Supplementary Material for "Quantile-adaptive probabilistic forecast combining"

Contents

S1 Additional methodological details

S1.1 Illustration of the CDFs obtained from horizontal, vertical, radial averaging, and their corresponding PDFs

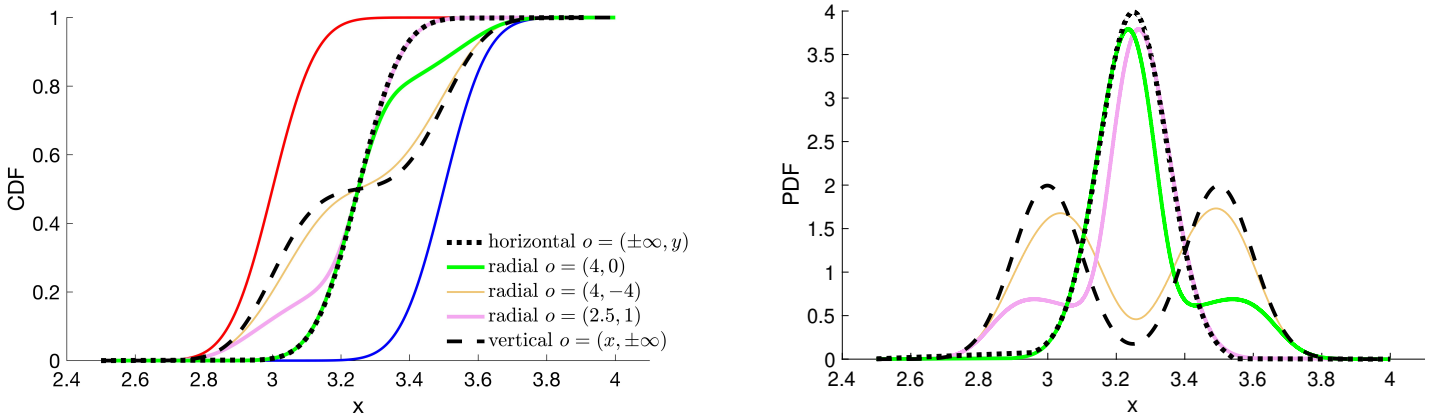


Figure S1: Horizontal, vertical and radial average forecasts (left) and their corresponding PDFs (right).

S1.2 Proof of Proposition 1

We provide the proof of Proposition 1 below, as it was deferred from the main manuscript.

Proof of Proposition 1. Write $F_i(x_i(\theta)) = -\tan\theta\,(x_i(\theta) - o_x) + o_y$. Then $x_i(\theta) = \frac{F_i(x_i(\theta)) - o_y}{-\tan\theta} + o_x$. Taking the derivative of x_i with respect to θ and rearranging yield that $\frac{\mathrm{d}x_i}{\mathrm{d}\theta} = \frac{\tan\theta \sec^2(\theta)(F_i(x_i(\theta)) - o_y))}{\tan\theta + f_i(x_i(\theta))}$. By the chain rule,

$$\begin{aligned}\hat{f}_{r,o}\left(\frac{1}{n}\sum_{i=1}^n x_i(\theta)\right) &= \frac{\mathrm{d}\left(\frac{1}{n}\sum_{i=1}^n F_i(x_i)\right)}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}\left(\frac{1}{n}\sum_{i=1}^n x_i\right)} = \left(\sum_{i=1}^n f_i(x_i(\theta))\frac{\mathrm{d}x_i}{\mathrm{d}\theta}\right)\Bigg/\left(\sum_{i=1}^n \frac{\mathrm{d}x_i}{\mathrm{d}\theta}\right)\\ &= \left(\sum_{i=1}^n \frac{f_i(x_i(\theta))(F_i(x_i(\theta)) - o_y)}{\tan\theta + f_i(x_i(\theta))}\right)\Bigg/\left(\sum_{i=1}^n \frac{F_i(x_i(\theta)) - o_y}{\tan\theta + f_i(x_i(\theta))}\right).\end{aligned}$$

When $o_y \to \pm\infty$,

$$\begin{aligned}&\left(\sum_{i=1}^n \frac{f_i(x_i(\theta))(F_i(x_i(\theta)) - o_y)}{\tan\theta + f_i(x_i(\theta))}\right)\Bigg/\left(\sum_{i=1}^n \frac{(F_i(x_i(\theta)) - o_y)}{\tan\theta + f_i(x_i(\theta))}\right)\\ &= \left(\sum_{i=1}^n \frac{f_i(x_i(\theta))F_i(x_i(\theta))}{\tan\theta + f_i(x_i(\theta))} - \sum_{i=1}^n \frac{f_i(x_i(\theta))o_y}{\tan\theta + f_i(x_i(\theta))}\right)\Bigg/\left(\sum_{i=1}^n \frac{F_i(x_i(\theta))}{\tan\theta + f_i(x_i(\theta))} - \sum_{i=1}^n \frac{o_y}{\tan\theta + f_i(x_i(\theta))}\right)\\ &\to \left(-o_y\sum_{i=1}^n \frac{f_i(x_i(\theta))}{\tan\theta + f_i(x_i(\theta))}\right)\Bigg/\left(-o_y\sum_{i=1}^n \frac{1}{\tan\theta + f_i(x_i(\theta))}\right)\\ &= \left(\sum_{i=1}^n \frac{f_i(x_i(\theta))}{\tan\theta + f_i(x_i(\theta))}\right)\Bigg/\left(\sum_{i=1}^n \frac{1}{\tan\theta + f_i(x_i(\theta))}\right).\end{aligned}$$

When $o_x \to \pm\infty$ such that $o_y/o_x \to 0$, for all i, $x_i(\theta) \to F_i^{-1}(p(\theta))$, where the function $p : [0, \pi/2] \to [0, 1]$ does not depend on i. $F_i(x_i(\theta))$ converges to the same value $p(\theta)$ for all i. Furthermore, $\tan\theta \to 0$. It then follows from the continuity of f_i's that $\hat{f}_{r,o}$ converges pointwise to the harmonic average of the PDFs.

When $o_y \to \pm\infty$ such that $o_x/o_y \to 0$, $\tan\theta \to \infty$ as $\theta \to \pi/2$, that

$$\begin{aligned}\hat{f}_{r,o}\left(\frac{1}{n}\sum_{i=1}^{n} x_i(\theta)\right) &= \left(\sum_{i=1}^{n} \frac{f_i(x_i(\theta))}{\tan\theta + f_i(x_i(\theta))}\right) \Bigg/ \left(\sum_{i=1}^{n} \frac{1}{\tan\theta + f_i(x_i(\theta))}\right) \\ &= \left(\sum_{i=1}^{n} \frac{f_i(x_i(\theta))}{\tan\theta} \cdot \frac{1}{1+\frac{f_i(x_i(\theta))}{\tan\theta}}\right) \Bigg/ \left(\sum_{i=1}^{n} \frac{1}{\tan\theta} \cdot \frac{1}{1+\frac{f_i(x_i(\theta))}{\tan\theta}}\right) \\ &\to \frac{1}{n}\sum_{i=1}^{n} f_i(x_i(\theta)).\end{aligned}$$

Furthermore, as the lines approach vertical lines, for all i, $x_i(\theta) \to x(\theta)$, where the function $x : [0, \pi/2] \to \mathbb{R}$ does not depend on i. Thus, the density of the radial average converges pointwise to the average of the PDFs.

The convergence of the CDFs follows immediately from the formula of $\hat{F}_{r,o}$. As $\theta \to \pi/2$, $x_1(\theta), \ldots, x_n(\theta) \to x$, for some $x \in \mathbb{R}$, and thus $\hat{F}_{r,o}(x) \to (F_1(x) + \cdots + F_n(x))/n$ pointwise. As $\theta \to 0$, $\hat{F}_{r,o}\left((x_1(\theta) + \cdots + x_n(\theta))/n\right), F_1(x_1(\theta)), \ldots, F_n(x_n(\theta)) \to p$, for some $p \in [0,1]$. Then $\hat{F}_{r,o}^{-1}(p) \to \hat{F}_h^{-1}(p)$ pointwise. The convergence to angular averaging follows by comparing the formula of CDFs for radial and angular averaging. □

S1.3 The relationship between radial, vertical and horizontal averaging

This section presents a new perspective on radial averaging. We show that the radial average forecast can be obtained by only averaging horizontally, or only averaging vertically after suitable transformations of the individual CDFs. This connects our method to generalized linear pools (Dawid et al., 1995; Ranjan & Gneiting, 2010; Gneiting & Ranjan, 2013), where forecasts are first transformed, then combined, and finally mapped back. Specifically, a generalized linear pool combines n CDF forecasts $F_1, \ldots, F_n$ as

$$G(x) = h^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} h\left(F_i(x)\right)\right), \tag{S1}$$

where h is a continuous and strictly monotone link function, with domain $[0,1]$ or a subset of $[0,1]$ and range $\mathbb{R}$ or a subset of $\mathbb{R}$, and h^{-1} denotes its inverse.

In what follows we introduce the transformations used in our setting and show how they lead to the vertical or horizontal averaging representations. Denote by $\mathcal{F}$ the set of all cumulative distribution functions on $\mathbb{R}$, and by $\mathcal{G}_I$ the set of all increasing functions with domain I.

Proposition S1.1. *Let $F_1, \ldots, F_n$ be the individual CDF forecasts, $\hat{F}_{r,o,o'}$ be the radial average forecast with focal points $o = (o_x, o_y)$ and $o' = (o_x, o_y')$, where $o_y' > o_y$, and $\hat{F}_{r,\widetilde{o},\widetilde{o}'}$ be the radial average forecast with focal points $\widetilde{o} = (\widetilde{o}_x, \widetilde{o}_y)$ and $\widetilde{o}' = (\widetilde{o}_x', \widetilde{o}_y)$, where $\widetilde{o}_x' > \widetilde{o}_x$. Then, there exist operators $T : \mathcal{F} \to \mathcal{F}$ and $S : \mathcal{G}_{[0,1]} \to \mathcal{G}_{(-\infty,\infty)}$ such that*

$$\hat{F}_{r,o,o'}(x) = T^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} TF_i\right)(x),\ x \in \mathbb{R},$$

and if the individual distributions are supported on $[\widetilde{o}_x, \widetilde{o}_x']$,

$$\hat{F}_{r,\widetilde{o},\widetilde{o}'}^{-1}(\alpha) = S^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} SF_i^{-1}\right)(\alpha),\ \alpha \in [0,1].$$

Proof. Without loss of generality, it suffices to consider the case $n = 2$, that is, two individual CDF forecasts F_1 and F_2; the general case $n > 2$ then follows by straightforward extension of the same argument.

Suppose the focal points are $o = (o_x, o_y)$ and $o' = (o_x, o'_y)$, where $o'_y > o_y$. The radial average forecast can be obtained by transforming the CDFs at each probability level, as illustrated in the left panel of Figure S2. Denote by $\mathcal{F}_I$ the collection of functions obtained by restricting a CDF to domain I, i.e., $\mathcal{F}_I := \{F|_I : F \in \mathcal{F}\}$. Given $F \in \mathcal{F}_{(-\infty,o_x)}$, the transformation, or more precisely, the operator, defined by the rays originating from the lower focal point is $T_1 : \mathcal{F}_{(-\infty,o_x)} \to \mathcal{F}_{(-\infty,o_x)}$, $(T_1F)^{-1}(\alpha) = o_x - \frac{(o_x - F^{-1}(\alpha))(o'_y - o_y)}{\alpha - o_y}$, for $\alpha < F(o_x)$. The arrowed lines in the left triangle in the left panel of the figure illustrates this transformation. The inverse of T_1 is $T_1^{-1} : \mathcal{F}_{(-\infty,o_x)} \to \mathcal{F}_{(-\infty,o_x)}$, $(T_1^{-1}\tilde{F})^{-1}(\alpha) = o_x - \frac{(o_x - \tilde{F}^{-1}(\alpha))(\alpha - o_y)}{o'_y - o_y}$, for $\tilde{F} \in \mathcal{F}_{(-\infty,o_x)}$. It can be easily verified that $(T_1^{-1}T_1F)^{-1}(\alpha) = F^{-1}(\alpha)$. Inspecting the left triangle, we see that

$$T_1\hat{F}_{r,o,o'}(x) = 1/2\,(T_1F_1(x) + T_1F_2(x))\,, \tag{S2}$$

or equivalently, $\hat{F}_{r,o,o'}(x) = T_1^{-1}\,(1/2(T_1F_1 + T_1F_2))\,(x)$, for all $x < o_x$, where the operator T_1 is applied to functions $\hat{F}_{r,o,o'}, F_1$ and F_2 restricted to the domain $(-\infty, o_x)$.

The right-hand side of (S2) is the vertical average of the transformed functions. This shows that the vertical average of the transformed functions is the transformed radial average of the original CDFs.

Similarly, as illustrated by the right triangle of the left panel of the figure, the transformation that applies to the right part of the distributions is $T_2 : \mathcal{F}_{[o_x,\infty)} \to \mathcal{F}_{[o_x,\infty)}$, $(T_2F)^{-1}(\alpha) = o_x + \frac{(o_x - F^{-1}(\alpha))(o'_y - o_y)}{o'_y - \alpha}$, for $\alpha \geq F(o_x)$.

Therefore, the operator $T : \mathcal{F} \to \mathcal{F}$ is given by

$$TF(x) = \begin{cases} T_1F(x), & \text{if } x < o_x, \\ T_2F(x), & \text{if } x \geq o_x. \end{cases}$$

for any $F \in \mathcal{F}$.

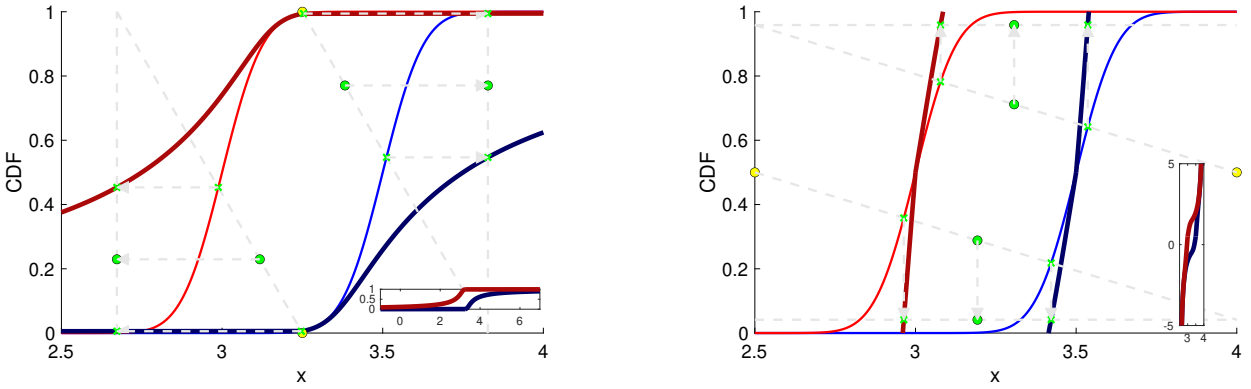


Figure S2: The thin lines are the two individual CDFs to be averaged, and the thick lines are their transformations. The inset of each panel provides a broader view of the transformed curves. The left panel illustrates that averaging vertically yields the radial average. Every quantile of the original CDFs is transformed to obtain the new thick lines, transforming back the vertical average of which yields the radial average of the original CDFs. The right panel illustrates that averaging horizontally yields the radial average. Every probability of the original CDFs is transformed to obtain the new thick lines, transforming back the horizontal average of which yields the radial average of the original CDFs.

When the focal points are $\widetilde{o} = (\widetilde{o}_x, \widetilde{o}_y)$ and $\widetilde{o}' = (\widetilde{o}'_x, \widetilde{o}_y)$, where $\widetilde{o}'_x > \widetilde{o}_x$, the radial average forecast can be obtained by another way of transformation, i.e., transforming each of the probabilities of the CDFs, as illustrated in the right panel of Figure S2. Averaging horizontally

the transformed functions, i.e., the thick lines, yields a function, which is the transformation of the radial average of the original CDFs. The transformations have a similar form to T_1 and T_2. Given $Q \in \mathcal{G}_{[0,\widetilde{o}_y)}$, the transformation defined by the rays originating from the left focal point is $S_1 : \mathcal{G}_{[0,\widetilde{o}_y)} \to \mathcal{G}_{(-\infty,\widetilde{o}_y)}$, $(S_1 Q)^{-1}(x) = \widetilde{o}_y - \frac{\left(\widetilde{o}_y - Q^{-1}(x)\right)(\widetilde{o}'_x - \widetilde{o}_x)}{x - \widetilde{o}_x}$, for $x < Q(\widetilde{o}_y)$. The inverse of S_1 is such that $(S_1^{-1}\widetilde{Q})^{-1}(x) = \widetilde{o}_y - \frac{\left(\widetilde{o}_y - \widetilde{Q}^{-1}(x)\right)(x - \widetilde{o}_x)}{\widetilde{o}'_x - \widetilde{o}_x}$, for $\widetilde{Q} \in \mathcal{G}_{(-\infty,\widetilde{o}_y)}$. The fact that $\left(S_1^{-1} S_1 Q\right)^{-1}(x) = Q^{-1}(x)$ verifies that S_1^{-1} is indeed the inverse operator of S_1. S_1 is illustrated as the arrowed lines in the lower triangle. Moreover,

$$S_1 \hat{F}^{-1}_{r,\widetilde{o},\widetilde{o}'}(\alpha) = 1/2 \left(S_1 F_1^{-1}(\alpha) + S_1 F_2^{-1}(\alpha) \right), \tag{S3}$$

for all $\alpha < \widetilde{o}_y$, or equivalently, $\hat{F}^{-1}_{r,\widetilde{o},\widetilde{o}'}(\alpha) = S_1^{-1}\left(1/2\left(S_1 F_1^{-1} + S_1 F_2^{-1}\right)\right)(\alpha)$. The right-hand side of (S3) is the horizontal average of the transformed functions. Similarly, as illustrated by the upper triangle of the right panel, the transformation that applies to the upper part of the distribution is such that $(S_2 Q)^{-1}(x) = \widetilde{o}_y + \frac{\left(\widetilde{o}_y - Q^{-1}(x)\right)(\widetilde{o}'_x - \widetilde{o}_x)}{\widetilde{o}'_x - x}$, for $x \geq Q(\widetilde{o}_y)$.

Therefore, the operator $S : \mathcal{G}_{[0,1]} \to \mathcal{G}_{(-\infty,\infty)}$ is given by

$$S F^{-1}(\alpha) = \begin{cases} S_1 F^{-1}(\alpha), & \text{if } \alpha < \widetilde{o}_y, \\ S_2 F^{-1}(\alpha), & \text{if } \alpha \geq \widetilde{o}_y. \end{cases}$$

for any $F \in \mathcal{F}$. □

S1.4 Illustration of angular and radial average forecasts

As we described in Section 5.3, for angular averaging, we optimize over 101 relative angles, and this level of discretization is adequate to locate the optimal angle. This is because, for angular averaging and our proposed radial averaging, the resulting average CDF varies smoothly as the parameter changes. For example, Figure S3 displays the angular average CDFs corresponding to the 101 relative angles for Florida in week 28 and for a forecast horizon of four weeks. The angular average CDFs corresponding to two consecutive integer relative angles are nearly identical, and this can be seen by the color changing smoothly from dark blue for 0 relative degrees to dark red for 100 relative degrees. Consequently, introducing additional relative angles is unnecessary, as it would not yield discernible improvement in performance.

We also display the radial average forecasts corresponding to the focal points on the grid presented in Figure 10 of the main manuscript. As illustrated by Figure S4, the grid contains five types of focal points, indicated by five colors. Each of them yields one type of the desired averaging patterns as depicted in Figure 4 of the main manuscript. For example, positioning the focal point at one of the red points would produce the radial averaging pattern as in Figure 4(c), which is abbreviated by V-H-V.

Figure S5 displays the radial average forecasts computed with the focal points in the grid. Each colored CDF corresponds to the focal points of the same color in Figure S4(right).

S2 Additional information for empirical study

S2.1 Data preprocessing

Let the sorted pairs $(q^{s,h,w}_{i,j}, \alpha_j)$ represent quantiles (of the h-week ahead forecast for location s from forecaster i at forecast origin w) and their associated probability levels for $j = 1, \ldots, 23$, i.e., $\alpha_1 = 0.01, \ldots, \alpha_{23} = 0.99$. In step 1, we define the lower bound of each distribution as $q^{s,h,w}_{i,0} = 2q^{s,h,w}_{i,1} - q^{s,h,w}_{i,2}$, if $2q^{s,h,w}_{i,1} \geq q^{s,h,w}_{i,2}$, or $q^{s,h,w}_{i,0} = 0$, if $2q^{s,h,w}_{i,1} < q^{s,h,w}_{i,2}$, and define the

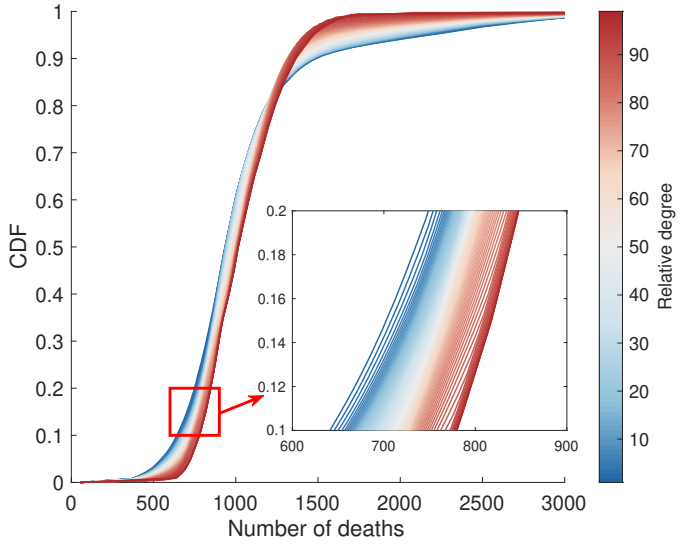


Figure S3: Illustration of the 101 angular average forecasts associated with 101 relative angles in week 28 for Florida. Adjacent CDFs are already very close to each other, indicating that 101 candidate angles provide sufficient resolution.

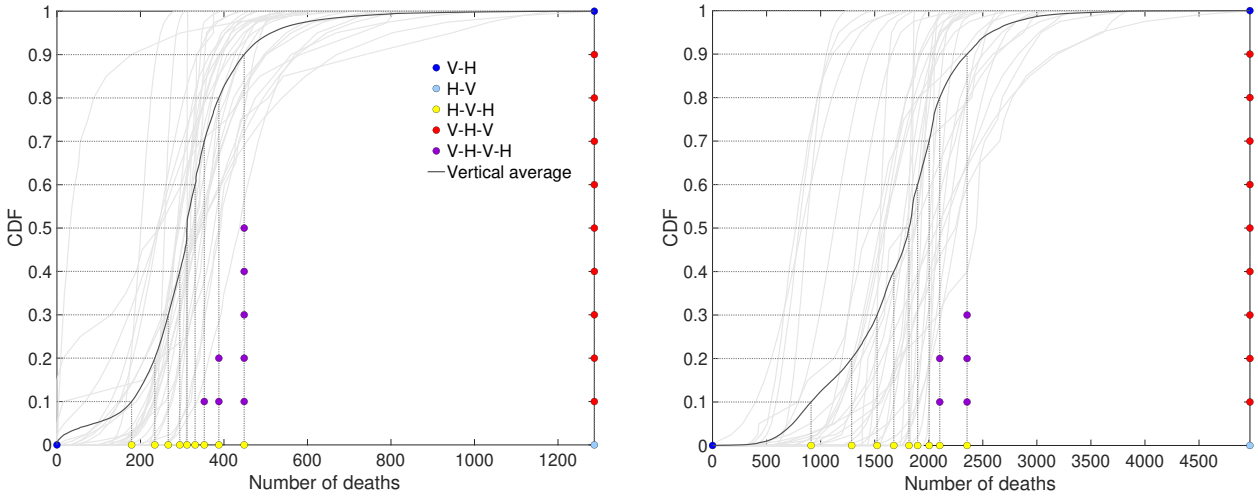
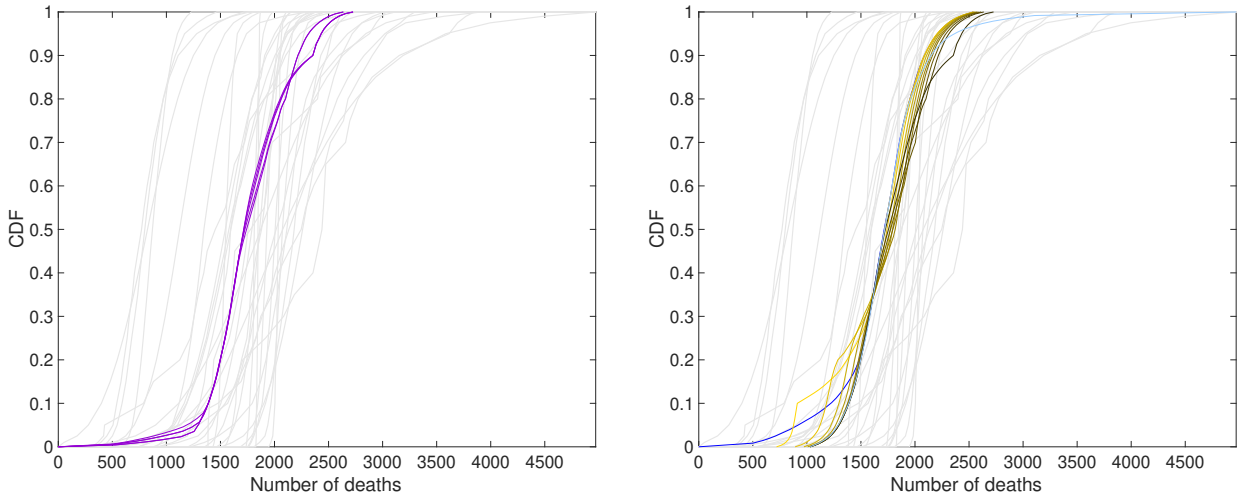
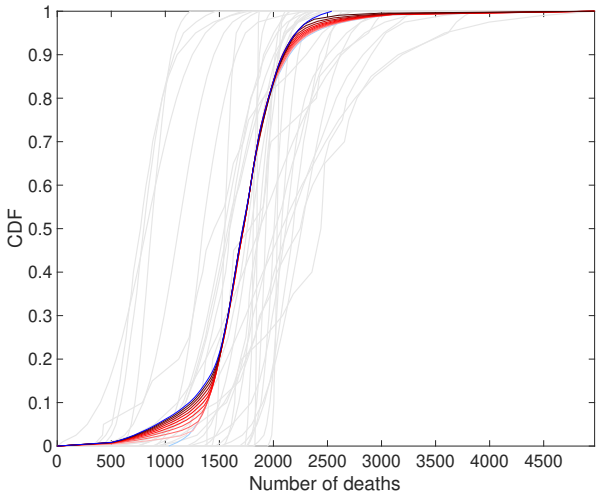


Figure S4: Illustration of the grid for o' and the corresponding radial average 1-week ahead forecasts for California in weeks 50 (left) and 40 (right).

upper bound of each distribution as $q^{s,h,w}_{i,24} = 2q^{s,h,w}_{i,23} - q^{s,h,w}_{i,22}$. In step 2, we observe that some forecasters submit equal predictions for different probability levels. For example, the one-week-ahead quantile forecasts submitted by team DDS on July 26, 2021 for Alaska are the same (which is 384) for probability levels $0.01, 0.025$ and 0.05. This means that the corresponding CDF $F : \mathbb{R}_+ \to [0, 1]$ takes multiple values at $x = 384$, and thus F is not a well-defined function. To solve this issue, we adjust consecutive equal values by adding small increments ϵ (e.g., $\epsilon = 10^{-5}$) until the quantiles are strictly increasing as the probability level increases. For example, we change the second 384 to be $384 + \epsilon$, and change the third 384 to be $384 + 2\epsilon$. In step 3, we define the h-week ahead CDF forecast of forecaster i for location s at forecast origin



(a) (b)



(c)

Figure S5: The radial average 1-week ahead forecasts produced by the five types of focal points for California in week 40. (a): Five radial average forecasts corresponding to the five purple choices of o' in Figure S4(right). (b): 11 radial average forecasts corresponding to the 11 yellow and blue choices of o' on the x-axis in Figure S4(right). (c): 11 radial average forecasts corresponding to the 11 red and blue choices of o' on the vertical line segment on the right in Figure S4 (right).

w as

$$F_i^{s,h,w}(x) = \begin{cases} 0, & x < q_{i,0}^{s,h,w} \\ \alpha_1 \cdot \frac{x - q_{i,0}^{s,h,w}}{q_{i,1}^{s,h,w} - q_{i,0}^{s,h,w}}, & q_{i,0}^{s,h,w} \le x < q_{i,1}^{s,h,w} \\ \frac{\alpha_j - \alpha_{j-1}}{q_{i,j}^{s,h,w} - q_{i,j-1}^{s,h,w}}(x - q_{i,j-1}^{s,h,w}) + \alpha_{j-1}, & q_{i,j-1}^{s,h,w} \le x < q_{i,j}^{s,h,w}, \quad j = 2, \ldots, 24 \\ 1, & x \ge q_{i,24}^{s,h,w}. \end{cases} \tag{S4}$$

S2.2 Additional empirical results

In our empirical study, we evaluate the performance of different forecast combining methods using the average MQS, averaged across all weeks in the out-of-sample period. Note that average scores may be dominated by a few extreme values. To complement the average score comparison, we also report in Table S1 the percentage of weeks in which horizontal, angular and radial averaging performs no worse than the benchmark. In fact, for all states and across the majority of weeks, our method delivers favorable performance relative to the benchmark.

Angular averaging may occasionally coincide with vertical averaging, since its parameter, i.e., the angle, can be optimized so that vertical averaging is actually performed in certain weeks. By contrast, horizontal and radial averaging generally do not coincide with vertical averaging; thus, for these two methods, the reported percentages essentially reflect the frequency of outperforming the benchmark.

Table S1: For each location and all lead times, the percentage (%) of weeks in the out-of-sample period in which each of horizontal, angular and radial averaging is no worse than vertical averaging in terms of MQS.

	AL	AK	AZ	AR	CA	CO	CT	DE	DC	FL	GA	HI	ID
Horizontal	75.9	36.4	80.9	78.4	76.5	74.7	49.4	32.7	45.1	69.1	77.8	30.2	68.5
Angular	86.4	83.3	57.4	78.4	74.7	80.9	63.6	69.1	61.1	66.0	75.3	67.9	75.3
Radial	81.5	71.6	84.0	73.5	80.9	85.8	80.9	74.7	79.6	80.9	80.2	88.3	87.7

	IL	IN	IA	KS	KY	LA	ME	MD	MA	MI	MN	MS	MO
Horizontal	72.2	75.9	64.2	62.3	66.7	59.3	52.5	64.8	74.1	75.9	77.2	67.9	61.1
Angular	77.2	76.5	79.0	75.9	63.6	74.7	56.2	75.3	79.0	84.0	73.5	70.4	67.3
Radial	84.6	88.9	88.3	73.5	67.9	90.1	86.4	84.0	82.7	88.9	92.0	82.7	89.5

	MT	NE	NV	NH	NJ	NM	NY	NC	ND	OH	OK	OR	PA
Horizontal	63.6	48.1	77.2	46.3	77.8	75.9	72.2	70.4	44.4	56.2	54.9	75.9	81.5
Angular	63.0	64.2	77.2	61.7	90.7	80.9	80.9	71.0	76.5	100.0	63.0	88.9	86.4
Radial	71.0	79.0	90.1	87.0	96.3	89.5	90.7	75.3	93.2	63.0	77.2	75.3	93.2

	RI	SC	SD	TN	TX	UT	VT	VA	WA	WV	WI	WY	US
Horizontal	42.6	72.8	51.9	66.0	69.8	69.8	29.0	67.3	71.6	71.6	76.5	42.6	87.0
Angular	70.4	67.3	74.1	71.6	64.2	69.1	59.3	82.7	72.2	87.7	75.9	67.3	82.1
Radial	84.0	86.4	96.9	75.3	79.0	83.3	89.5	77.8	74.7	83.3	80.9	86.4	93.8

As we mentioned in relation to Figure 12 of the main manuscript, in two separate weeks, a forecaster submitted CDFs that were significantly different from the forecasts of other forecasters. These were weeks 45 and 47. For example, the left panel of Figure S6 illustrates all individual 2-week ahead CDF forecasts (as gray curves) for Connecticut in week 45. Among the set of CDF forecasts, most were closely aligned with each other. However, a single forecaster submitted a CDF forecast that was far to the right of the forecasts from the other forecasters (shown in blue). The optimal focal point for our radial averaging method was learned based on past information. For week 45, the optimal focal point for radial averaging was a single focal point at the upper bound. This means that, at all probability levels, the extremely large quantile forecasts (all equal to approximately 2300) were included in the averaging process, thus inflating the radial average forecast. Similarly, the horizontal and angular average CDF

forecasts were also pulled upward due to the influence of the extreme CDF forecast. Vertical averaging was less affected by this outlying forecast, because vertical averaging averages probabilities, and the probabilities for most mortality numbers given by the outlying forecast were 0. It is also worth noting that, when such an extreme forecast occurred for the second time, in week 47 (right panel of Figure S6), radial averaging performed well, as it had successfully adjusted its parameter due to the first occurrence of the extreme data.

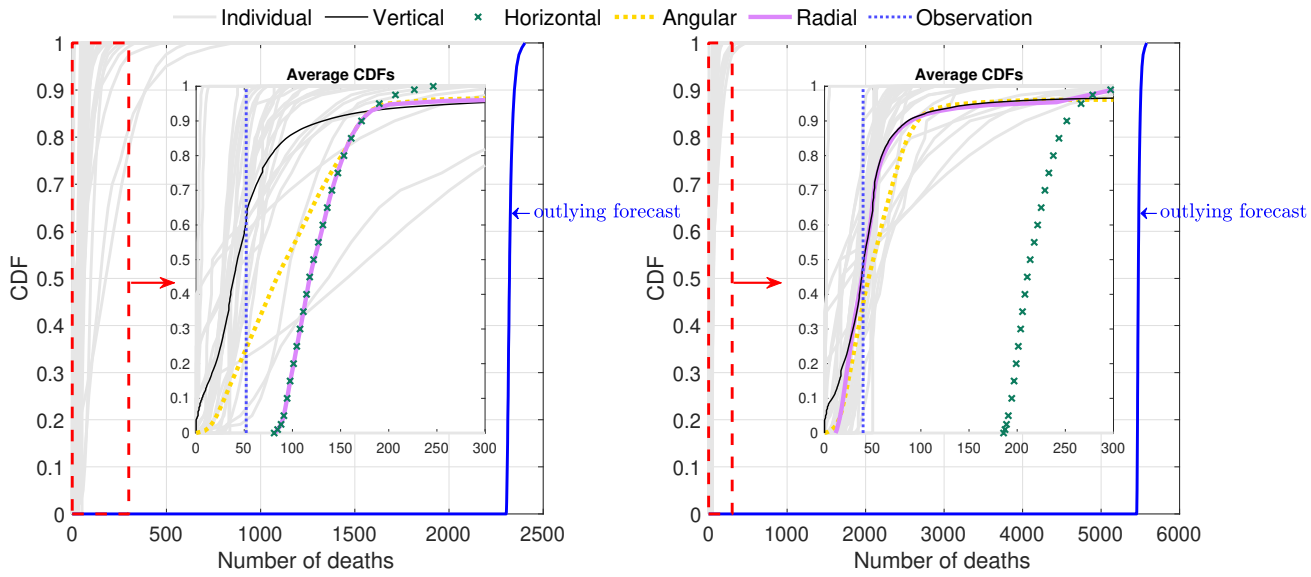


Figure S6: Individual 2-week ahead CDF forecasts, and average CDF forecasts for Connecticut in weeks 45 (left panel) and 47 (right panel). In both weeks, a forecaster submitted substantially large forecasts for all quantiles, which are labeled as outlying forecast. For each panel, the inset provides a magnified view of the region outlined by the red dashed rectangle and displays the four different average forecasts and the 2-week ahead observation.

We provide a visualization of the average CDF forecasts we obtain. As an example, Figure S7 shows the forecasts for the out-of-sample period for New York (state) obtained from the vertical, horizontal, angular and radial averaging methods. Each bar in the figure represents a CDF forecast, with the highest value of the bar being the upper bound of the distribution and the lowest value being the lower bound of the distribution. The vertical and angular average exhibit more dispersed distributions compared with the horizontal and radial average (note the difference in the scaling of the y-axes of the four plots). On the other hand, horizontal and radial averaging produce sharper forecasts, with the probability mass concentrating within smaller intervals. Furthermore, one team submitted much larger forecasts for all quantiles than those submitted by other teams in weeks 45 and 47. These outlying forecasts lead to undesirable upper bounds for the vertical and angular average forecasts and undesirable lower bounds for the horizontal average. By contrast, radial averaging performs well for these two weeks.

S2.3 Choosing the number of focal points via cross-validation

While our method automatically determines the number of focal points, a cross-validation procedure could be used if there is a preference to select this number explicitly. The number of focal points would be treated as a hyperparameter and determined via an expanding window cross-validation procedure. In our implementation of this, for each state, week and forecast horizon, we considered three candidates: one focal point, at most two, and at most three. For each training set, we constructed 20 internal validation steps by successively enlarging an internal training window by one week. All validation steps were performed within the training set to ensure a fair comparison with angular averaging, which does not involve a validation

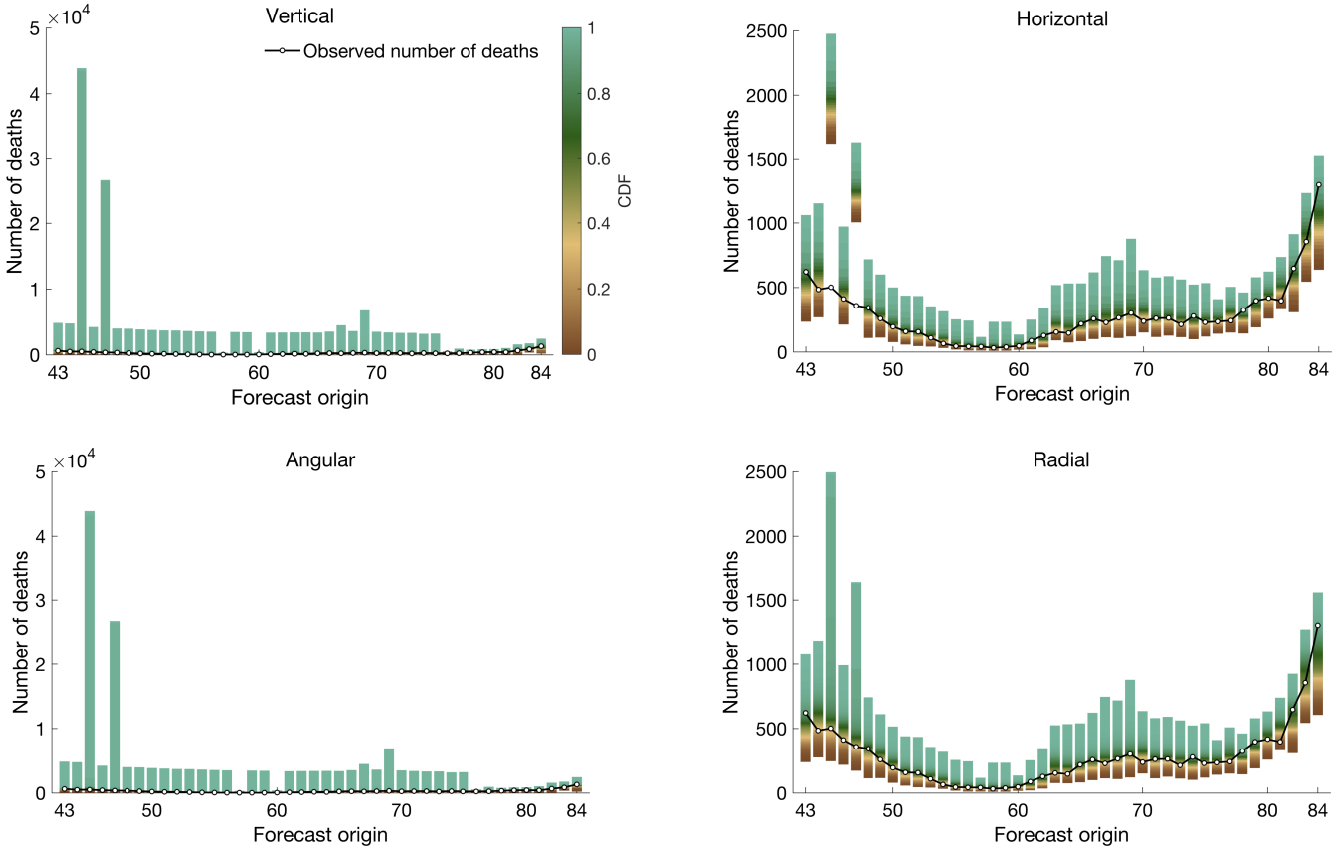


Figure S7: One-week ahead forecasted CDFs using vertical, horizontal, angular and radial averaging methods for the out-of-sample period for New York (state). Each bar displays an average CDF forecast at a forecast origin, with the colors representing the probability that the predicted number of death is no larger than the corresponding values on the vertical axis, which are the quantiles of the distribution. The line represents the observed number of deaths.

stage. At each step, we produced an h-week ahead CDF forecast for each candidate number of focal points. The MQS was then computed for this forecast using the observation at the corresponding horizon. Repeating this process over all 20 validation weeks yielded a series of MQS values for each candidate number of focal points. The MQS was then averaged across all validation weeks for each candidate. Finally, the number of focal points that minimized the average MQS was selected.

The out-of-sample MQS and MQSS results are presented in Table S2. Comparing these results with Tables 1 and 2 of the paper, we see that radial averaging with cross-validation (in Table S2) outperforms vertical, horizontal and angular averaging (in Tables 1 and 2), while it underperforms our implementation of radial averaging in the paper (in Tables 1 and 2), which selects from up to three focal points and directly optimizes over the entire grid. This perhaps is not surprising given that we were only able to use relatively short validation samples in our implementation of cross-validation.

S3 Review of angular combining of distributional forecasts

Angular combining Taylor & Meng (2025) is a method for aggregating multiple CDF forecasts that generalizes the commonly used horizontal and vertical combining approaches. In horizontal combining, the quantile functions are averaged, while in vertical combining, the CDFs are averaged directly. Each method has advantages and drawbacks, and the choice between them

Table S2: For radial averaging with cross-validation, MQS and MQSS averaged across the four lead times, and for each lead time, averaged across the out-of-sample periods for the different groupings of the 52 locations.

Lead time	MQS					MQSS (%)				
(week)	U.S.	High	Medium	Low	All	U.S.	High	Medium	Low	All
Average	697.20	56.90	23.44	7.28	42.05	10.56	1.57	3.69	4.69	3.47
1	554.30	46.22	21.15	6.76	34.90	6.75	4.68	2.14	1.32	2.80
2	618.71	51.03	23.37	7.20	38.57	10.74	6.48	3.79	3.71	4.79
3	734.24	60.50	23.18	7.37	43.89	11.07	-0.08	3.03	5.70	3.07
4	881.56	69.83	26.04	7.79	50.84	12.27	0.30	5.70	7.67	4.76

depends on the true data generating process. Angular combining provides a flexible alternative by introducing an angle parameter θ, that creates a continuous spectrum of combination methods, effectively bridging the gap between horizontal and vertical averaging.

Horizontal averaging corresponds to an angle of $0°$, and vertical averaging corresponds to an angle of $90°$. By selecting an angle $\theta \in [0, 90°]$, angular combining ensures that the resulting function remains a valid CDF, while allowing practitioners to exploit useful aspects of both horizontal and vertical averaging. More specifically, for a fixed θ, angular averaging considers a family of straight lines $y = -\tan\theta(x - c)$, with $c \in (-\infty, \infty)$. Each of the lines intersects with the graph of each of the individual CDFs $\{F_i\}_{i=1}^n$. The angular averaging method averages the x- and y-coordinates of these intersection points to produce a point on the combined CDF given by: $(1/n \sum_{i=1}^n x_i(c), 1/nF_i(x_i(c)))$, where $x_i(c)$ is the x-coordinate of the intersection point of the line $y = -\tan\theta(x - c)$ with $y = F_i(x)$. In this formulation of the method, c parameterizes the CDF produced by angular averaging.

References

Dawid, A. P., DeGroot, M. H., Mortera, J., Cooke, R., French, S., Genest, C., Schervish, M., Lindley, D. V., McConway, K., & Winkler, R. L. (1995). Coherent combination of experts' opinions. *Test*, *4*, 263–313.

Gneiting, T., & Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics*, *7*, 1747–1782.

Ranjan, R., & Gneiting, T. (2010). Combining probability forecasts. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *72*, 71–91.

Taylor, J. W., & Meng, X. (2025). Angular combining of forecasts of probability distributions. *Management Science*, Advance online publication. https://doi.org/10.1287/mnsc.2024.05558.