Combining predictive distributions for the time-to-frost

FOCUS DRIVEN STATISTICAL
INFERENCE WITH COMPLEX DATA

Céline Cunen

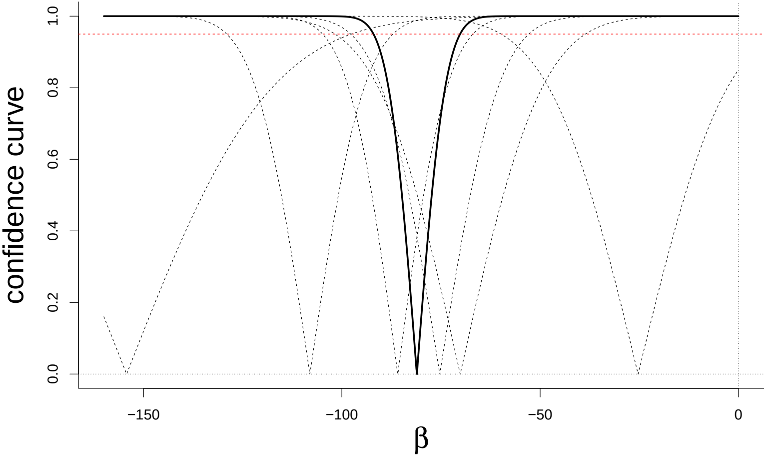
Godt Hjort! 05/12/2023

Focus Driven Statistical Inference with Complex Data

FocuStat 2014–2018

Combination of information, then

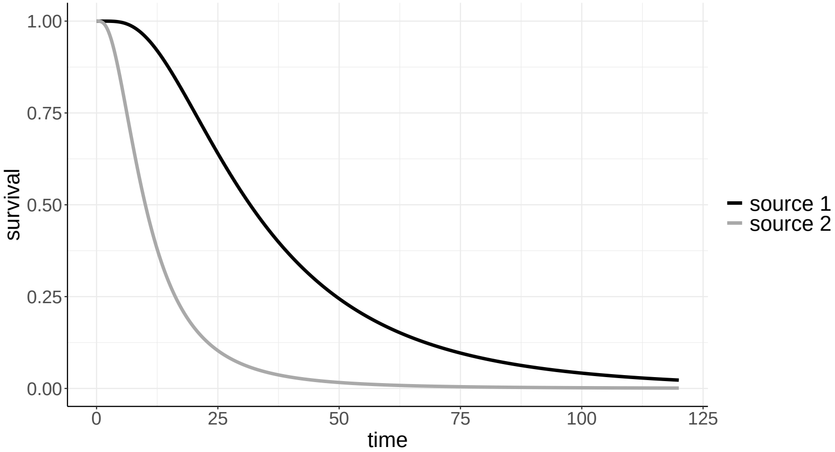
Cunen, C. and Hjort, N.L. (2021). Combining information across diverse sources: The II-CC-FF paradigm. *Scandinavian Journal of Statistics*.



Now, combination of predictions

This is joint work with people at NR, in particular Thea Roksvåg, Claudio Heinrich-Mertsching and Alex Lenkoski.

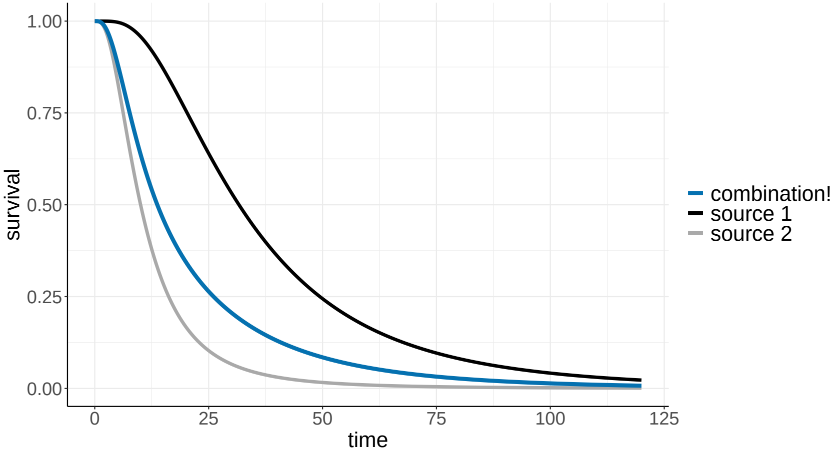
Say we are interested in predicting the time to some event,



Now, combination of predictions

This is joint work with people at NR, in particular Thea Roksvåg, Claudio Heinrich-Mertsching and Alex Lenkoski.

Say we are interested in predicting the time to some event,



The problem

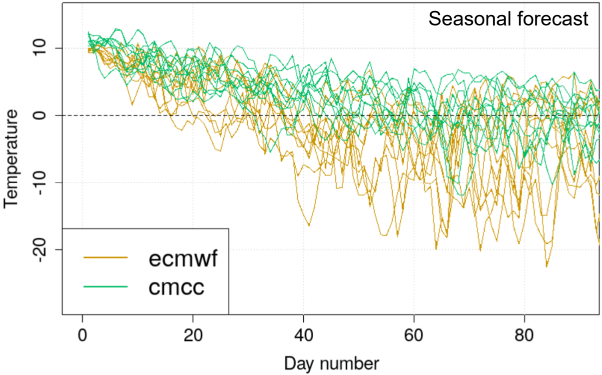
Say we are in Hamar, on the 30th of September 2023: when will we have the first frost?
We have access to two sources of information:

1. the **seasonal forecast**: issued on September 1st.
2. the **subseasonal forecast**: issued on September 30th.

The problem

Say we are in Hamar, on the 30th of September 2023: when will we have the first frost?
We have access to two sources of information:

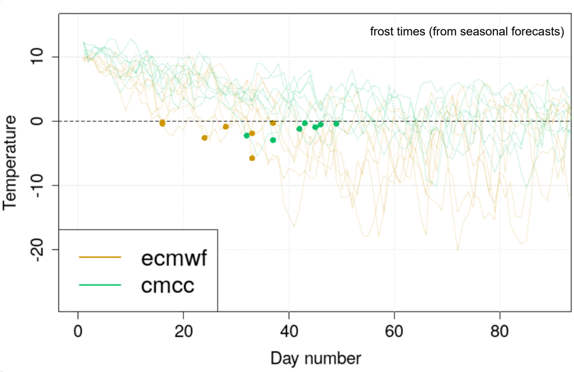
1. the **seasonal forecast**: issued on September 1st.
2. the **subseasonal forecast**: issued on September 30th.



The problem

Say we are in Hamar, on the 30th of September 2023: when will we have the first frost?
We have access to two sources of information:

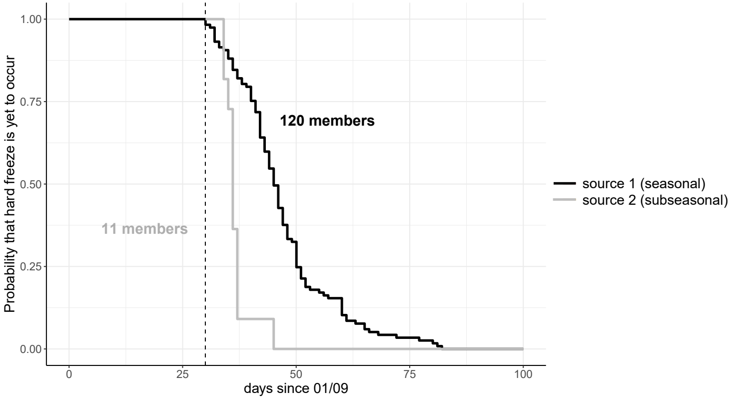
1. the **seasonal forecast**: issued on September 1st.
2. the **subseasonal forecast**: issued on September 30th.



The problem

Say we are in Hamar, on the 30th of September 2023: when will we have the first frost?
We have access to two sources of information:

1. the **seasonal forecast**: issued on September 1st.
2. the **subseasonal forecast**: issued on September 30th.

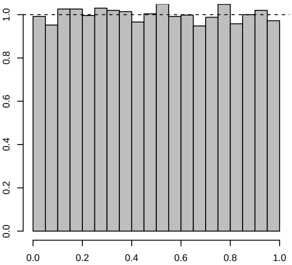


How can we best combine them?

What we want

A combined probabilistic forecast $\widehat{S}_{\text{comb}}(t)$ (or $\widehat{F}_{\text{comb}}(t)$) which is

1. calibrated! $\widehat{F}_{\text{comb}}(T) \sim \mathrm{Unif}(0, 1)$
2. as precise as possible.



Four combination methods

1. LP – the linear pool

$$F_{\text{comb}}(t) = \hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t).$$

Four combination methods

1. LP – the linear pool

$$F_{\mathrm{comb}}(t) = \hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t).$$

2. HB – hazard blending

$$S_{\mathrm{comb}}(t) = \prod_{i:t_i \le t} (1 - \lambda_i), \text{ with } \lambda_i = \frac{\hat{\omega} e_{1,i} + (1 - \hat{\omega}) e_{2,i}}{\hat{\omega} n_{1,i} + (1 - \hat{\omega}) n_{2,i}}.$$

Four combination methods

1. LP – the linear pool

$$F_{\text{comb}}(t) = \hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t).$$

2. HB – hazard blending

$$S_{\text{comb}}(t) = \prod_{i:t_i \leq t} (1 - \lambda_i), \text{ with } \lambda_i = \frac{\hat{\omega} e_{1,i} + (1 - \hat{\omega}) e_{2,i}}{\hat{\omega} n_{1,i} + (1 - \hat{\omega}) n_{2,i}}.$$

3. BP – the Beta pool

$$F_{\text{comb}}(t) = B_{\hat{\alpha},\hat{\beta}}\{\hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t)\}.$$

with $B_{\alpha,\beta}\{\cdot\}$ the cdf of the Beta distribution.

Four combination methods

1. LP – the linear pool

$$F_{\text{comb}}(t) = \hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t).$$

2. HB – hazard blending

$$S_{\text{comb}}(t) = \prod_{i:t_i \leq t} (1 - \lambda_i), \text{ with } \lambda_i = \frac{\hat{\omega} e_{1,i} + (1 - \hat{\omega}) e_{2,i}}{\hat{\omega} n_{1,i} + (1 - \hat{\omega}) n_{2,i}}.$$

3. BP – the Beta pool

$$F_{\text{comb}}(t) = B_{\hat{\alpha},\hat{\beta}}\{\hat{\omega} F_1(t) + (1 - \hat{\omega}) F_2(t)\}.$$

with $B_{\alpha,\beta}\{\cdot\}$ the cdf of the Beta distribution.

4. GFF - the Gaussian forecast filter

$$F_{\text{comb}}(t) = \Phi \left\{ \frac{\hat{\omega} \Phi^{-1}\{F_1(t)\} + (1 - \hat{\omega}) \Phi^{-1}\{F_2(t)\} - \hat{\mu}}{\hat{\sigma}} \right\}.$$

with $\Phi\{\cdot\}$ the standard normal cdf.

How do we estimate the combination parameters?

For example with the GFF method we have (ω, μ, σ).

We need data: realised times-to-frost and corresponding forecasts from the two sources. Here we will use historical data.

We consider two options:

- maximum likelihood;
- minimising the integrated brier score.

In the following we will look at LP, GFF and BP with ML, and HB with minIBS.

Maximum likelihood

For the GFF method the log-likelihood looks like

$$\ell(\omega, \mu, \sigma) = \sum_{i=1}^{n} \log \left(\phi \left[\frac{\omega \Phi^{-1}\{F_{1,i}(t_i)\} + (1-\omega)\Phi^{-1}\{F_{2,i}(t_i)\} - \mu}{\sigma} \right] \right.$$
$$\left. \frac{1}{\sigma} [\omega \Phi^{-1\prime}\{F_{1,i}(t_i)\} f_{1,i}(t_i) + (1-\omega)\Phi^{-1\prime}\{F_{2,i}(t_i)\} f_{2,i}(t_i)] \right),$$

where we have n years of data: with t_i the realised time-to-frost and

$$(F_{1,i}(t_i), F_{2,i}(t_i), f_{1,i}(t_i), f_{2,i}(t_i)),$$

the predictive CDFs and densities from each source.

Minimising IBS

Minimise

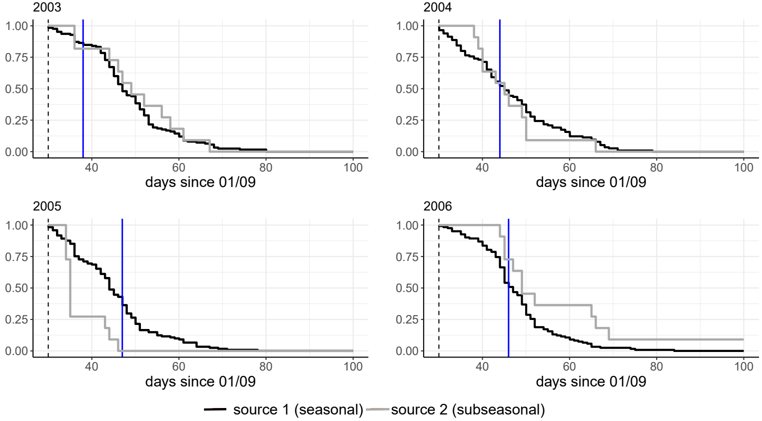
$$\mathrm{ibs}(\omega) = \sum_t \frac{1}{n} \sum_{i=1}^{n} [\mathbb{I}\{t_i > t\} - S_{\mathrm{comb},i}(t)]^2,$$

where $S_{\mathrm{comb},i}(t)$ is the combined predictive survival curve from year i and t_i is the realised time to frost in year i.

Back to Hamar

We have 20 years of historical data:

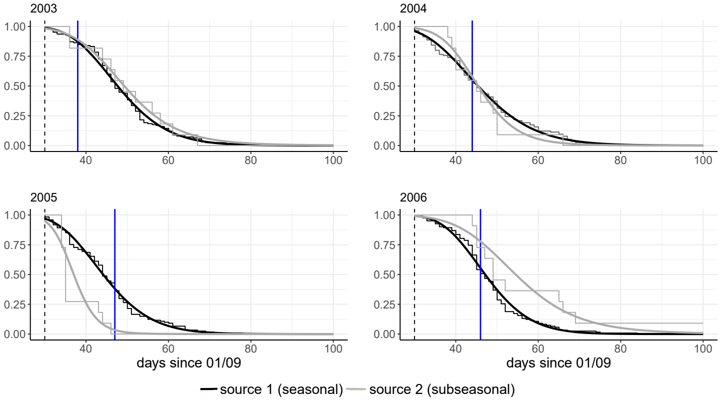
- the realised frost date in each year;
- the forecasts that were made (seasonal + subseasonal).



and so on...

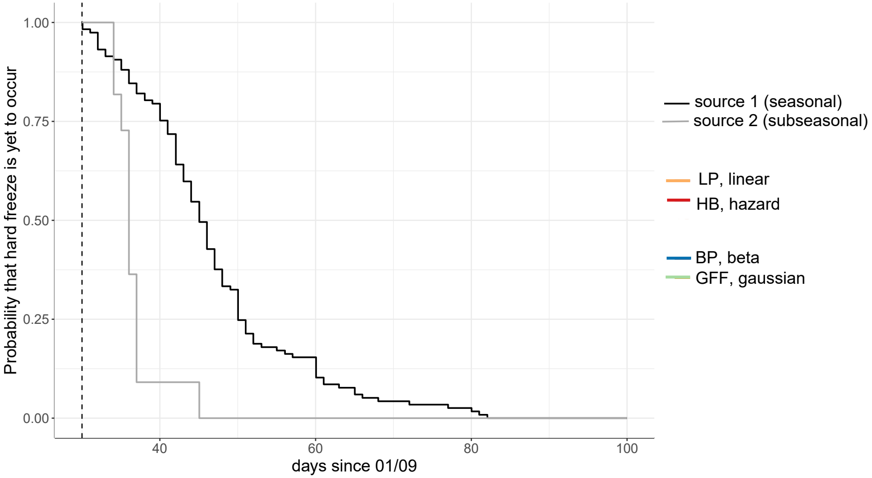
Back to Hamar

Within each year, we have to decide which predictive distributions to use:

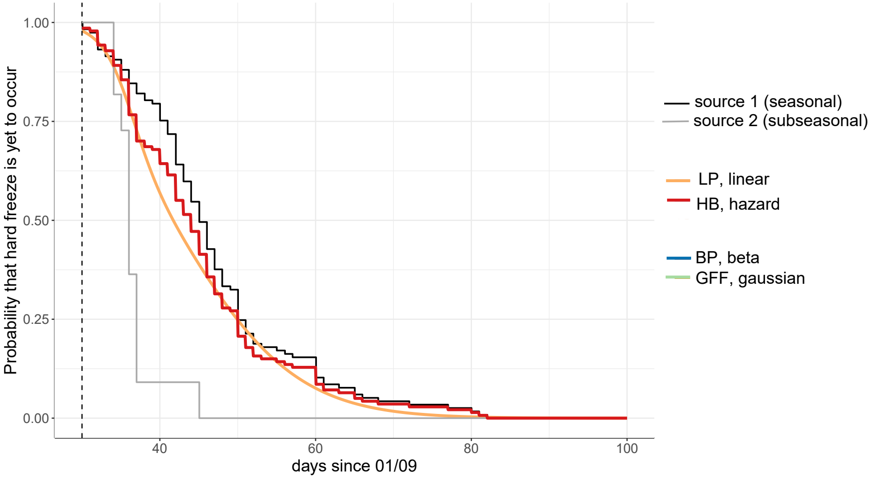


$$F_{comb}(t) = \Phi\left[\frac{\hat{\omega}\Phi^{-1}\{\widehat{F}_1(t)\} + (1-\hat{\omega})\Phi^{-1}\{\widehat{F}_2(t)\} - \hat{\mu}}{\hat{\sigma}}\right].$$

Back to Hamar – combined forecasts on 30/09/2023



Back to Hamar – combined forecasts on 30/09/2023

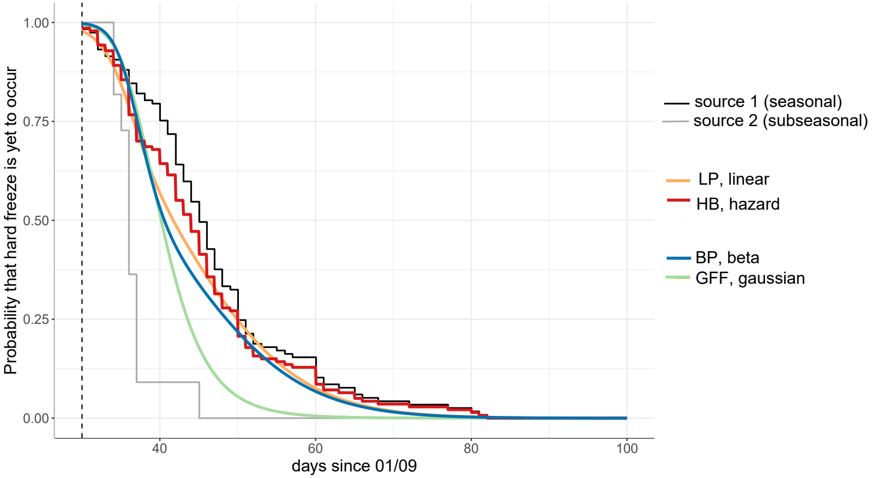


We get

$$\widehat{\omega}_{LP} = 0.73 \quad \text{(weight on seasonal)}$$

$$\widehat{\omega}_{HB} = 0.32$$

Back to Hamar – combined forecasts on 30/09/2023



We get

$$\widehat{\omega}_{\mathrm{LP}} = 0.73 \quad \text{(weight on seasonal)}$$

$$\widehat{\omega}_{\mathrm{HB}} = 0.32$$

$$\widehat{\omega}_{\mathrm{BP}} = 0.45 \quad \hat{\alpha} = 1.48 \quad \hat{\beta} = 1.01$$

$$\widehat{\omega}_{\mathrm{GFF}} = 0.46 \quad \hat{\mu} = 0.36 \quad \hat{\sigma} = 0.94$$

Simulations

$$\begin{aligned}\text{Truth: } & T_i | (x_{1,i}, x_{2,i}) \sim \mathrm{LogNormal}(\xi_i, \tau_0^2) \\ & \text{with } \xi_i = \xi_0 + x_{1,i} + x_{2,i}, \\ & \text{and } X_{1,i} \sim \mathrm{N}(0, \tau_1^2) \quad X_{2,i} \sim \mathrm{N}(0, \tau_2^2).\end{aligned}$$

Simulations

$$\text{Truth: } T_i|(x_{1,i}, x_{2,i}) \sim \mathrm{LogNormal}(\xi_i, \tau_0^2)$$
$$\text{with } \xi_i = \xi_0 + x_{1,i} + x_{2,i},$$
$$\text{and } X_{1,i} \sim \mathrm{N}(0, \tau_1^2) \quad X_{2,i} \sim \mathrm{N}(0, \tau_2^2).$$

$$\text{Source (1): } Y_{i,j}|x_{1,i} \sim \mathrm{LogNormal}(\xi_0 + x_{1,i}, \tau_0^2 + \tau_2^2), \quad n_1 = 100,$$
$$\text{Source (2): } Z_{i,j}|x_{2,i} \sim \mathrm{LogNormal}(\xi_0 + x_{2,i} + b, \tau_0^2 + \tau_1^2), \quad n_2 = 20.$$

with $\tau_0 = \tau_1 = 0.4$.

Simulations

$$\begin{aligned} \text{Truth: } & T_i|(x_{1,i}, x_{2,i}) \sim \mathrm{LogNormal}(\xi_i, \tau_0^2) \\ & \text{with } \xi_i = \xi_0 + x_{1,i} + x_{2,i}, \\ & \text{and } X_{1,i} \sim \mathrm{N}(0, \tau_1^2) \quad X_{2,i} \sim \mathrm{N}(0, \tau_2^2). \end{aligned}$$

$$\begin{aligned} \text{Source (1): } & Y_{i,j}|x_{1,i} \sim \mathrm{LogNormal}(\xi_0 + x_{1,i}, \tau_0^2 + \tau_2^2), \quad n_1 = 100, \\ \text{Source (2): } & Z_{i,j}|x_{2,i} \sim \mathrm{LogNormal}(\xi_0 + x_{2,i} + b, \tau_0^2 + \tau_1^2), \quad n_2 = 20. \end{aligned}$$

with $\tau_0 = \tau_1 = 0.4$.

Simulation scenarios:

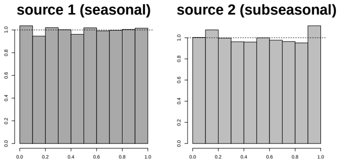
- Both sources calibrated or source (2) biased: $b = 0$ or $b = -0.5$
- Balanced or unbalanced sources: $\tau_2 = 0.5$ or $\tau_2 = 0.2$
- Loads or little training data: $n = 1000$ or $n = 20$

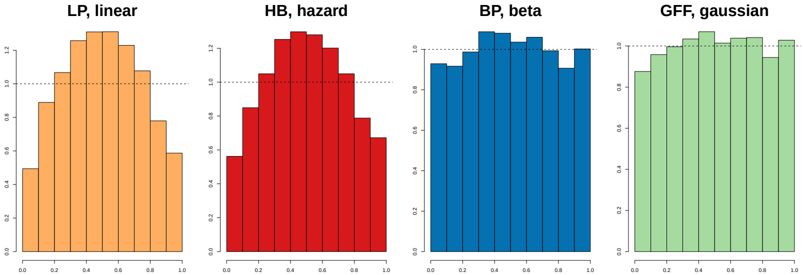
Lessons from simulations

1. Extra combo parameters fix calibration, when both sources are calibrated
2. Extra combo parameters fix calibration, when there is bias
3. More complex combo-methods are often best, but all combination is bad when the sources are unbalanced and there is little training data

1. Extra combo parameters fix calibration, ...

when both sources are calibrated





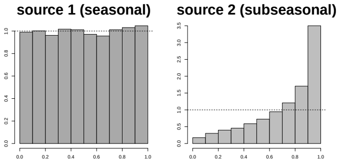
$\widehat{\omega}_{\mathrm{LP}} = 0.32$ $\widehat{\omega}_{\mathrm{HB}} = 0.13$

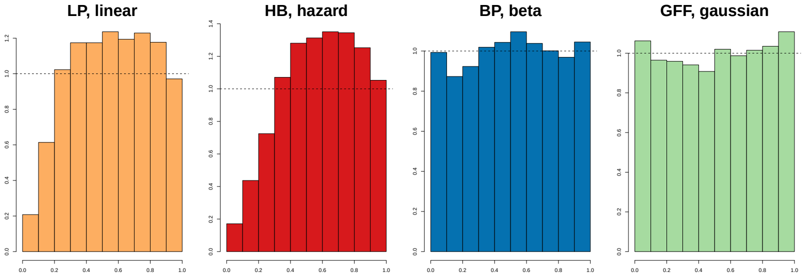
$\widehat{\omega}_{\mathrm{BP}} = 0.43$ $\hat{\alpha} = 1.43$ $\hat{\beta} = 1.53$

$\widehat{\omega}_{\mathrm{GFF}} = 0.45$ $\hat{\mu} = -0.05$ $\hat{\sigma} = 0.89$

2. Extra combo parameters fix calibration, ...

when there is bias





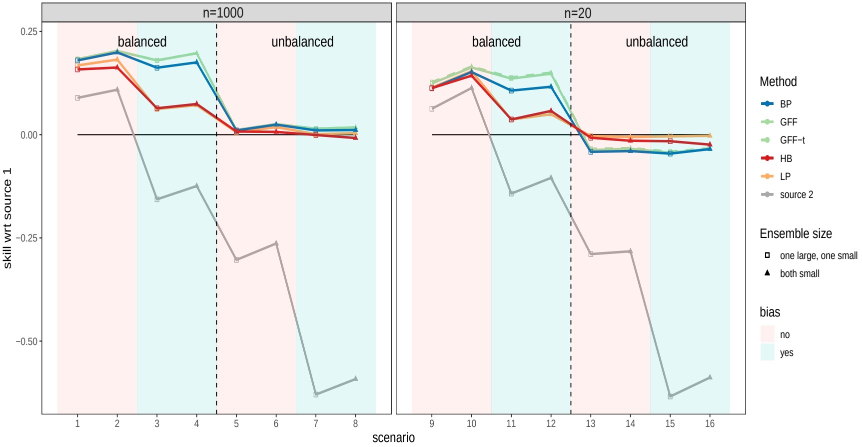
$\widehat{\omega}_{\mathrm{LP}} = 0.73 \qquad \widehat{\omega}_{\mathrm{HB}} = 0.26$

$\widehat{\omega}_{\mathrm{BP}} = 0.42 \quad \hat{\alpha} = 2.32 \quad \hat{\beta} = 1.33$

$\widehat{\omega}_{\mathrm{GFF}} = 0.46 \quad \hat{\mu} = 0.50 \quad \hat{\sigma} = 0.87$

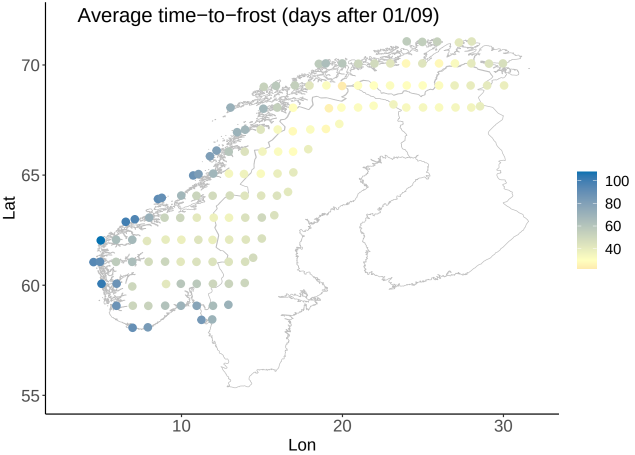
3. More complex combo-methods are often best, ...

but all combination is bad when the sources are unbalanced and there is little training data.



IBS skill wrt to source 1.

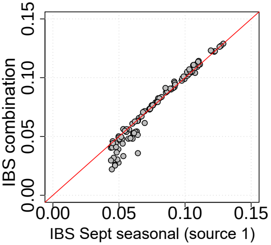
Real data

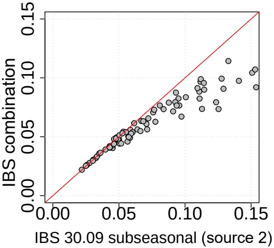


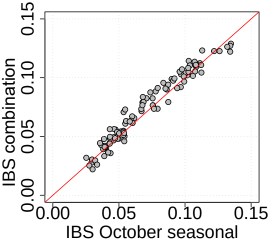
20 years of data, at 138 locations

Some real data results

Combination forecast beats each of the individual sources!

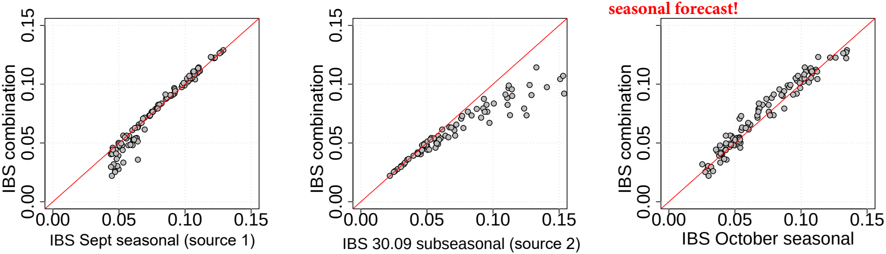






Some real data results

Combination forecast on par with next seasonal forecast!



The hazard blending method has a similar predictive performance as the **seasonal forecast from October 1st**, which will only be available two weeks later.

Good start!

A lot more to do

- More complex estimation schemes for weights and combination parameters:
  - Model for combo parameters in space; time-varying weights;
  - Covariates (elevation).
- We have a lot more forecasts than two:
  - Seasonal forecasts on 01/09, 01/10,...
  - Subseasonal forecasts in 02/09, 09/05, 09/09, 12/09, 16/09, 19/09, 23/09, 26/09, ...
- Other applications?

Thank you!

- Aastveit, K. A., Mitchell, J., Ravazzolo, F., & Van Dijk, H. K. (2018). The evolution of forecast density combinations in economics. *Tinbergen Institute Discussion Paper*.
- Gneiting, T., & Ranjan, R. (2013). Combining predictive distributions. *Electronic Journal of Statistics*.
- Roksvåg, T., Lenkoski, A., Scheuerer, M., Heinrich-Mertsching, C., & Thorarinsdottir, T. L (2023). Probabilistic prediction of the time to hard freeze using seasonal weather forecasts and survival time methods. *Quarterly Journal of the Royal Meteorological Society*.