μ s	2.0 μ s	0.5 μ s
PRF mode	Dual	Single	Dual
PRF	1000/800 Hz	300 Hz	1200/900 Hz

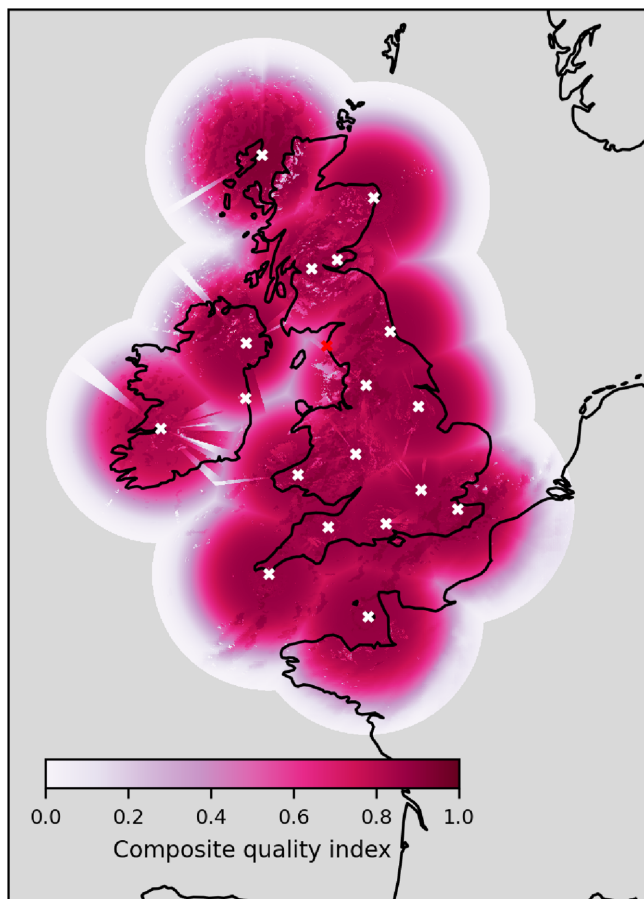


FIGURE 1 | Coverage and quality of the British Isles radar composite, 2018-11-30 00:00 UTC. The colour shading shows the quality index for the composite. Grey shading indicates the region beyond the maximum range of the radars. White crosses indicate the locations of the 18 WSRs that contribute to the final product. The red cross shows the location of the NXPOL-1 radar during RAIN-E. The black line is the coastline.

channels, differential reflectivity, cross-polar correlation coefficient, and differential phase shift were collected during the campaign (Bennett 2021).

Quantitative precipitation estimates were generated from this volume data by applying a multi-step processing chain. This included correction for pointing offset, the replacement of persistent echoes with Doppler clutter filtered data, partial beam-blockage correction, non-meteorological echo removal using fuzzy logic, and attenuation correction using the Z-PHI method prior to QPE estimation using the Marshall-Palmer relationship (as also used by the UK Met Office) (Dufton and Collier 2015; Dufton 2016; Dufton et al. 2023; Neely III et al. 2021). The open-source Py-ART library is used during this processing, and to produce several of the plots within the paper along with the cmweather package for colormaps (Helmus and Collis 2016; Sherman et al. 2024). Although previous studies have shown the value of adding dual-polarisation algorithms into such a QPE processing chain, they have been omitted here to focus on the details of quality estimation and merging.

RAiN-E was funded by the Environment Agency to assess the benefits of placing a weather radar in Cumbria, given previous severe flooding in the region and the current coverage of the UK radar network (Figure 1). The RAIN-E project demonstrated the value of observations of weather systems from NXPOL-1 as they move into Cumbria from the west over the Irish Sea, but these observations were only available to the Environment Agency outside of their existing systems. Here we look at the scientific feasibility of integrating NXPOL-1 QPE from RAIN-E with the existing UK radar composite QPE as a precursor to any deeper integration of third-party data into UK flood forecasting operations.

2.2 | The UK Met Office Precipitation Composite

The UK Met Office produces a composite surface rainfall product for the British Isles. The composite merges data from 18 weather radars (Figure 1) onto a 1 km resolution Cartesian grid using the British National Grid (BNG) reference system at 5 min intervals. The radars conduct a volume scan comprising multiple elevations with a combination of long pulse width and short pulse width used during the volume where only the long pulse scans are used for the precipitation composite (see Table 1). The processing chain has evolved significantly over the last two decades, particularly since the dual-polarisation upgrade of the Met Office network radars was completed in 2018. The Met Office single-site radar observations are processed to correct for beam blockage, clutter contamination, non-meteorological echo removal, attenuation and vertical profile of reflectivity (VPR) gradients prior to QPE generation with either the Marshall-Palmer Z-R relation (Marshall et al. 1955) or where possible an R-KDP relationship (Harrison et al. 2015, 2014; Lewis et al. 2007). After precipitation estimation, the data is also adjusted to account for orographic enhancement, and a mean-field bias adjustment is applied using data from surrounding operational rain gauges. Once each radar's QPE has been generated in the polar coordinate system, these are then gridded individually before being composited

into the final data product. A quality index is used for compositing QPE data from the individual radars. The quality index calculation incorporates the impact of signal-to-noise ratio, attenuation and the height and distance of each radar voxel from the radar (Sandford and Gaussiat 2012; Harrison et al. 2012). Figure 1 shows an example of the composite quality index score for the entire composite from 2018-11-30 at 00:00 UTC. Although the precipitation composite is published in the Centre for Environmental Data Analysis (CEDA) Archive, the quality information is not, with the data used in this study being shared directly from the Met Office archive under licence.

2.3 | Rain Gauge Networks

Rain gauge data from the Environment Agency (152 gauges), Scottish Environment Protection Agency (74 gauges), and Isle of Man flood hub (14 gauges) hydrometric networks were used in this study. In total 240 gauges in the combined network lie within 150 km range of the NXPol-1's deployment in Sandwith. The majority of these gauges were tipping-bucket type, with some weighing gauges also included within the hydrometric networks. Data from these gauges were available as accumulations (in mm) at 15-min time resolution. Rain gauge locations within the study area are shown in Figure 2.

3 | A Quality Methodology for NXPOL-1

A quality index combines a range of individual quality factors to provide a singular measure of radar data quality at any given radar voxel (three-dimensional equivalent of a pixel). The range of quality factors used varies depending on the particularities

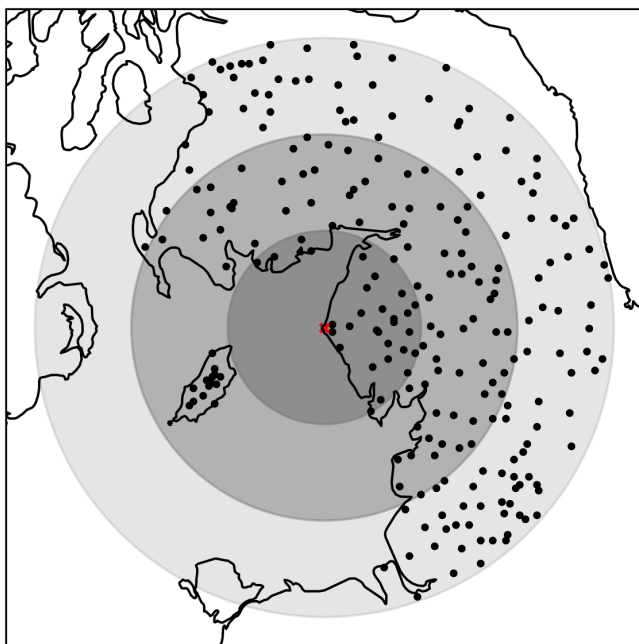


FIGURE 2 | Location of rain gauges (black dots) within the study domain. Range from the NXPol-1 radar (red cross) at 50 km increments is shown by the decreasing grayscale intensity shading. The coastline is shown as a solid black line.

of any given use case. Ośródk et al. (2014) describe in detail their approach for the Polish weather radar network which has since been incorporated into the BALTRAD system (Michelson et al. 2018) with a similar approach being taken in this work to account for the factors that most impact the overall quality of NXPol-1 QPE. In this section, these factors are described and formulated into individual quality scores before being combined using a geometric mean to determine the final quality score for each individual radar voxel.

3.1 | Individual Quality Indices

The quality factors used here logically follow from the processing methodology applied to NXPol-1 data to generate QPE, along with considerations of the geometric environmental factors which also impact radar data quality (beam altitude and distance from the radar). They are chosen to represent the main factors impacting the quality of NXPol-1 QPE, particularly the impact of attenuation (which is greater at X-band) and the impact of partial beam blockage (which is a significant factor given the location of the radar). The factors are based on previous implementations of quality indices, with minor modifications to account for differences between the NXPol-1 processing chain and those studies, particularly that by Ośródk et al. (2014).

3.1.1 | Persistent Clutter Regions (Q_M)

This factor accounts for the impact of regions of persistent clutter within the radar domain, where radar echoes are observed the majority of the time. Although NXPol-1 implements a Doppler clutter filter within the signal processor to correct radar moments for static clutter, previous work has shown this to lead to over-filtering along the zero isodop, along with remnant clutter echoes within the filtered data (Dufton 2016). To avoid over-filtering, the current QPE processing uses unfiltered data except in regions where persistent echoes have been observed. In this case, persistent echoes are based on a monthly frequency over threshold analysis (75% of total monthly echoes in excess of 10 dBZ). Given these regions are still subject to additional quality control checks, including non-meteorological echo filtering, the quality score for this stage of the processing is defined as follows:

$$Q_M = \begin{cases} 1, & \text{if unfiltered data used.} \\ 0.9, & \text{if filtered data used.} \end{cases} \quad (1)$$

This provides a small negative penalty to regions of persistent echoes, which therefore use Doppler-filtered data, even if they ultimately remain as meteorological echoes during the remaining processing. Given that Doppler-filtered data is sometimes considered to be of higher quality than non-filtered data (Ośródk et al. 2014), the penalty applied here is minor, reflecting previous experience of using data from NXPol-1.

3.1.2 | Non-Meteorological Echoes (Q_{N-MET})

Spurious echoes are often observed with WSRs. This factor accounts for the fact that when they are the dominant echoes it

is not also possible to accurately estimate precipitation. Three processing functions are used to remove non-meteorological and other erroneous echoes. Firstly, second-trip echoes are removed using a variable threshold approach based on signal quality index (SQI) and ray-to-ray variability in differential phase-shift. Then, non-meteorological echoes (insects, noise, ground clutter and unknown) are removed using a fuzzy logic classifier. Finally, isolated speckles (connected echoes smaller than 4 voxels) are removed from the remaining data (Dufton and Collier 2015; Dufton 2016). As each of these steps results in a clear binary inclusion/exclusion decision for any observed echoes the following approach is then taken for defining quality:

$$Q_{N-MET} = \begin{cases} 1, & \text{if echoes retained.} \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

As Q_{N-MET} is equal to 0 where echoes are removed, it can act as an absolute quality factor when combined with any other quality metric, provided they are combined using multiplication (a geometric mean, for example) rather than addition.

3.1.3 | Partial Beam Blockage (Q_{Pbb})

This factor accounts for uncertainties within the corrections for partial beam blockage, where the radar beam is blocked by topography or other features. During the processing partial beam blockage is corrected using a DTM-derived (SRTM 1-arc sec global data (NASA 2013)) beam-blockage model using the open-source Wradlib processing library (Heistermann et al. 2013; Bech et al. 2007, 2003). During the field campaign, it was found that the antenna pointing of NXPol-1 varied and required correction (Dufton et al. 2023). To reflect the uncertainty in both the beam blockage correction and the pointing correction where the accuracy of beam-blocked echoes is reduced, there is a corresponding reduction in quality. The quality is defined to reflect these uncertainties as follows:

$$Q_{Pbb} = \begin{cases} 1 - BBF, & \text{if } BBF \leq 0.9. \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

where BBF is the beam-blocked fraction.

3.1.4 | Attenuation (Q_A)

Attenuation is the reduction in signal power as the beam propagates through the atmosphere due to scattering and absorption. Although attenuation can be estimated and corrected for, these corrections have uncertainties which impact the quality of the WSR observations. Here attenuation is corrected using an implementation of the Z-PHI methodology to calculate path-integrated attenuation where differential phase shift is iteratively smoothed to remove backscatter differential phase (Testud et al. 2000; Hubbert and Bringi 1995; Dufton and Collier 2015; Wallbank et al. 2022). Although specific attenuation is only calculated below the melting layer with this method, the path-integrated attenuation (PIA) within that region still limits the quality of all data from above the melting layer; therefore, the

quality factor is based on PIA rather than specific attenuation. The quality metric here scales linearly with PIA between a lower limit of 1 dB and an upper limit of 10 dB as follows:

$$Q_A = \begin{cases} 1, & \text{if } PIA \leq 1. \\ 0, & \text{if } PIA \geq 10. \\ 1 - \frac{PIA - 1}{10 - 1}, & \text{otherwise.} \end{cases} \quad (4)$$

The potential for additional attenuation within the melting layer and ice-phase regions of the precipitation is not considered when using this metric, although these effects are known to be small when compared to attenuation from rain.

3.1.5 | Beam Height (Q_H)

The altitude (height above mean sea level) of the radar beam impacts the quality of surface precipitation estimates due to changes in the atmosphere between the point of observation and the surface, along with the potential for advection of observations as the observed hydrometeors fall (Sandford et al. 2017; Lack and Fox 2007; Berne et al. 2004). Beam height can also contribute to the quality of regions where hydrometeors are not observed when considering surface precipitation, as increased observation height increases the probability of the radar beam overshooting precipitation.

In this work, the height quality metric is defined as a decreasing exponential function of height (h) relative to a fixed height (H) of the form introduced by Zhang et al. (2011) for the US multi-sensor QPE system as follows:

$$Q_H = \exp\left(-\frac{h^2}{H^2}\right) \quad (5)$$

where H is set at a fixed value of 1500 m in this study and h is measured in metres.

3.1.6 | Range (Q_D)

Distance from the radar impacts the quality of radar QPE due to the broadening of the radar beam, reducing the resolution of the data and increasing the chance of non-uniform beam filling. To account for these effects, the decreasing exponential function of Zhang et al. (2011) is used to calculate the distance quality metric as follows:

$$Q_D = \exp\left(-\frac{d^2}{D^2}\right) \quad (6)$$

where d is the distance from the radar in metres, and D is a fixed reference distance set at 120,000 m in this study.

3.1.7 | Minimum Detectable Reflectivity (Q_{MDR})

The minimum detectable reflectivity (MDR) depends on the radar calibration constant, receiver sensitivity, and

range following the standard radar equation (Brangi and Chandrasekar 2001). It is a useful factor for hydrological applications as it impacts the ability to detect light rainfall, particularly at far range. Assessment of the quality of undetected regions of the radar domain must consider the possibility of precipitation existing with an equivalent reflectivity factor lower than is possible to detect with the radar system. While the minimum detectable observed reflectivity is only a function of the radar equation, both partial beam blockage and attenuation act to increase the minimum detectable true reflectivity and must be accounted for when considering this quality factor. As such, the following equation is used to calculate a quality factor for MDR in this study.

$$Q_{MDR} = \begin{cases} 1, & \text{if } MDR \leq 10. \\ 0, & \text{if } MDR \geq 15. \\ 1 - \frac{15 - MDR}{15 - 10}, & \text{otherwise.} \end{cases} \quad (7)$$

where MDR is calculated as follows:

$$MDR = C + 20 \log_{10} \left(\frac{r}{1000} \right) + \alpha_g r + PIA_r + PBB_r \quad (8)$$

This represents the observable minimum detectable reflectivity accounting for gaseous attenuation (α_g is 0.000024 dB/m), rainfall attenuation (PIA) and beam blockage (PBB). Here C is -34 dBZ for the field campaign and r is the range from the radar in metres.

3.2 | Combining Individual Indices

While each individual quality index provides information relating to a single aspect of radar performance, when combined, they gain more utility as a tool for manipulating and compositing radar data. Here, they are combined to provide information about detected echoes (i.e., where reflectivity from any source is observed) and those regions where no echoes are observed. While the latter may seem less informative, it is important to quantify prior to the applications discussed in Section 4.

3.2.1 | Quality of Detected Echoes– Q_{Detect}

With the exception of the minimum detectable reflectivity, all of the factors in Section 3.1 contribute to the overall quality of detected echoes. Here, those factors are combined as follows:

$$Q_{Detect} = Q_{N-MET} \cdot \sqrt[5]{Q_M \cdot Q_{Pbb} \cdot Q_A \cdot Q_H \cdot Q_D} \quad (9)$$

The use of a geometric mean rather than an arithmetic mean also allows both Q_A and Q_{Pbb} , when equal to zero, to act as absolute flags of quality in situations of extreme signal loss. The non-meteorological echo index is excluded from the geometric mean as it is a binary score. Its inclusion in the mean would only increase the quality score of all retained echoes, while filtered echoes would still have a score of zero.

3.2.2 | Quality of Undetected Range Voxels– $Q_{Undetect}$

The most significant factors controlling the quality of undetected regions (i.e., the likelihood that those regions do not contain precipitation) are the minimum detectable reflectivity, the occurrence of beam blockage and the height of the radar beam. The quality of undetected regions is defined here as the dot product of these indices as follows:

$$Q_{Undetect} = Q_{MDR} \cdot Q_{Pbb} \cdot Q_H \quad (10)$$

A simple dot product was chosen rather than a geometric mean to reflect the increasing potential for missed echoes when any one of these factors has reduced quality. Height is included as a factor here due to the final products of interest in this study being surface precipitation rates rather than just observations at a given location within the atmosphere. Beam blockage has been included as a factor, even though it already contributes to MDR due to initial tests of the compositing methodology. During these initial tests, beam-blocked sectors were clearly more evident in rainfall accumulations if the importance of partial beam blockage was not increased as a factor in the quality of undetected regions.

3.2.3 | Total Quality Index– Q_{IND}

Here, the total quality index, Q_{IND} , is defined as follows:

$$Q_{IND} = \begin{cases} Q_{Detect}, & \text{where reflectivity observed} \\ Q_{Undetect}, & \text{otherwise} \end{cases} \quad (11)$$

which is simply to say that the detected quality is used for all voxels which initially contained an echo, even if it was subsequently filtered and the undetected quality is used elsewhere.

4 | Application of the Total Quality Index

Following the calculation of the total quality index, it can be applied to improve the derivation of surface precipitation estimates for the NXPOL-1 radar (Sections 4.1 and 4.2) and to combine those estimates with the national precipitation composite (Section 4.3). As the quality index is calculated for each voxel of each radar volume, it adapts to changes in the conditions, including the extent of visible clutter and the amount of attenuation. This adaptation can be utilised to improve the single site QPE from NXPOL-1 by dynamically adjusting the lowest usable data for each point and the weighting of each individual voxel within a gridding routine, allowing localised changes in quality to be accounted for. This dynamic behaviour also improves the compositing of single site data into a larger composite (Sandford and Gaussiat 2012).

4.1 | Dynamic Lowest Usable Elevation

Previous studies using NXPOL-1 for QPE have defined a static map of the lowest usable radar sweep (elevation angle) across the domain, taking into account the minimum detectable reflectivity modified by partial beam blockage and regions of persistent clutter (Wallbank et al. 2022; Neely III et al. 2021). This lowest

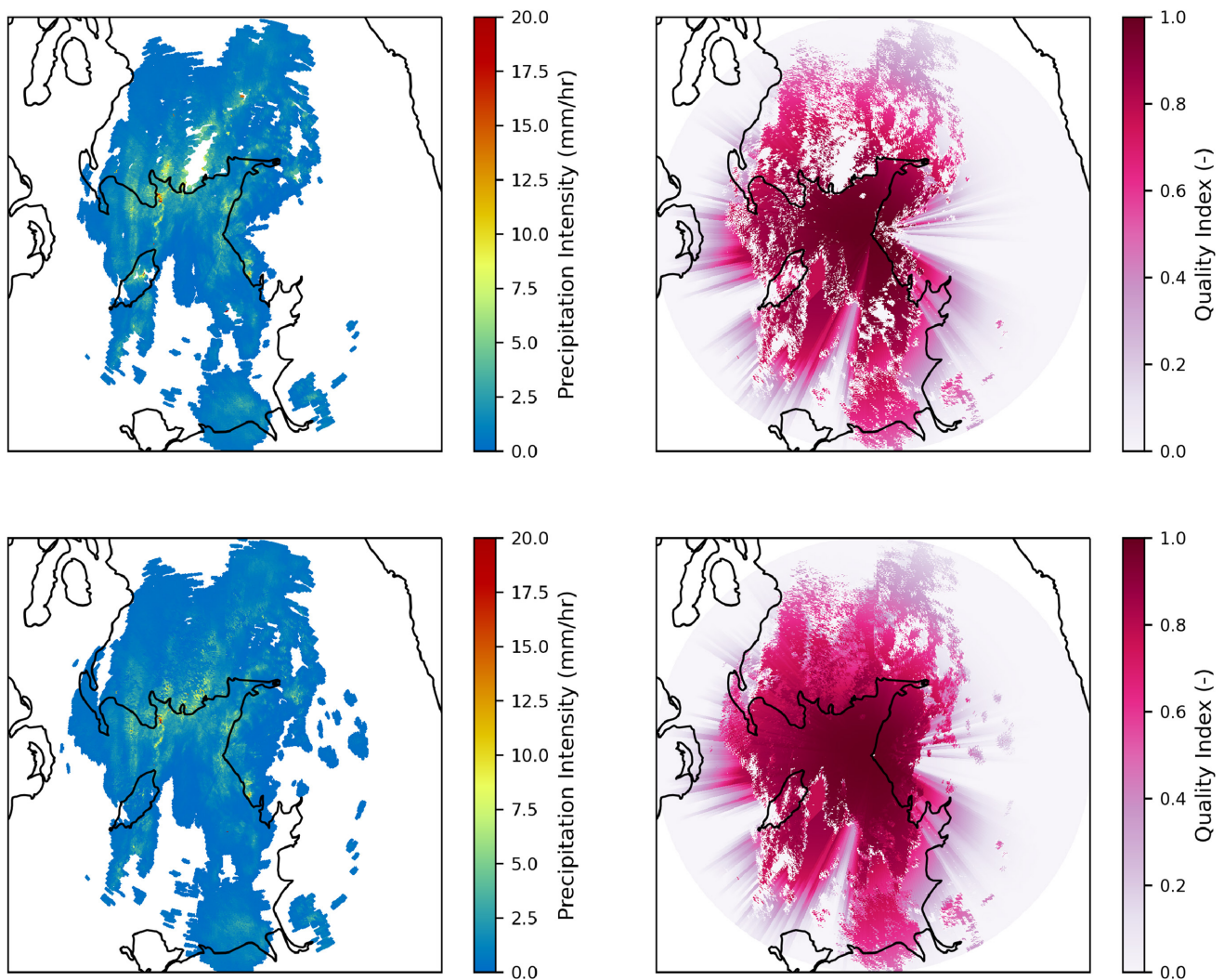


FIGURE 3 | Radar precipitation intensity (left panels) and Q_{IND} (right panels) for radar sweep at 2018-11-01 01:03GMT. The top row shows data from the lowest elevation sweep in the radar scan volume (0.5° nominal), while the bottom row shows the results of applying the dynamic flattening approach to generate a lowest usable field. Solid lines indicate the coastline. NXPol-1 is located in the centre of the domain.

usable elevation map was then fixed for the entire field campaign analysis, however it has been noted that a dynamic approach which also considers the impact of attenuation on the minimum detectable reflectivity and the possibility of variations in the extent of clutter regions due to changing conditions would be beneficial. In this work, the new quality indices defined for both the detected and undetected regions of all individual volumes can be used to dynamically flatten the radar volume into the lowest usable precipitation product. For each individual range/azimuth pair the flattening uses the elevation angle with the highest quality index, subject to the following conditions:

1. Detected precipitation at low elevations can not be replaced with undetected voxels from higher elevations, even if their quality is higher.
2. Filtered precipitation at low levels (i.e., $Q_{detect}=0$) can only be replaced with detected precipitation from higher levels (rather than undetected precipitation).
3. Undetected precipitation at the lowest elevation will not be replaced with undetected precipitation from a higher

elevation with higher quality, for that, the quality of undetected precipitation remains that of the lowest elevation even if subsequent elevations also have no detected precipitation.

This leads to a 2D product which retains low quality in terms of undetected (0) regions at the lowest elevation, considering the potential for overshooting, low-quality observations of shallow echoes, and filtering of shallow echoes. Figure 3 highlights the benefits of this approach when compared to just relying on the lowest collected elevation sweep. Comparing the lowest scan data (top panels of the figure) to the dynamic flattening approach (bottom panels), you can clearly see the increase in coverage of the rainfall intensity field. Of particular note is the large region of infilling to the north-west of the radar location where ground clutter contamination has been filtered from the lowest sweep but is then infilled within the dynamic lowest elevation precipitation product. In the example, you can see that the infilled region has a lower quality than the surrounding regions, which are from the lowest collected sweep. This is particularly important considering the next two uses of the quality index.

4.2 | Quality Weighted Gridding

Given the UK composite is produced at 1 km resolution using the British National Grid, the same approach has been taken for NXPOL-1 data. This allows a like for like comparison with the national composite while also being suitable to use as input into the hydrological models commonly used in the UK (Wallbank et al. 2022). In previous studies, the gridding has used an areal weighted arithmetic mean of all radar voxels from the flattened precipitation product which overlap with the target grid box. In this study, the gridding scheme has been modified to include quality within the weighting scheme. The computational approach has also changed to improve calculation times. Here, the polar, flattened data is interpolated using nearest neighbour interpolation to an intermediate grid with 10 times the resolution of the target grid (100 m intermediate grid for 1 km target in this case), with Q_{IND} also being interpolated to this grid. The target output for each grid cell is then the quality-weighted mean of the intermediate points within that cell (100 in this case) as follows:

$$R_i = \frac{\sum_{j=1}^N Q_j \cdot R_j}{\sum_{j=1}^N Q_j} \quad (12)$$

where i is the index of the target grid cell, j is the index of each of the intermediate grid points within that cell and N is the total number of intermediate points within the cell (100 in this case). In practice, this is achieved through convolutions of Q , R and Q with a 10×10 unit kernel.

Taking this approach has improved the gridding of NXPOL-1 precipitation estimates both in terms of processing speed and data quality. As Q_{IND} dynamically adjusts, localised filtered data is ignored within the scheme, provided the filtering does not fill the entire grid cell. At the same time, more weight is given to higher quality data (less attenuation, lower height, for example), which is particularly relevant where data has been infilled from higher elevations within a grid cell.

4.3 | Merging With a National Composite

Merging NXPOL-1 and UKMO gridded precipitation datasets together requires them to have a common spatial and temporal index. While they natively have a common spatial grid the temporal indices are different (in terms of both resolution and timestamps). To allow merging on a common temporal index, the NXPOL-1 data was linearly interpolated in time to match the finer temporal resolution of the UKMO dataset. Where NXPOL-1 data was missing for an extended period of time, it was infilled for a maximum of 5 min using the nearest (in time) available valid data before time steps were then set to missing. After this adjustment it is then necessary to consider how to merge them when both precipitation intensity and quality information are available for both products.

Recent studies have compared several techniques for compositing gridded WSR data, ranging from simply taking the highest available precipitation rate for a given cell to complex weighted combinations incorporating quality data along with other weighting factors (Barbieri et al. 2022; Jurczyk et al. 2019). Although Jurczyk et al. (2019) suggest using a weighted combination including

quality and distance from the radar is most effective (for their use case), they also show that this only marginally outperforms using only the data with the highest quality (Q_{MAX}) for a given pixel. In this case the national composite data available does not contain information about which radar the data has come from, therefore calculating distance metrics for weighting is not feasible while the composite itself is based on using Q_{MAX} as the determining factor (Sandford and Gaussiat 2012). Given the nature of this study, taking the maximum quality as the determining factor in merging research data into the national composite can act as a proof of concept that could be expanded upon in future work. As the quality indices of the two QPE products are calculated using different methodologies, using maximum quality is not guaranteed to be an optimal approach and this will be reviewed in the discussion.

Figure 4 shows a single example of merging precipitation fields using Q_{MAX} . In this case several isolated showers are observed off the coast of Cumbria. These showers (regions of high-intensity precipitation) have a greater extent when observed by NXPOL-1 (right panels) when compared to the national composite (left panels). The quality index for NXPOL-1 (lower, left panel) is strongly influenced by partial beam blockage, particularly to the east of the radar. The showers have a high quality index as they are located close to the radar, in an unblocked region that is not subject to filtering or ground clutter contamination. In contrast this region is further from the Met Office radars leading to a comparably lower quality index, particularly offshore (lower right panel). Given the proximity of these showers to NXPOL-1, and considering all other quality factors, they are merged into the combined product (central panels). Similarly, the national composite contributes the moderate intensity precipitation observed in the far north and far south of the domain (which has a slightly higher intensity in the national composite) to the combined product, as you would also expect from the quality panels. However the precipitation due east of the NXPOL-1 location is not merged into the combined product due to the impact of partial beam blockage (and higher radar elevations) on quality scoring in that region (NXPOL-1 quality is below that of the UKMO product). It is clear in this example that despite using only Q_{MAX} and not averaging the fields, significant discontinuities are not observed within the combined product, which reflects the similarities between the two independent WSR QPE products.

5 | Product Evaluation

A two-dimensional gridded surface precipitation product at 1 km/5-min resolution was produced for the entire campaign, which merged NXPOL-1 research data into the national composite (limited to the field campaign domain) using precipitation data with the maximum total quality index for each grid cell. To evaluate this product, it can be compared to the original Cartesian gridded radar products, particularly the national composite, as well as to surface observations of precipitation. The evaluation considers both whether the method outlined in the paper represents a viable methodology for merging research radar data into wider radar composites and whether, in this case, adding NXPOL-1 data into the UKMO composite improves the resulting surface precipitation estimates. In the first instance, precipitation accumulations from the three products can be compared to provide a visual indication of the effectiveness of the methodology (Section 5.1), which

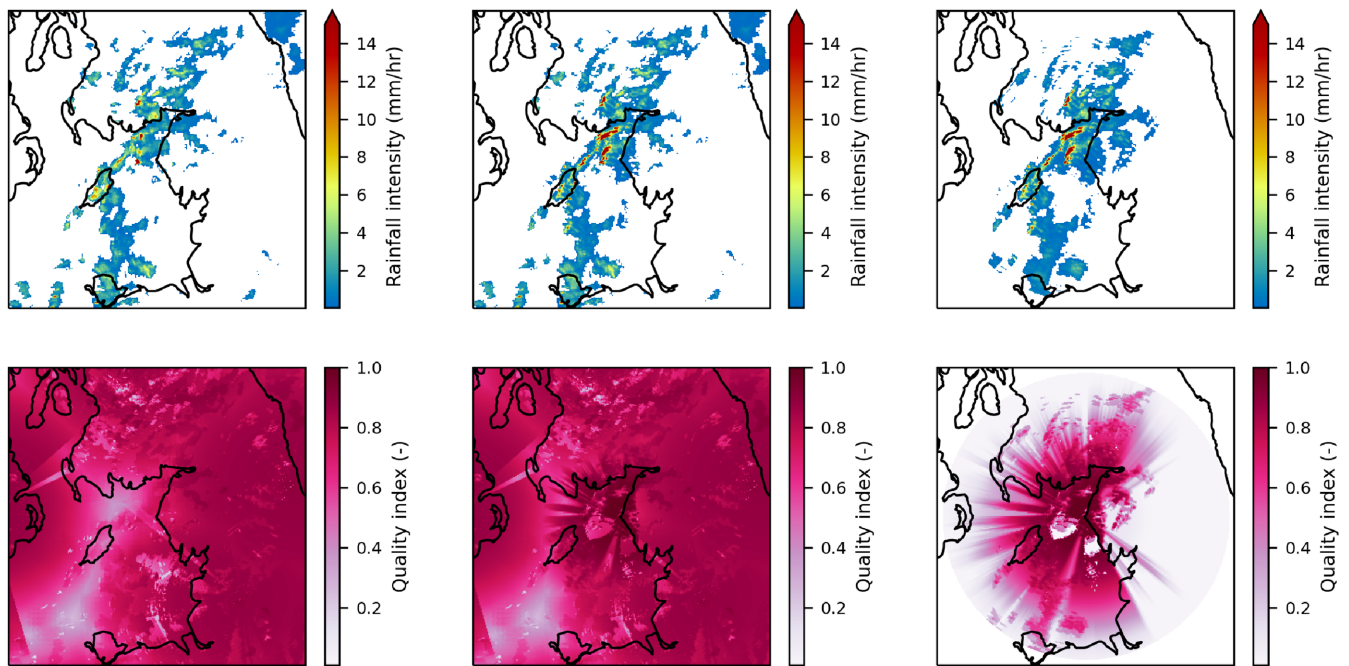


FIGURE 4 | Example of merging single time-step precipitation fields (top row), 2018-11-12 06:15 UTC and associated instantaneous quality index (bottom row). Data is from the national composite (left column), the merged precipitation product (central column) and NXPOL-1 (right column). Each panel is a 310 km box centred on the location of NXPOL-1. The solid black lines are the coastlines of the British Isles. Only precipitation in excess of 0.01 mm/h intensity is plotted.

has previously been used as an effective qualitative assessment of WSR processing performance (Harrison et al. 2014). Comparison to rain gauge observations then provides a quantitative evaluation of the products (Section 5.2) using commonly utilised statistical metrics (Dufton et al. 2023; Barbieri et al. 2022; Hyndman and Koehler 2006).

5.1 | Precipitation Accumulations

Visual comparisons of areal precipitation datasets often provide the first indication of their quality, particularly if a reference dataset of known quality is available. Although data from NXPOL-1 has previously been compared to surface observations a significant proportion (> 25%) of the domain observes over the sea where ground truth data is not available (Figure 2). Accumulating precipitation in time is a well-established technique for the evaluation of radar data as it often highlights errors due to beam blockage and clutter in the datasets, and is equally applicable over the sea as it is for observations over land (Harrison et al. 2014). Daily accumulations were produced for all the products for the entire campaign, allowing visual evaluation of the entire dataset. These accumulations are a simple time integration of the precipitation rates from the radar products over a 24 h period beginning at 00:00 without intermediate advection being applied to either dataset. Here a single example is presented in Figure 5 before a wider, qualitative assessment of the three datasets based on 697 daily accumulations from the 2 year data record is discussed.

Figure 5 shows daily precipitation accumulations over the domain for the 12th November 2018 for each of the three radar

datasets, along with the mean quality score for the day. During the day a low pressure system passed across the north of the UK, travelling west to east, with two general periods of precipitation crossing the domain, first in the early hours of the morning and then again in the evening. Considering the national composite (left panels) first, the highest rainfall accumulations occur over the high topography in Cumbria and southern Scotland, which is also the case within the NXPOL-1 accumulation. The most obvious artefacts within the WSR accumulations are coincident with where they have a lower quality score. This indicates that the quality methodology is generally characterising the pertinent features well; for example, the hole in the accumulation off the Cumbrian coast, where wind turbine interference is removed, is highlighted by a reduction in mean quality as is the beam blockage off the Cumbrian coast. The other reductions in mean quality are regions of ground clutter in upland areas to the north of the domain (in Scotland) and to the east over the Pennines. However there are noticeable discontinuities within the accumulation in the south of the domain due to beam blockage, which are not coincident with decreases in the mean quality observed. The other noticeable feature is the reduction in quality with distance highlighting the transition zones between radars, the most obvious of which runs down the Irish Sea and over the Isle of Man.

The NXPOL-1 accumulation (right panels) indicates the new quality methodology is characterising the main features within the precipitation product, with the eastern blockages due to the Cumbrian fells being indicated with a significant reduction in quality in that region. While the mean quality field indicates significant removal of sea clutter, this is not evident within the accumulation, suggesting either that the fuzzy logic classifier successfully retained precipitation within that region or that the

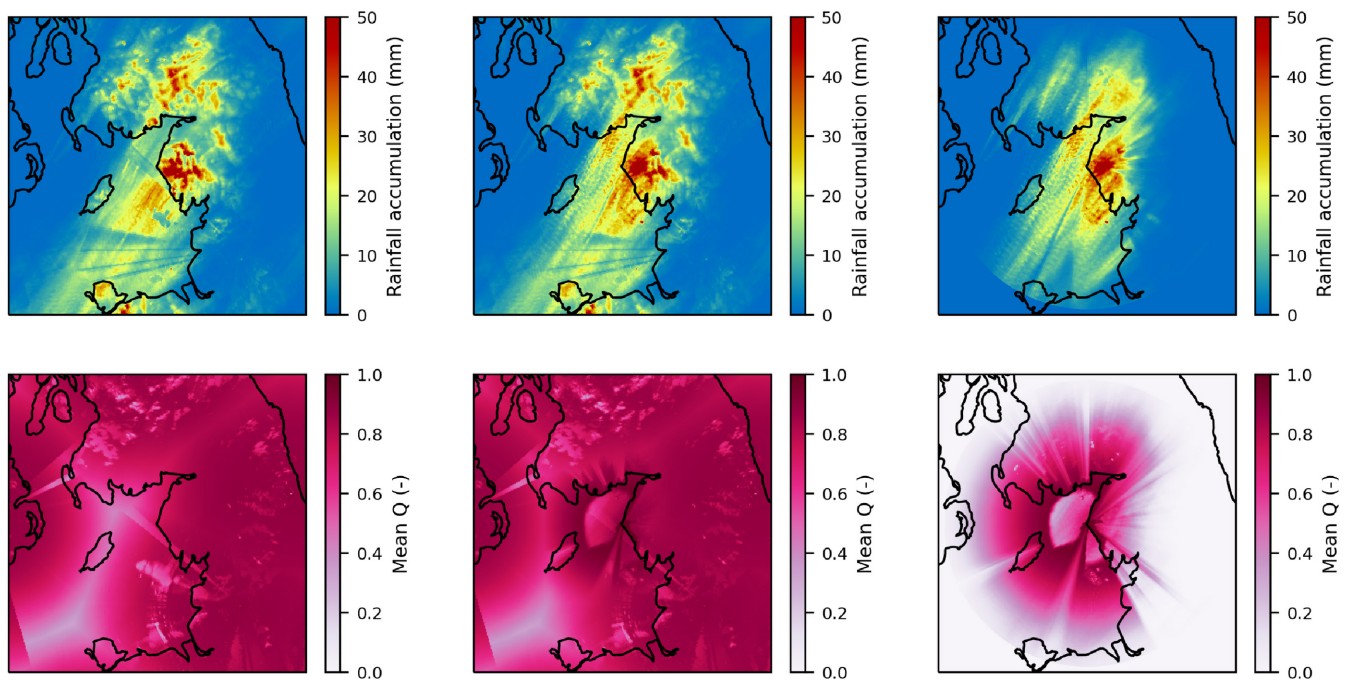


FIGURE 5 | Gridded surface precipitation accumulations for 2018-11-12 (top row) and temporal mean quality for the day (bottom row). Data is from the national composite (left column), the merged precipitation product (central column) and NXPOL-1 (right column). Precipitation accumulations are daily totals in mm, the mean quality is the arithmetic mean through time for that day. The black lines show the coastlines of the British Isles.

infill from higher elevations provided equivalent estimates (i.e., no discontinuities) even though they were from higher altitudes. Similarly to the national composite, there is one noticeable region of beam blockage within the accumulation to the south-west, which is not reflected in the quality field.

The merged accumulation maintains the characteristics of the underlying datasets, with the most noticeable difference to the national composite being in the Irish Sea, where the daily precipitation accumulations increase from around 10 to 30 mm. As for the national composite, the approach taken here has not introduced any of the obvious discontinuities that can occur within radar composites at the boundaries between radars with smooth transitions moving both northwards and eastwards. The dynamic nature of the merging processes is best highlighted by the region of sea clutter in the centre of the domain, where again, there is no visible impact on the daily precipitation accumulation and the daily mean Q has increased when compared to either of the stand-alone radar products (suggesting a blending through time rather than always using one product in the same location). Unfortunately, the blockages to the south of the domain are not in-filled during the merging process, nor is the fine blockage in the NXPOL-1 dataset, as neither is reflected in the quality scores of their respective datasets.

While a single daily example is informative, both quality scores are time variant depending on atmospheric conditions, and the whole dataset should be considered to provide a more informed assessment. In total, daily accumulations from the 697 days where NXPOL-1 was operating were considered and the following conclusions have been drawn:

- The general behaviour seen in the example continues within the whole dataset, particularly the smooth transitions

between products in the merged accumulations, the presence of blockages which are not reflected in the quality scores and the difficulty of QPE estimation in regions affected by wind turbines.

- The national composite exhibits very strong temporal variations. During some periods within the dataset, the quality “troughs” between radars almost vanish, with almost the entire domain having quality scores in excess of 0.8.
- The merged dataset varies most from the national composite in the Irish Sea, while the east of the domain almost exclusively uses data from the national composite. Averaged across the whole domain, NXPOL-1 data is used 14% of the time in the merged product.

Although accumulations provide a powerful visual assessment of precipitation products, obtaining a quantitative measure of performance across the whole domain is not feasible with this approach. To provide a quantitative measure of performance, rain gauges can provide an external reference dataset for statistical analysis. These statistical measures can provide further verification of the quality and merging methods employed here.

5.2 | Rain Gauges

To provide a quantitative assessment of the new precipitation product, each of the radar precipitation estimates has been compared to rain gauge observations across the region. For this analysis the statistical error, ϵ (Equation 13), is defined as the difference between the hourly precipitation estimates from the radar composite products, R , and those from the rain gauges, G .

$$\epsilon = R - G \quad (13)$$

Each of the following six statistical error metrics is computed for verification, where angular brackets denote time averaging in the following equations. In all cases, the metrics are computed for each individual rain gauge, with the radar QPE Cartesian grid box containing that gauge being used for comparison (as opposed to the polar data or any other spatially averaged/interpolated approach). The statistics are the root mean square error (RMSE) in millimetres of rainfall per hour (mm/h), the percentage bias (PBias), the weighted mean absolute percentage error (wMAPE), the mean absolute scaled error (MASE), the Nash-Sutcliffe efficiency (NSE) and the Pearson's correlation coefficient (r).

$$\text{RMSE} = \sqrt{\langle \epsilon^2 \rangle} \quad (14)$$

$$\text{PBias} = 100 \cdot \left\langle \frac{\epsilon}{G} \right\rangle \quad (15)$$

$$\text{wMAPE} = 100 \cdot \frac{\langle |\epsilon| \rangle}{\langle G \rangle} \quad (16)$$

$$\text{MASE} = \frac{\langle |\epsilon| \rangle}{\langle |G_i - \bar{G}| \rangle} \quad (17)$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (\epsilon_i)^2}{\sum_{i=1}^n (G_i - \bar{G})^2} \quad (18)$$

$$r = \frac{\sum_{i=1}^n (R_i - \bar{R})(G_i - \bar{G})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (G_i - \bar{G})^2}} \quad (19)$$

These metrics provide a complimentary assessment of the statistical similarity of hourly accumulations from the radar products at 1 km resolution and the surface rain gauge observations. Previous studies have shown that perfect statistical equivalence should not be expected, given the temporal and spatial sampling differences between the datasets. However, the statistical measures are an effective tool for comparing different radar QPEs. Figure 6 summarises the distribution of all the statistical metrics across the 240 rain gauges for each of the three gridded radar products. As stand-alone products, it is clear that the national composite outperforms the NXPOL-1 dataset across the domain, particularly in terms of correlation.

Given the national composite contains data from multiple radars, each of which is adjusted to account for VPR and orographic enhancement while also being corrected with a temporally varying mean-field bias adjustment through comparison to rain gauges, this disparity is unsurprising with that dataset better coping with dynamic changes compared to the less sophisticated approach taken with the research data. There is also a clear difference in behaviour when looking at the percentage bias, with the majority of the gauges being underestimated by NXPOL-1 and overestimated by the national composite; this behaviour also impacts the distribution of the results for the wMAPE and the MASE. Analysing the results spatially indicates that NXPOL-1's performance is strongly correlated with the distance and height of the radar observations. This is less so for the national composite as it already contains data from multiple radars, thereby reducing the distance to the closest radar for most gauges, and due to its VPR correction improving

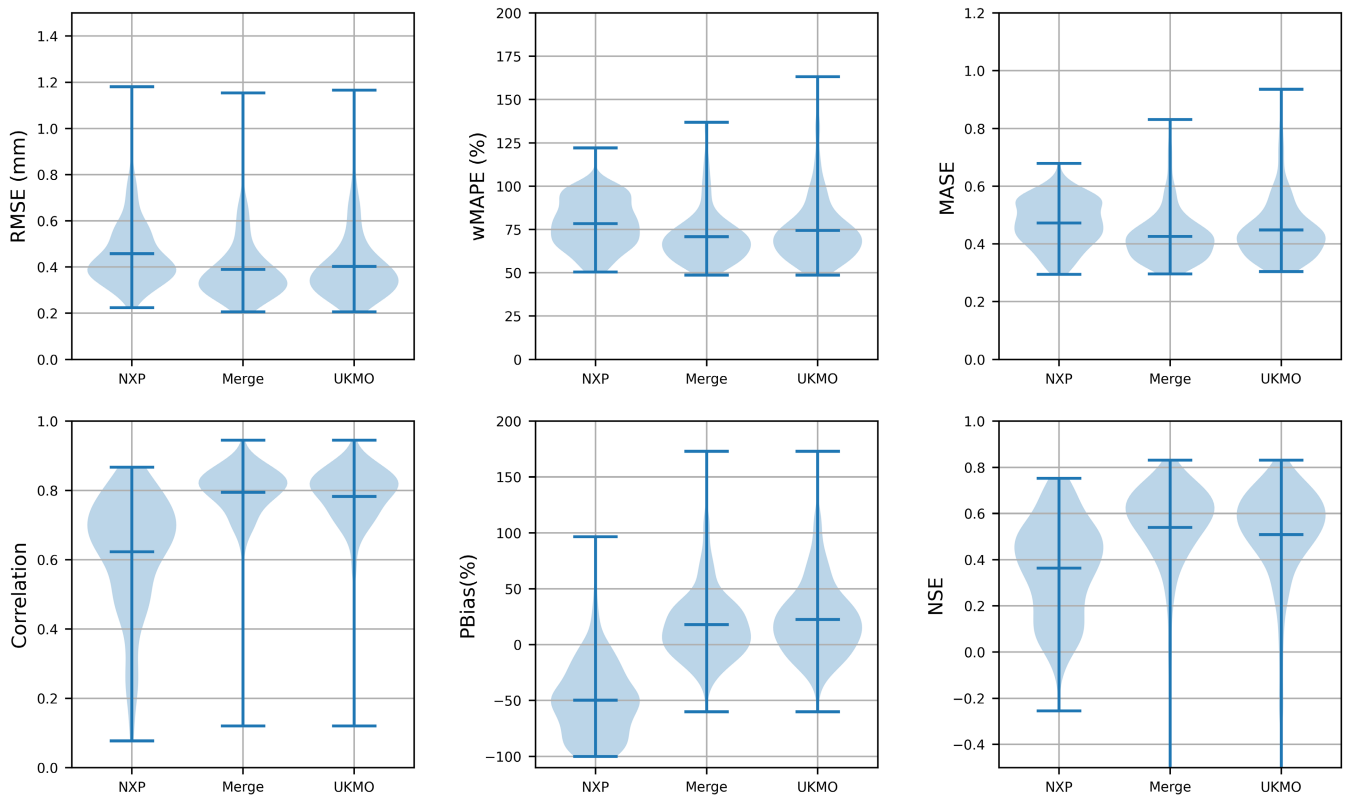


FIGURE 6 | Violin plots showing the 6 statistical metrics when comparing radar products to radar observations. The plots show results for RMSE (top-left), wMAPE (top-centre), MASE (top-right), correlation (bottom-left), percentage bias (bottom-centre) and the Nash-Sutcliffe efficiency (bottom-right). Each plot contains the comparison for NXPOL-1 (left violin), the merged product (centre violin) and the national composite (right violin).

performance where it is only possible to observe precipitation at relatively high altitudes.

Despite these marked differences, there is still an improvement in each of the metrics when the datasets are merged using the quality scores. This is seen as a small shift in the mean of each distribution and a shift in the shape of the distribution towards improved statistical representation. These shifts indicate that the quality merging approach developed here successfully makes a dynamic selection of precipitation data that improves the overall statistical performance of the final dataset, despite the obvious performance differences between NXPOL-1 and the national composite.

In the context of the RAIN-E campaign, the significant objective of this study was to develop a methodology that would allow integration of the research data while not degrading the performance of the newly merged product when compared to the national composite. To evaluate this, the performance of the new product is directly compared to that of the original national composite in Figure 7. The figure shows that within a limited geographic area to the north and west of the radar (enclosed

within the higher topography beyond) performance improves when QPEs from NXPOL-1 are merged into the national composite, while outside that region, the performance largely remains unchanged. This is seen through an increase in both the correlation of hourly radar accumulations with rain gauge accumulations and an increase in the Nash-Sutcliffe efficiency along the coastlines of both Cumbria and south-west Scotland and across the Isle of Man. Both the RMSE and MASE decrease in these regions too, and again show no increases (decreasing performance) in the regions to the north and east of the domain.

6 | Discussion

The quality methodology and applications detailed in Sections 3 and 4 provide a framework that has been shown to successfully merge research radar precipitation estimates into a national composite despite the differences in radar hardware, scan strategy and processing methodologies. This process leads to an improvement in the statistical performance of the resulting Cartesian QPE product when rain gauges are used as a surface benchmark.

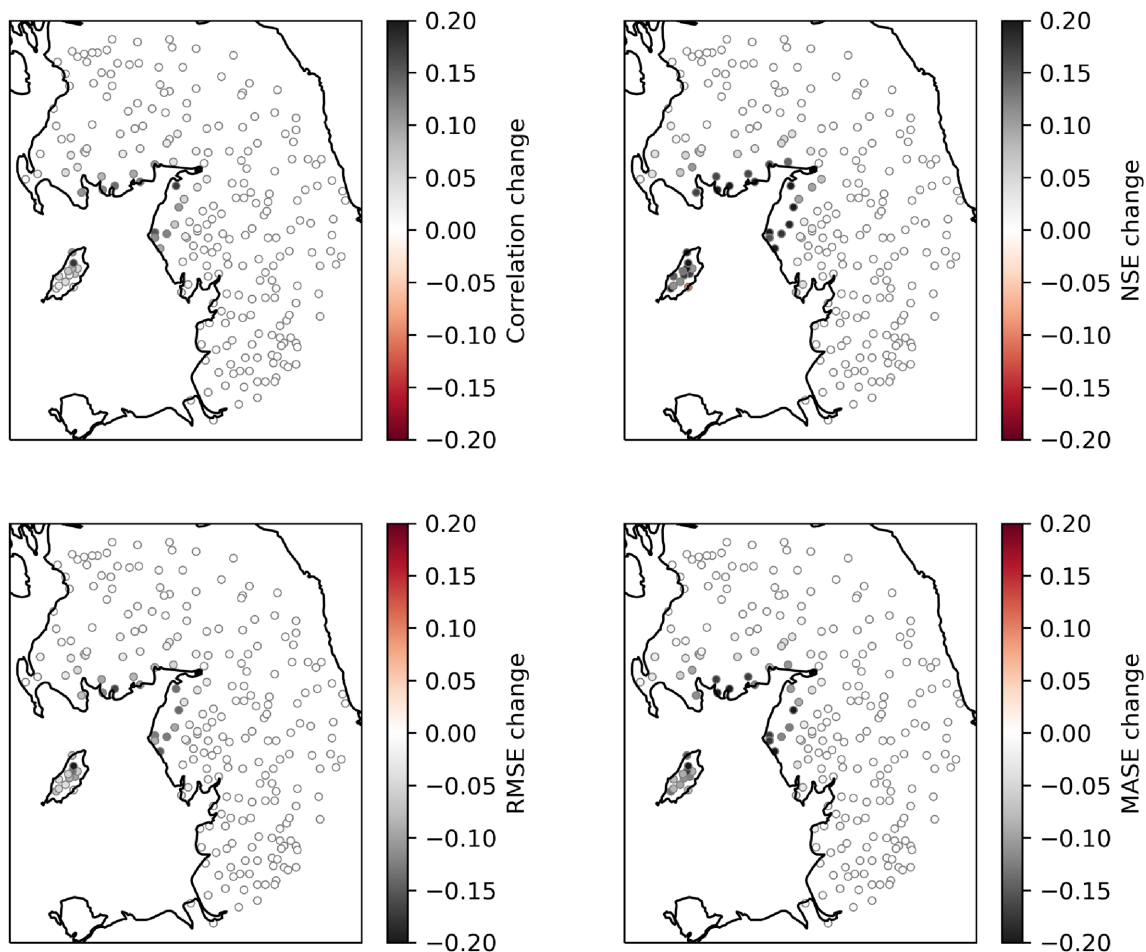


FIGURE 7 | Spatial distribution of changes in 4 statistical metrics between the new merged precipitation product and the national composite. The change in correlation is shown on the top left plot, the change in Nash-Sutcliffe efficiency is shown on the top right plot, the change in RMSE is shown on the bottom-left plot, and the change in MASE is shown on the bottom-right plot. Each gauge is shown by an open circle, which is coloured to show the change in the statistical measure (new merged minus national composite). Positive changes represent an improved statistical representation between radar and gauge observations on the top row, while negative changes represent an improved statistical representation on the bottom row, with the colour bars reflecting this difference such that black is always an improvement and red a worsening in statistical performance.

As seen in Figure 7, the methodology developed would allow integration of NXPOL-1 research data while not degrading the performance of the newly merged product when compared to the national composite; however, there is a question of whether modification of the quality indices to more closely align with the national method would improve the results further. Section 6.1 outlines modifications to the methodology following discussions with the UK Met Office, which create closer alignment with their own quality methodology while retaining the overall structure of the intuitive approach outlined in Section 3. This modified quality metric is then evaluated using the same measures as in Section 5.

Section 6.2 then discusses the wider applicability of the methods developed along with potential future research questions which naturally follow from this proof of concept.

6.1 | Closer Alignment With the UKMO Quality Index

One of the primary differences in the quality metrics for NXPOL-1 and the UK Met Office composite as observed in Section 5.1 is the temporal variation of quality seen within the UK Met Office precipitation composite. This impacts the extent to which NXPOL-1 data is merged into the new product during periods of elevated UK Met Office precipitation composite quality. Sandford and Gaussiat (2012) discuss the current quality approach at the Met Office in terms of using physically realistic error estimates to define Q , however they leave open the question of how to deal with height and range as additional factors impacting quality. Since then, internal work at the UK Met Office has led to both of these factors being incorporated, with their quality factor varying as a function of the melting layer height (T. Darlington and S. Best, personal communication, 2024). In addition, the UK Met Office take a more conservative approach to beam blockage correction than used in the NXPOL-1 QPE, and their physically based approach to quality leads to an inherent maximum quality value of 0.93 being observed at any point during the dataset. To account for these differences, a new quality methodology was devised for NXPOL-1, which included the following changes.

Firstly Q_H and Q_D have been changed to both be a function of the melting layer height. The functional forms have been chosen to match the behaviour of the Met Office quality calculation below the melting layer while maintaining both height and distance as individual factors. The functional forms chosen to ensure that quality for NXPOL-1 decreases rapidly above the melting layer in contrast to the Met Office quality, reflecting the fact that NXPOL-1 has no VPR correction while the UK Met Office precipitation composite does. Equation (5) has been replaced with the following new formulation for Q_H :

$$Q_H = \frac{1}{1 + \exp\left(-\frac{\Delta H}{0.1}\right)} \quad (20)$$

where ΔH is the difference between the bright band top and the voxel altitude measured in kilometres (i.e., the height of the bright-band minus the radar altitude). In this case, the bright-band top is taken as a single value for the entire radar domain as used

within the NXPOL-1 attenuation processing (Dufton 2016) while a lower limit of 1500m is imposed as is also the case in the Met Office's own quality scheme (Darlington and Best). In addition, Equation (6) has been replaced with the following for calculating Q_D :

$$Q_D = \frac{\exp(-d^2)}{d_0^2} \quad (21)$$

where d_0 is a scaling factor dependent on the height of the melting layer, calculated as follows:

$$d_0 = 100000 + (50 \cdot \text{MLH}) \quad (22)$$

where MLH is the height of the melting layer in metres.

Changing both Q_H and Q_D to factor in the melting layer height ensures that the quality index for NXPOL-1 varies temporally in a manner consistent with the temporal fluctuations seen in the Met Office quality index. Another change made to the quality calculation was to reduce the acceptable level of partial beam blockage from 0.9 to 0.75, in essence reformulating Equation (3) to the following:

$$Q_{\text{Pbb}} = \begin{cases} 1 - \text{BBF}, & \text{if } \text{BBF} \leq 0.75. \\ 0, & \text{otherwise.} \end{cases} \quad (23)$$

Finally, to account for the fact that the Met Office quality index never exceeds 0.93, the final computed NXPOL-1 quality index has been scaled by a fixed adjustment factor (of 0.93) when compared to the UK Met Office quality for merging to the combined composite.

An example of the new quality index for NXPOL-1 is shown in Figure 8 for the same time as shown originally in Figure 4. In contrast to the original quality method the new technique has a much narrower transition zone from high to low quality in undetected regions (to the south and west in this example), thus increasing the likely zone of influence of the NXPOL-1 radar when incorporating it into the national composite. This effect will increase when the melting layer is higher given the new Q_H is now calculated as a function of melting layer height. The same change will also impact detected regions, especially in the absence of partial beam blockage and attenuation. In the case of the NXPOL-1 stand-alone product the changes will lead to an increase in infilling from higher elevation angles when those observations are below the melting layer as the new Q_H has a very shallow response to altitude changes below the melting layer.

To understand the effectiveness of these changes, the entire dataset has then been reprocessed to allow these changes to also impact the flattening and gridding process when generating Cartesian NXPOL-1 QPE, before being re-evaluated as done in Section 5 for the original method. After these changes, NXPOL-1 data is used 21% of the time in the merged product, when averaged across the whole domain, in contrast to 14% of the time in the original version.

The results of the statistical comparison are shown in Figure 9, which is directly analogous to Figure 6. The same trends as

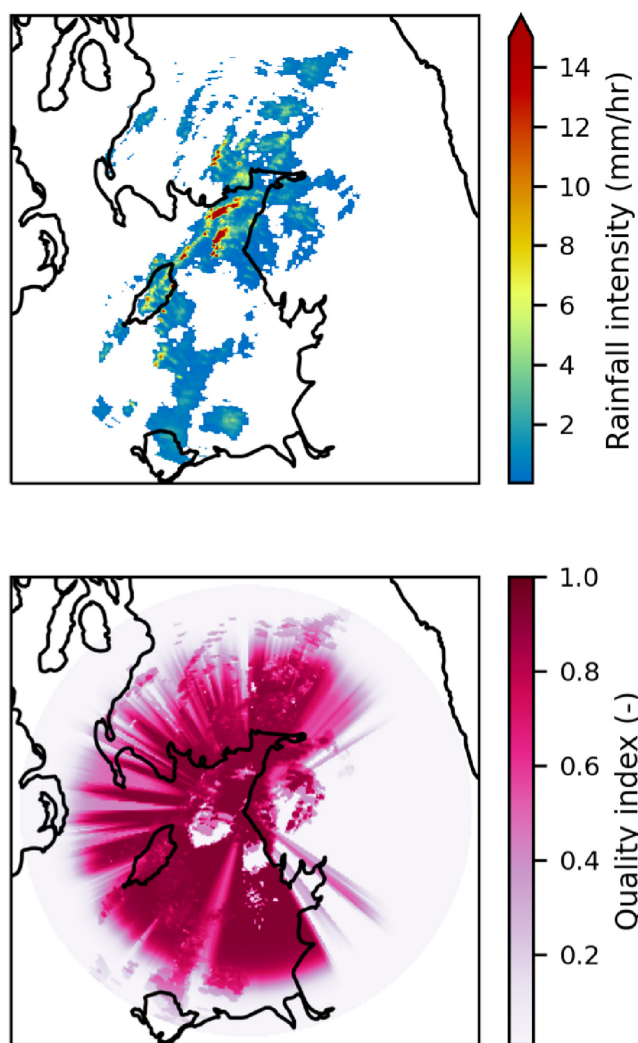


FIGURE 8 | Example of the new quality index (bottom row) and associated precipitation field (top row) for 2018-11-12 06:15 UTC, as previously shown for the original method in Figure 4. Both panels are 310 km squares centred on the NXPOL-1 radar.

previously noted for the original method are visible for the new approach, with the new Cartesian product being a statistical improvement on either of the original datasets as seen when using the original method.

The new distributions shown in Figure 9 are nearly identical to those in Figure 6, with the means of the distributions differing by less than 3% when comparing the two NXPOL-1 products. When considering changes to the statistical metrics for individual rain gauges there is a near even split in the number showing improvement and degradation for both the NXPOL-1 dataset and the merged dataset with the majority differing only slightly. However, there are some very subtle differences between the original and the new results when analysed closely. Firstly, the highest performing rain gauges in terms of correlation coefficient and Nash-Sutcliffe efficiency all show a reduction in the statistical accuracy of the gridded NXPOL-1 QPE when using the new, more aligned quality scoring. However, for the 24 rain gauges that have a correlation in excess of 0.8 using the first methodology the average decrease is still only 0.02 (with the maximum decrease being less than

0.04). While a noticeable effect in the data analysis when considering the overall changes it still remains a small variation in accuracy between the two methods. In contrast to this effect, the final merged product now has a subtle, positive (in the sense of improving goodness of metric) shift in the statistical distributions, indicating that while the targeted approach has degraded the original underlying dataset, it has improved the final dual source merged product. Again these changes are all small, with only the average of percentage bias varying by more than 3%. While this highlights the importance of having an overall objective in mind when considering the development of quality indices, the statistical comparisons show the two methods to be similar in performance for the majority of the reference region.

6.2 | Wider Applications and Future Research

The intuitive quality methodology developed in Section 3 has been shown to work effectively both as a tool for manipulating NXPOL-1 QPE into a Cartesian grid and when used to actively compare to the Met Office quality index for dual source QPE merging. The comparison with rain gauges indicates that the quality index developed successfully characterises the radar performance issues so that they can effectively integrate with the existing methodology of the UK Met Office. This indicates that the seven factors outlined characterise the biggest sources of error for NXPOL-1 such that poor-performing data is not inadvertently merged when creating the dual-source QPE product. To paraphrase Ośródk and Szturc (2015) this is only possible because the NXPOL-1 processing is complete and efficient enough to produce QPE of comparable accuracy to the national composite with which it is being merged, even though the processing chains are not identical in scope or method.

Following this, it is clear that the method developed can not be viewed as directly transferable to another radar system, but the approach taken can be considered as a guide. Given that each system and processing chain is unique, what is important is that the quality scheme developed effectively characterises each stage within the processing while acknowledging any deficiencies and/or omissions therein. The results of this study indicate that if that is the case, then it is possible to combine radar QPE from different sources effectively, even when the quality calculations are not directly aligned (due to processing differences or knowledge gaps, for example). This has significant potential to allow research radar data to be incorporated into wider, existing composites allowing their use to enhance now-casting or flood modelling across a wider area, provided that the existing composite datasets are published with a quality index. We encourage that to be the case wherever possible. At the same time, it is also preferable that the precise methodology of both QPE generation and quality determination are also published to allow further research in this area. Doing so would allow third-party data to be incorporated in this way, even when it is not possible to fully replicate the processing applied to large-scale (national/international) composites.

When undertaking any such studies, it is worth considering the final objective of the study. As we have seen here, the best quality methodology for one outcome, deriving QPE for NXPOL-1, may not necessarily be the best for another outcome, merging that

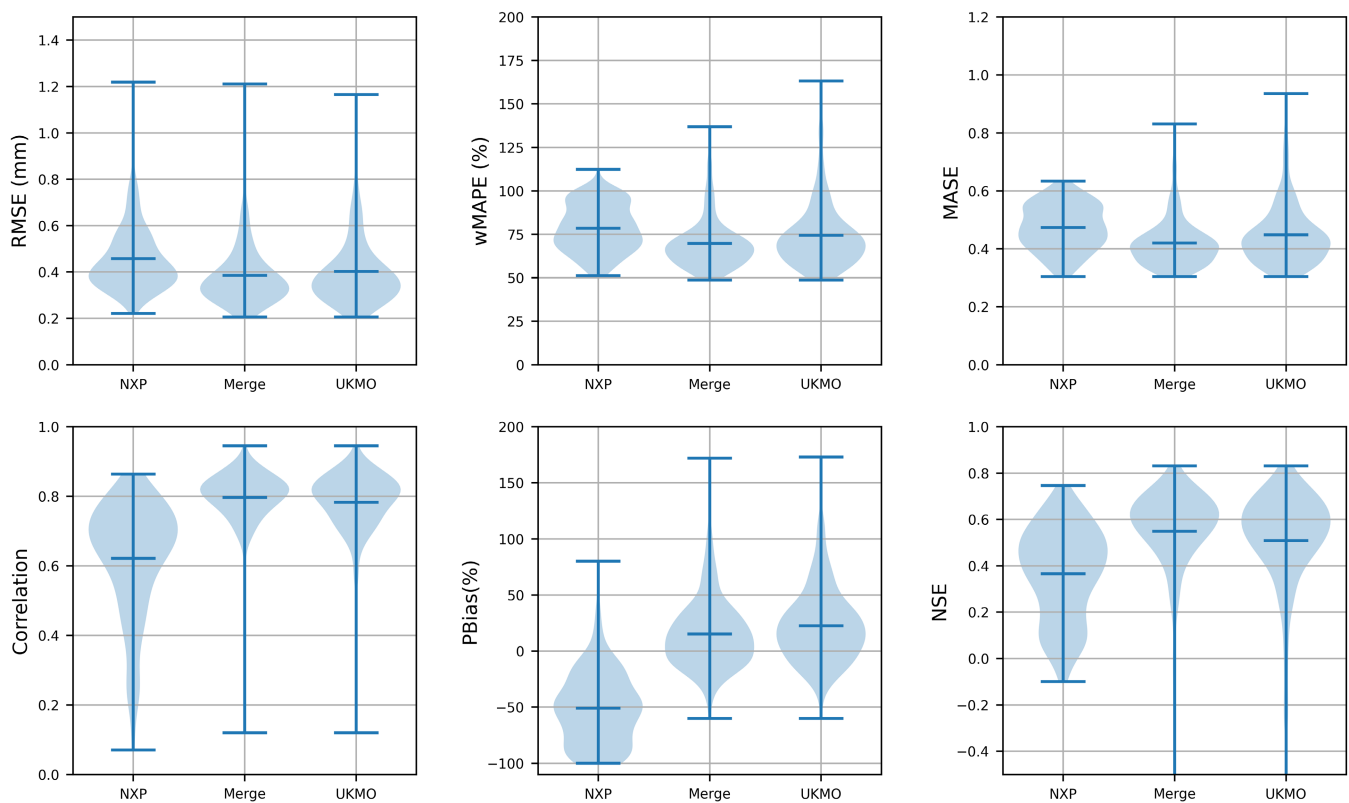


FIGURE 9 | Violin plots showing the six statistical metrics when comparing radar products to radar observations for the 2nd version of the quality methodology. As for Figure 6, each plot contains the comparison for NXPol-1 (left violin), the merged product (centre violin) and the national composite (right violin) and the plots are RMSE (top-left), wMAPE (top-centre), MASE (top-right), correlation (bottom-left), percentage bias (bottom-centre) and the Nash-Sutcliffe efficiency (bottom-right). The results for the national composite are identical on these plots to Figure 6.

QPE with existing data in this case. Targeting the approach to the specific application as done in Section 6.1 may lead to better final results should the available information allow, while it is even possible that taking different approaches for the X-band processing and then the merging could be more optimal. It is encouraging that even the more broadly applicable approach taken in Section 3 still produces effective results (almost identical to the targeted approach), indicating its viability for combining QPE datasets even in situations where imperfect knowledge is available.

One potential unknown during the study was how effective using maximum quality alone as a decision factor in the merging would be given the different quality calculations being used for the Met Office and NCAS radar precipitation estimates. In this case it has proven effective due to the comparable underlying accuracy of the two datasets (in optimal conditions) and the very similar underlying principles being applied to each quality calculation. It is likely that maximum quality would be effective as a determining factor in other use cases provided these two conditions are also considered, though it may not be the optimal choice.

While taking the intuition-led approach of Ośródko et al. (2014) and adjusting it to fit the specifics of the NXPol-1 radar has provided a viable proof of concept, another possible avenue of future work in this area is the introduction of machine learning to optimise the quality factor parameters more effectively. In this work, the quality methodology was developed entirely independently of the statistical performance analysis; however, given the length of the dataset available

along with the number and distribution of surface observations, there is potential for applying optimisation schemes for each of these factors. This is beyond the scope of this initial study but could provide interesting future insight into the efficacy of optimising quality calculation as opposed to taking the intuitive approach.

7 | Conclusions

This study demonstrates how weather surveillance radar QPE products produced by different organisations can be effectively merged using a quality-index approach. The quality method developed for NXPol-1 QPE clearly characterises the main factors influencing the accuracy of the dataset. Given the UK Met Office also takes a compatible, multi-factor, best-endavours approach to quality scoring, it was possible to successfully merge the research radar data into the national composite and improve its performance as evaluated by both surface observations and visual analysis of the resulting data product. The method presented is easily adaptable to other WSR QPE datasets, even in situations where a single operator may not generate them. While the fact remains that it is almost impossible to provide a fully comprehensive and objective quality-index estimation methodology (Jurczyk et al. 2019), taking the intuitive, best-endavours approach shown here provides valuable information for dynamic merging over and above just using factors relating to the geographic position (distance and height) of the WSR radars. Although the methodology clearly works in this case, merely

defining quality indices does not guarantee this outcome; it still requires that the myriad factors impacting radar QPE accuracy have been accounted for as effectively as possible during the data processing such that the QPE estimates from the data products to be merged have comparable accuracy levels during ideal conditions, as eloquently expressed by Ośródk and Szturc (2015). The results shown here indicate that it is these underlying principles that are most important for effective merging rather than the precise mathematical formulation of the quality metrics themselves.

Improving local and regional precipitation composites by merging in ad-hoc WSR observations from different sources has the potential to improve local flood forecasting and incident management efforts. In particular, regions where existing WSR networks have reduced coverage would particularly benefit from this approach.

Author Contributions

David R. L. Dufton: writing – original draft, writing – review and editing, data curation, methodology, software, investigation, formal analysis, visualization, validation, conceptualization, project administration, funding acquisition. **Tamora D. James:** software, writing – review and editing, data curation, resources, investigation. **Mark Whitting:** writing – review and editing, project administration, conceptualization, supervision, resources. **Ryan R. Neely III:** writing – review and editing, conceptualization, funding acquisition, supervision.

Acknowledgements

The authors would like to acknowledge all those who assisted with the radar deployment at Sandwith, particularly David Snaith from the Environment Agency and Lindsay Bennett at NCAS. The authors would also like to thank their collaborators at UKCEH (both in this project and in the Hydro-JULES project) and colleagues at both NCAS and the Environment Agency for the many valuable discussions on the subject of this work. In addition we would like to thank Tim Darlington and the radar research and development team at the UK Met Office for their discussions during the work and for providing access to the national radar rainfall and quality metrics used in the study. Finally the authors would like to acknowledge the Atmospheric Measurement and Observation Facility (AMOF) a Natural Environment Research Council (UKRI-NERC) funded facility (NE/Y005376/1), for providing the NXPol-1 data used in this study, and SEPA and the Isle of Man flood hub for providing additional rain gauge data used in this study.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request. Both the Met Office radar composite rainfall and the NXPol-1 polar data from RAiNE are available from the CEDA archive (<https://catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350/> and <https://doi.org/10.5285/bd860cadea1344ee83af32887a135d67>, respectively).

References

Barbieri, S., S. Di Fabio, R. Lidori, F. L. Rossi, F. S. Marzano, and E. Picciotti. 2022. "Mosaicking Weather Radar Retrievals From an Operational Heterogeneous Network at C and X Band for Precipitation Monitoring in Italian Central Apennines." *Remote Sensing* 14, no. 2: 248. <https://doi.org/10.3390/rs14020248>.

Bech, J., B. Codina, J. Lorente, and D. Bebbington. 2003. "The Sensitivity of Single Polarization Weather Radar Beam Blockage Correction to Variability in the Vertical Refractivity Gradient." *Journal of Atmospheric and Oceanic Technology* 20, no. 6: 845–855. [https://doi.org/10.1175/1520-0426\(2003\)020<0845:TSOSPW>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)020<0845:TSOSPW>2.0.CO;2).

Bech, J., U. Gjertsen, and G. Haase. 2007. "Modelling Weather Radar Beam Propagation and Topographical Blockage at Northern High Latitudes." *Quarterly Journal of the Royal Meteorological Society* 133, no. 626: 1191–1204. <https://doi.org/10.1002/qj.98>.

Bennett, L. 2021. "RAIN-E: NCAS Mobile X-Band Radar Scan Data From Sandwith, Near Whitehaven in Cumbria, UK, Version 1." Medium: application/xml. <https://catalogue.ceda.ac.uk/uuid/bd860cadea1344ee83af32887a135d67>.

Berne, A., G. Delrieu, H. Andrieu, and J. D. Creutin. 2004. "Influence of the Vertical Profile of Reflectivity on Radar-Estimated Rain Rates at Short Time Steps." *Journal of Hydrometeorology* 5, no. 2: 296–310. [https://doi.org/10.1175/1525-7541\(2004\)005<0296:IOTVPO>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0296:IOTVPO>2.0.CO;2).

Bringi, V. N., and V. Chandrasekar. 2001. *Polarimetric Doppler Weather Radar: Principles and Applications*. Cambridge University Press Google-Books-ID: 4r7e6assEHwC.

Cole, S. J., and R. J. Moore. 2009. "Distributed Hydrological Modelling Using Weather Radar in Gauged and Ungauged Basins." *Advances in Water Resources* 32, no. 7: 1107–1120. <https://doi.org/10.1016/j.advwatres.2009.01.006>.

Dufton, D., L. Bennett, J. R. Wallbank, and R. R. Neely. 2023. "Correcting for Mobile X-Band Weather Radar Tilt Using Solar Interference." *Remote Sensing* 15, no. 24: 5637. <https://doi.org/10.3390/rs15245637>.

Dufton, D. R. L. 2016. "Quantifying Uncertainty in Radar Rainfall Estimates Using an X-Band Dual Polarisation Weather Radar." phd. University of Leeds. <https://etheses.whiterose.ac.uk/15486/>.

Dufton, D. R. L., and C. G. Collier. 2015. "Fuzzy Logic Filtering of Radar Reflectivity to Remove Non-Meteorological Echoes Using Dual Polarization Radar Moments." *Atmospheric Measurement Techniques* 8, no. 10: 3985–4000. <https://doi.org/10.5194/amt-8-3985-2015>.

Fabry, F. 2018. *Radar Meteorology: Principles and Practice*. Cambridge University Press Google-Books-ID: 0aRwCQAQBAJ.

Germann, U., M. Boschetti, L. Clementi, et al. 2022. "Weather Radar in Complex Orography." *Remote Sensing* 14, no. 3: 503. <https://doi.org/10.3390/rs14030503>.

Golding, B. W. 1998. "Nimrod: A System for Generating Automated Very Short Range Forecasts." *Meteorological Applications* 5, no. 1: 1–16. <https://doi.org/10.1017/S1350482798000577>.

Harrison, D., S. Georgiou, N. Gaussiat, and A. Curtis. 2014. "Long-Term Diagnostics of Precipitation Estimates and the Development of Radar Hardware Monitoring Within a Radar Product Data Quality Management System." *Hydrological Sciences Journal* 59, no. 7: 1277–1292. <https://doi.org/10.1080/02626667.2013.841316>.

Harrison, D., K. Norman, T. Darlington, et al. 2015. "The Evolution of the Met Office Radar Data Quality Control and Product Generation System: RADARNET." In: *Proceedings of 37th Conference on Radar Meteorology*. Norman, Oklahoma, USA: American Meteorological Society. 14B.2. <https://ams.confex.com/ams/37RADAR/webprogram/Paper275684.html>.

Harrison, D. L., K. Norman, C. Pierce, and N. Gaussiat. 2012. "Radar Products for Hydrological Applications in the UK." *Proceedings of the Institution of Civil Engineers: Water Management* 165, no. 2: 89–103. <https://doi.org/10.1680/wama.2012.165.2.89>.

He, X., F. Vejen, S. Stisen, T. O. Sonnenborg, and K. H. Jensen. 2011. "An Operational Weather Radar-Based Quantitative Precipitation Estimation and Its Application in Catchment Water Resources Modeling." *Vadose Zone Journal* 10, no. 1: 8–24. <https://doi.org/10.2136/vzj2010.0034>.

- Heistermann, M., S. Jacobi, and T. Pfaff. 2013. "Technical Note: An Open Source Library for Processing Weather Radar Data (Wradlib)." *Hydrology and Earth System Sciences* 17, no. 2: 863–871. <https://doi.org/10.5194/hess-17-863-2013>.
- Helmus, J., and S. Collis. 2016. "The Python ARM Radar Toolkit (Py-ART), a Library for Working With Weather Radar Data in the Python Programming Language." *Journal of Open Research Software* 4, no. 1: e25. <https://doi.org/10.5334/jors.119>.
- Hosseini, S. H., H. Hashemi, R. Larsson, and R. Berndtsson. 2023. "Merging Dual-Polarization X-Band Radar Network Intelligence for Improved Microscale Observation of Summer Rainfall in South Sweden." *Journal of Hydrology* 617: 129090. <https://doi.org/10.1016/j.jhydrol.2023.129090>.
- Hubbert, J., and V. N. Bringi. 1995. "An Iterative Filtering Technique for the Analysis of Copolar Differential Phase and Dual-Frequency Radar Measurements." *Journal of Atmospheric and Oceanic Technology* 12, no. 3: 643–648. [https://doi.org/10.1175/1520-0426\(1995\)012<0643:AIFTF T>2.0.CO;2](https://doi.org/10.1175/1520-0426(1995)012<0643:AIFTF T>2.0.CO;2).
- Huuskonen, A., E. Saltikoff, and I. Holleman. 2014. "The Operational Weather Radar Network in Europe." *Bulletin of the American Meteorological Society* 95, no. 6: 897–907. <https://doi.org/10.1175/BAMS-D-12-00216.1>.
- Hyndman, R. J., and A. B. Koehler. 2006. "Another Look at Measures of Forecast Accuracy." *International Journal of Forecasting* 22, no. 4: 679–688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>.
- Junyent, F., V. Chandrasekar, D. McLaughlin, E. Insanici, and N. Bharadwaj. 2010. "The CASA Integrated Project 1 Networked Radar System." *Journal of Atmospheric and Oceanic Technology* 27: 61–78. <https://doi.org/10.1175/2009JTECHA1296.1>.
- Jurczyk, A., J. Szturc, and K. Ośródk. 2019. "Quality-Based Compositing of Weather Radar Derived Precipitation." *Meteorological Applications* 27, no. 1: e1812. <https://doi.org/10.1002/met.1812>.
- Kox, T., C. Lüder, and L. Gerhold. 2018. "Anticipation and Response: Emergency Services in Severe Weather Situations in Germany." *International Journal of Disaster Risk Science* 9, no. 1: 116–128. <https://doi.org/10.1007/s13753-018-0163-z>.
- Lack, S. A., and N. I. Fox. 2007. "An Examination of the Effect of Wind-Drift on Radar-Derived Surface Rainfall Estimations." *Atmospheric Research* 85, no. 2: 217–229. <https://doi.org/10.1016/j.atmosres.2006.09.010>.
- Lewis, H., D. L. Harrison, and M. Kitchen. 2007. "Local Vertical Profile Corrections Using Data From Multiple Scan Elevations." In: 33rd Conference on Radar Meteorology. American Meteorological Society. 2.6. https://ams.confex.com/ams/33Radar/techprogram/paper_123355.htm.
- Marshall, J. S., W. Hirschfeld, and K. L. S. Gunn. 1955. "Advances in Radar Weather." In *Advances in Geophysics*, edited by H. E. Landsberg, vol. 2, 1–56. Elsevier. <https://www.sciencedirect.com/science/article/pii/S0065268708603106>.
- Michelson, D., A. Henja, S. Ernes, et al. 2018. "BALTRAD Advanced Weather Radar Networking." *Journal of Open Research Software* 6, no. 1: 12. <https://doi.org/10.5334/jors.193>.
- NASA. 2013. "NASA Shuttle Radar Topography Mission (SRTM)." Shuttle Radar Topography Mission (SRTM) Global. Distributed by OpenTopography. <https://doi.org/10.5069/G9445JDF>. itemType: dataset.
- Neely, R. R., III, L. Bennett, A. Blyth, et al. 2018. "The NCAS Mobile Dual-Polarisation Doppler X-Band Weather Radar (NXPol)." *Atmospheric Measurement Techniques* 11, no. 12: 6481–6494. <https://doi.org/10.5194/amt-11-6481-2018>.
- Neely, R. R., III, L. Parry, D. Dufton, L. Bennett, and C. Collier. 2021. "Radar Applications in Northern Scotland (RAiNS)." *Journal of Hydrometeorology* 22, no. 2: 483–498. <https://doi.org/10.1175/JHM-D-19-0184.1>.
- Ośródk, K., and J. Szturc. 2015. "Quality-Based Generation of Weather Radar Cartesian Products." *Atmospheric Measurement Techniques* 8, no. 5: 2173–2181. <https://doi.org/10.5194/amt-8-2173-2015>.
- Ośródk, K., J. Szturc, and A. Jurczyk. 2014. "Chain of Data Quality Algorithms for 3-D Single-Polarization Radar Reflectivity (RADVOL-QC System)." *Meteorological Applications* 21, no. 2: 256–270. <https://doi.org/10.1002/met.1323>.
- Pejic, V., J. Soderholm, K. Mühlbauer, V. Louf, and S. Trömel. 2022. "Five Years Calibrated Observations From the University of Bonn X-Band Weather Radar (BoXPOL)." *Scientific Data* 9, no. 1: 551. <https://doi.org/10.1038/s41597-022-01656-0>.
- Ravuri, S., K. Lenc, M. Willson, et al. 2021. "Skilful Precipitation Nowcasting Using Deep Generative Models of Radar." *Nature* 597, no. 7878: 672–677. <https://doi.org/10.1038/s41586-021-03854-z>.
- Ryzhkov, A., P. Zhang, P. Buković, J. Zhang, and S. Cocks. 2022. "Polarimetric Radar Quantitative Precipitation Estimation." *Remote Sensing* 14, no. 7: 1695. <https://doi.org/10.3390/rs14071695>.
- Saltikoff, E., K. Friedrich, J. Soderholm, et al. 2019. "An Overview of Using Weather Radar for Climatological Studies: Successes, Challenges, and Potential." *Bulletin of the American Meteorological Society* 100: 1739–1752. <https://doi.org/10.1175/BAMS-D-18-0166.1>.
- Saltikoff, E., G. Haase, L. Delobbe, et al. 2019. "OPERA the Radar Project." *Atmosphere* 10, no. 6: 320. <https://doi.org/10.3390/atmos10060320>.
- Sandford, C., and N. Gaussiat. 2012. "Use of a Radar Quality Index to Mitigate the Effects of Attenuation at C-Band in the UK Composite." In *Proceedings of 7th European Conference on Radar in Meteorology and Hydrology*. Meteo France.
- Sandford, C., A. Illingworth, and R. Thompson. 2017. "The Potential Use of the Linear Depolarization Ratio to Distinguish Between Convective and Stratiform Rainfall to Improve Radar Rain-Rate Estimates." *Journal of Applied Meteorology and Climatology* 56, no. 11: 2927–2940. <https://doi.org/10.1175/JAMC-D-17-0014.1>.
- Sherman, Z., M. Grover, R. Jackson, et al. 2024. "Effective Visualization of Radar Data for Users Impacted by Color Vision Deficiency." *Bulletin of the American Meteorological Society* 105: E1479–E1489. <https://doi.org/10.1175/BAMS-D-23-0056.1>.
- Testud, J., E. Le Bouar, E. Oblis, and M. Ali-Mehenni. 2000. "The Rain Profiling Algorithm Applied to Polarimetric Weather Radar." *Journal of Atmospheric and Oceanic Technology* 17, no. 3: 332–356. [https://doi.org/10.1175/1520-0426\(2000\)017<0332:TRPAAT>2.0.CO;2](https://doi.org/10.1175/1520-0426(2000)017<0332:TRPAAT>2.0.CO;2).
- Villarini, G., and W. F. Krajewski. 2010. "Review of the Different Sources of Uncertainty in Single Polarization Radar-Based Estimates of Rainfall." *Surveys in Geophysics* 31, no. 1: 107–129. <https://doi.org/10.1007/s10712-009-9079-x>.
- Wallbank, J. R., D. Dufton, R. R. Neely III, L. Bennett, S. J. Cole, and R. J. Moore. 2022. "Assessing Precipitation From a Dual-Polarisation X-Band Radar Campaign Using the Grid-To-Grid Hydrological Model." *Journal of Hydrology* 613: 128311. <https://doi.org/10.1016/j.jhydrol.2022.128311>.
- Werner, M., and M. Cranston. 2009. "Understanding the Value of Radar Rainfall Nowcasts in Flood Forecasting and Warning in Flashy Catchments." *Meteorological Applications* 16, no. 1: 41–55. <https://doi.org/10.1002/met.125>.
- Zhang, J., K. Howard, C. Langston, et al. 2011. "National Mosaic and Multi-Sensor QPE (NMQ) System: Description, Results, and Future Plans." *Bulletin of the American Meteorological Society* 92, no. 10: 1321–1338. <https://doi.org/10.1175/2011BAMS-D-11-00047.1>.