Weather forecasts nowadays rely heavily numerical weather prediction (NWP) models (Bauer et al., 2015), and commonly an ensemble of predictions is used to represent uncertainties (Buizza, 2019). NWP models do not output forecasts at specific sites (that most customers require) but instead output “average” values for grid boxes. If sub-grid variability (i.e., the variation seen amongst all point values observed within the same model gridbox) is low, then raw NWP forecasts are representative of point values. But if sub-grid variability is high raw NWP forecasts inevitably fail. The scale mismatch between in situ observations and gridded NWP forecasts is called representativeness error (Janjić et al., 2018).

The two most common strategies to address sub-grid variability problems are using km-scale NWP models (e.g., ∼2 km), or using calibrated post-processing (PP) techniques to statistically convert from gridbox to point forecasts4–6. For predicting rainfall, the parameter central to this article, highresolution models, whilst showing much more realistic-looking spatial patterns, and exhibiting improvements in forecast skill7,8, have limited geographical coverage. This is because of computational constraints, which imply that one such model might only cover ~0.2% of the world. For global coverage PP techniques are a better prospect, and they have historically performed well in improving forecasts of dry weather6,9, but as previous authors themselves acknowledge, many challenges and issues remain.

Ours is a non-local gridbox-analogue approach, formulated via the principles of conditional verification10, with some structural similarities to quantile regression forests11,12.

We call the method “ecPoint”—“ec” for ECMWF, i.e. the European

Centre for Medium-Range Weather Forecasts, and “Point” for point

forecasts.

Sub-grid variability in rainfall is itself very variable (Fig. 1) and

Relates closely to the weather situation. There are clear-cut physical reasons for this. Dynamics-driven (large-scale) rainfall, often

Related to atmospheric fronts, arises from steady ascent of moist

Air across regions typically larger than GM scales (Fig. 1a). As

Rainfall rates mirror ascent rates, rainfall rate sub-grid variability

Tends to be small. Conversely instability-driven rainfall (i.e.

Showers/convection) arises from localised pockets of rapid ascent,

Which are typically hundreds of metres to kilometres across. So

During convection rainfall rate sub-grid variability, on GM scales,

Can be very large indeed13.

The embodiment of ecPoint is that features of the NWP gridbox

Forecast output (and other global datasets) can tell us what degree of

Sub-grid variability to expect. For example, NWP output commonly

Subdivides rainfall into dynamics-driven and convective, and then for

Convective cases shower movement speed can be approximated by

(e.g.) the 700 hPa wind speed. So by using the convective rainfall

Fraction and the 700 hPa wind speed (two “governing variables”) we

Can distinguish each of the three “gridbox-weather-types” on Fig. 1,

To anticipate a priori the expected sub-grid variability, and accordingly convert each forecast for each gridbox into a probabilistic point

Rainfall prediction (going from red to blue on Fig. 1 PDFs). To our

Knowledge this general approach, based on first principles of precipitation generation, has not been used before except in a limited

Way for nowcasting14 (Table 1, row 11). Another powerful feature of

Our approach is that each gridbox-weather-type is also associated, via

Calibration, with a gridscale bias-correction factor.

The logic outlined above could be successfully applied to a

Single deterministic forecast but in NWP ensembles furnish the

Most useful predictions2,15. So instead we apply separately to each

Ensemble member, creating an ensemble of probabilistic realisations (or “ensemble of ensembles”) that we merge to give the final

Probabilistic point forecast. Whilst Fig. 1 shows just three gridbox-weather-types, ECMWF’s current ecPoint-Rainfall system

Uses 214 such types, defined in decision tree form.