Achieving seamless verification across sub-seasonal time scales from weather to climate

Paul Dirmeyer[1] & Trent Ford[2]

[1]Center for Ocean-Land-Atmosphere Studies
George Mason University
Fairfax, Virginia, USA

[2]Southern Illinois University
Carbondale, Illinois, USA

How we (still) do forecasting

- Traditional weather forecasts – day-by-day, deterministic or probabilistic, out to about 7 days with current models

COLA Center for Ocean-Land-Atmosphere Studies GEORGE MASON UNIVERSITY

How we (still) do forecasting

- Traditional weather forecasts – day-by-day, deterministic or probabilistic, out to about 7 days with current models

- Arbitrarily we jump to time-averages after some lead time
  - Usually step by weeks, months, etc.

How we (still) do forecasting

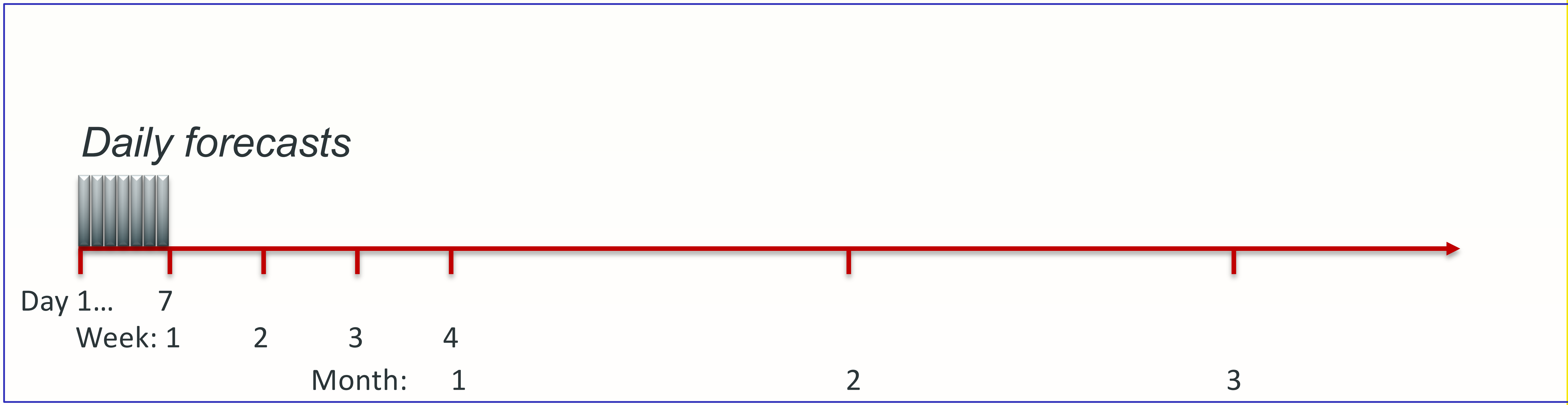
- Traditional weather forecasts – day-by-day, deterministic or probabilistic, out to about 7 days with current models

- Arbitrarily we jump to time-averages after some lead time
  – Usually step by weeks, months, etc.

- These divisions dictated by the construct of our calendar, not the nature of phenomena – Is this the best? Can we be more flexible?

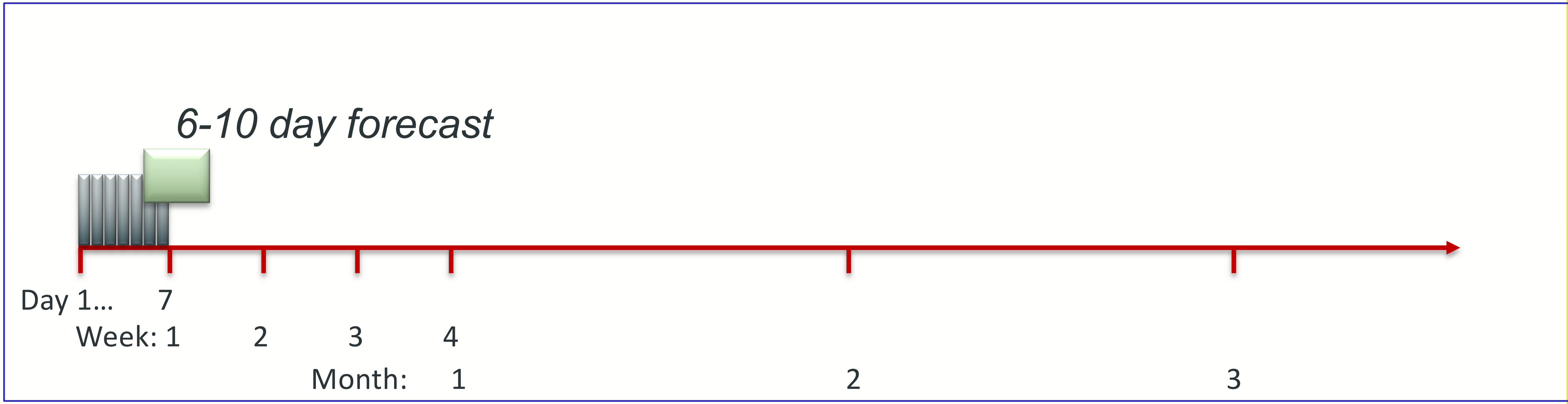
For example: NCEP

- The longer the forecast lead, the wider the "window".



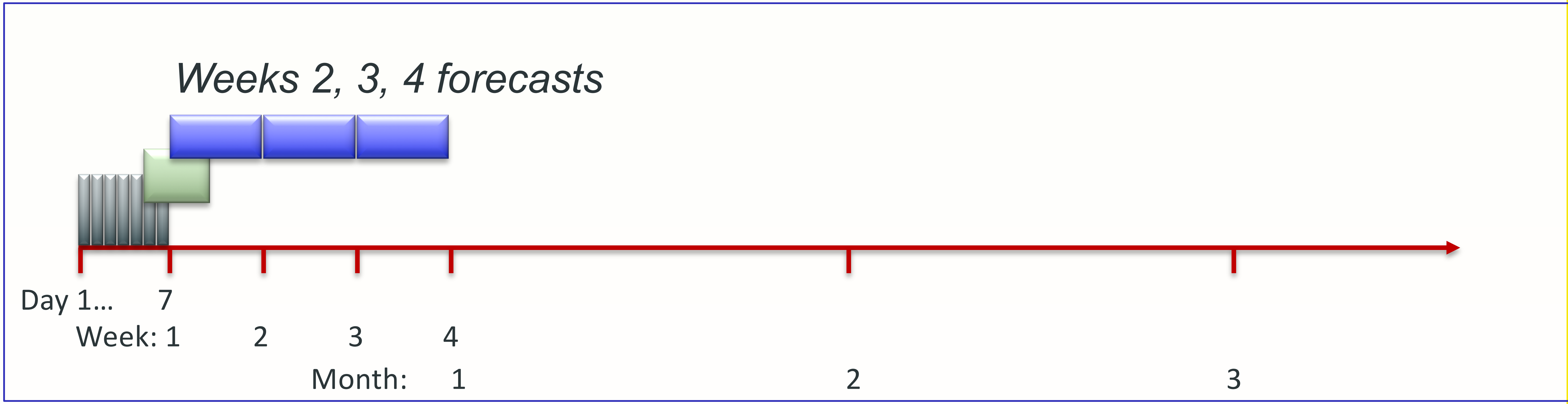
For example: NCEP

- The longer the forecast lead, the wider the "window".



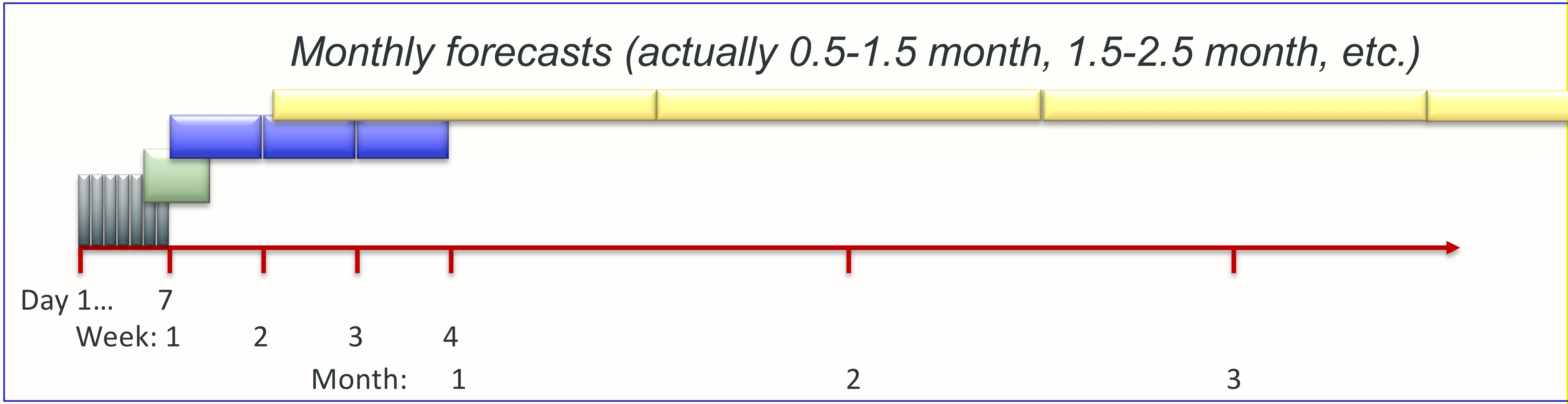
For example: NCEP

- The longer the forecast lead, the wider the "window".



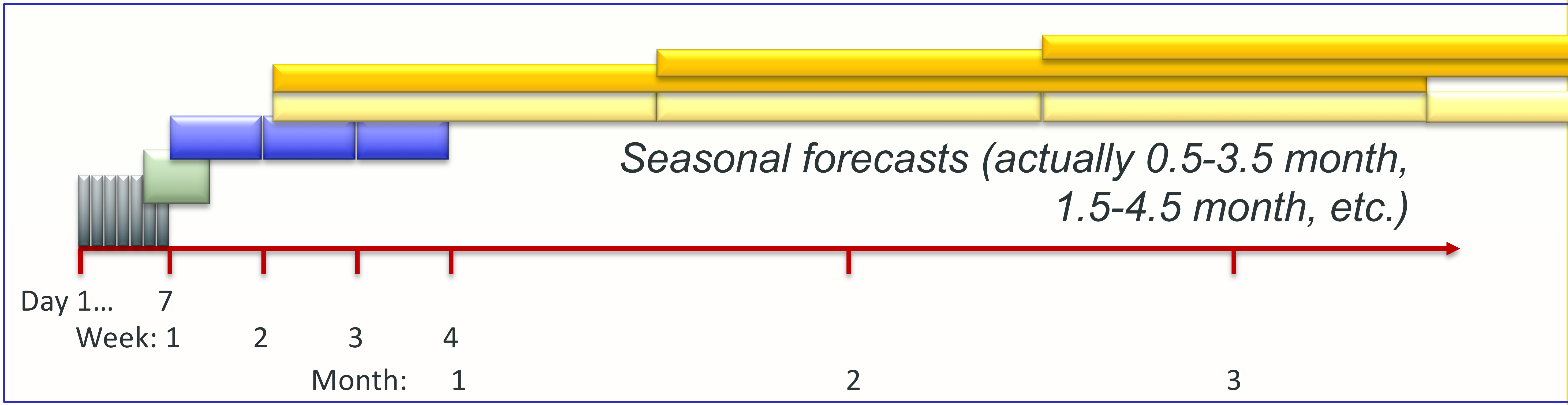
For example: NCEP

- The longer the forecast lead, the wider the "window".



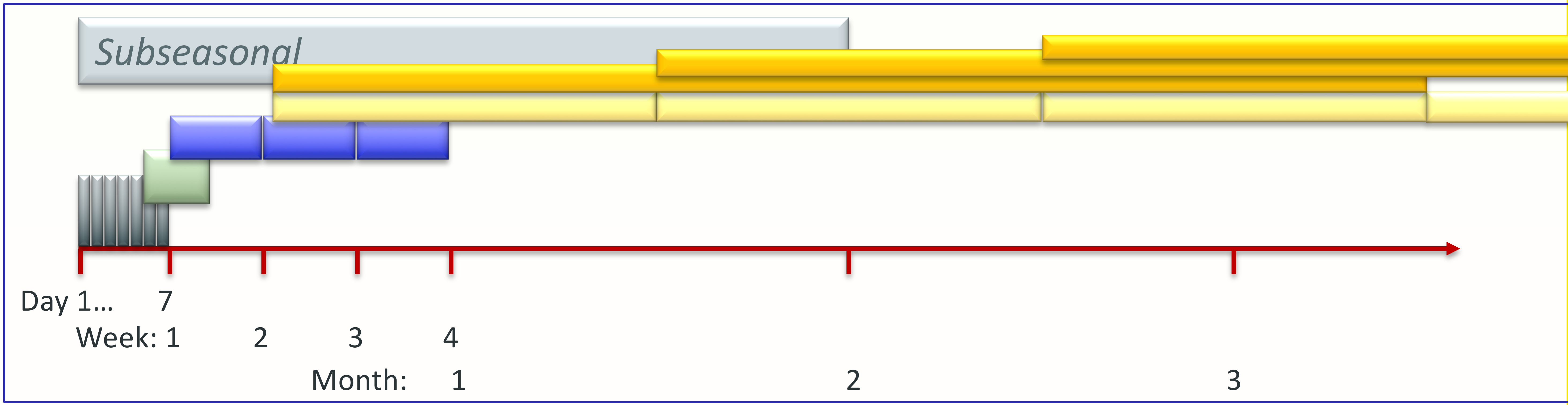
For example: NCEP

- The longer the forecast lead, the wider the "window".



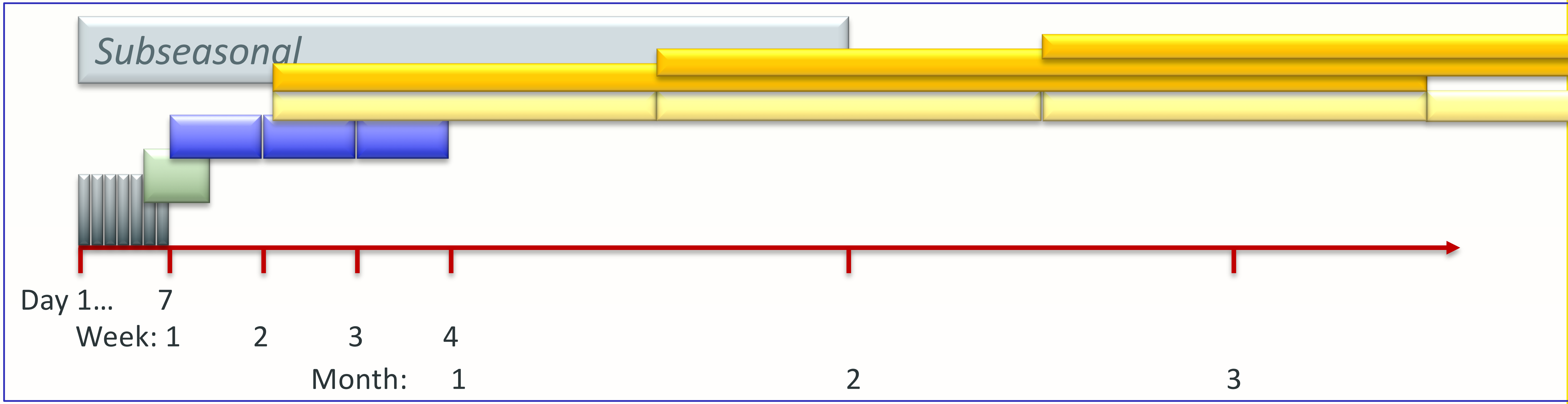
For example: NCEP

- The longer the forecast lead, the wider the "window".



For example: NCEP

- The longer the forecast lead, the wider the "window".



- This is intuitive – we measure error relative to the distance to the target.
- Begs the question: what constitutes a successful forecast?

Seamless Validation (Poisson Weighting)

- As forecast lead times get longer, precise timing is less of an issue for useful predictions.

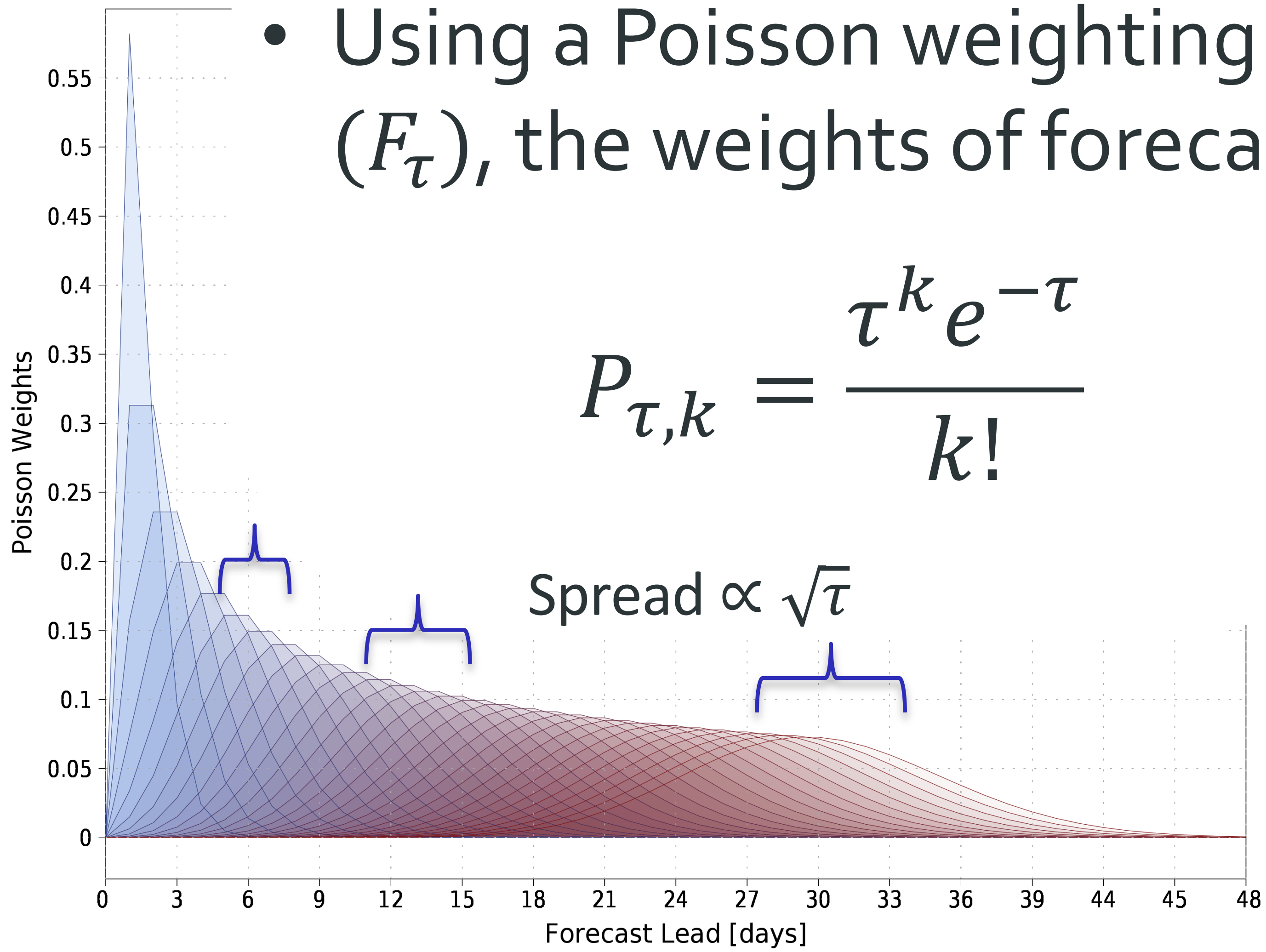
Seamless Validation (Poisson Weighting)

- As forecast lead times get longer, precise timing is less of an issue for useful predictions.

- Using a Poisson weighting in time for a forecast at lead τ: (F_τ), the weights of forecast daily means at each lead k (A_k):

$$P_{\tau,k} = \frac{\tau^k e^{-\tau}}{k!}$$

$$F_\tau = \frac{\sum_{k=1}^{N} A_k P_{\tau,k}}{\sum_{k=1}^{N} P_{\tau,k}}$$



Ford, T. W., P. A. Dirmeyer and D. O. Benson, 2018: *npj Climate Atmos. Sci.*, doi: 10.1038/s41612-018-0027-7.

Weakness of Poisson weighting approach

- The Poisson function has an advantageous shape and behavior for this approach to seamless forecasting, but it could be better.
  - Not flexible – its shape is what it is.
  - Even at 1-day lead, it is a blend of forecasts from several lead times, not a deterministic forecast only for day 1.

Weakness of Poisson weighting approach

- The Poisson function has an advantageous shape and behavior for this approach to seamless forecasting, but it could be better.
  – Not flexible – its shape is what it is.
  – Even at 1-day lead, it is a blend of forecasts from several lead times, not a deterministic forecast only for day 1.

- Would like to have an approach that transitions smoothly from day-to-day forecasts to a time-average (Poisson → Normal distribution).

Kronecker delta

- The area under the curve of the Poisson function always equals 1, making it ideal as a weighting function.
- The Kronecker delta:

$$\delta_{\tau,k} = \begin{cases} 0 & \tau \neq k \\ 1 & \tau = k \end{cases}$$

also has this characteristic, and is effectively the mathematical definition of the weight of a deterministic forecast.

Kronecker delta

- The area under the curve of the Poisson function always equals 1, making it ideal as a weighting function.
- The Kronecker delta:

$$\delta_{\tau,k} = \begin{cases} 0 & \tau \neq k \\ 1 & \tau = k \end{cases}$$

  also has this characteristic, and is effectively the mathematical definition of the weight of a deterministic forecast.
- Thus, a linear combination of Poisson and Kronecker weighting functions, with weights for each function that sum to 1, is well behaved as a versatile, compound weighting function.

Blending $P_{\tau,k}$ and $\delta_{\tau,k}$

- We want the deterministic forecast represented by $\delta_{\tau,k}$ to last for some number of days before transitioning to the $P_{\tau,k}$ weighting.
- The 2-parameter Hill equation is ideal for this:

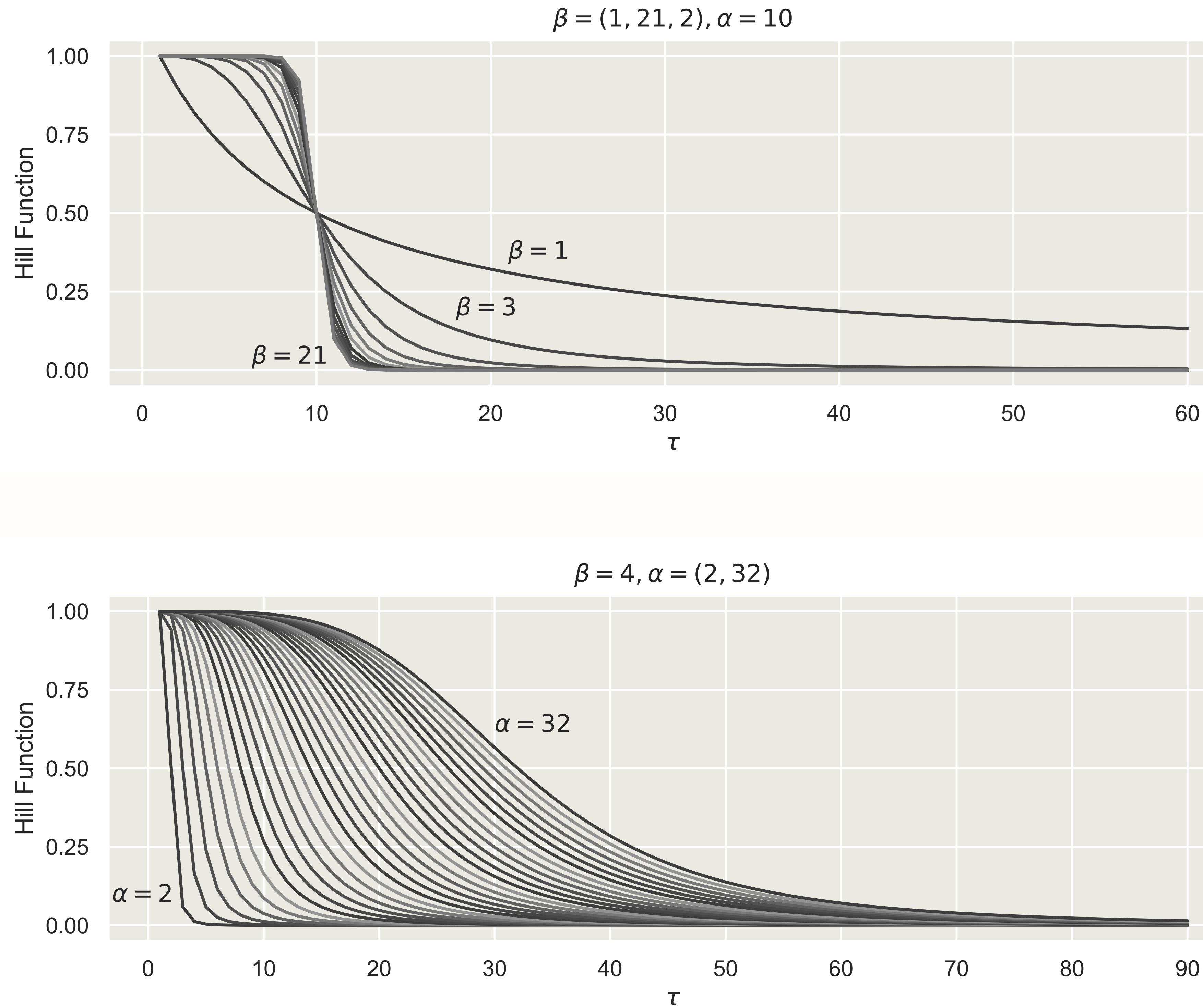
$$H_\tau = \frac{1}{\left(\frac{\tau - 1}{\alpha - 1}\right)^\beta + 1}$$

  - α determines the transition point (50/50 weighting between $P_{\tau,k}$ and $\delta_{\tau,k}$)
  - β gives the abruptness of the transition.
- The compound weight is: $W_{\tau,k} = H_\tau \delta_{\tau,k} + (1 - H_\tau) P_{\tau,k}$

Hill, A. V., 1910: *J. Physiol.*, doi:10.1113/jphysiol.1910.sp001386

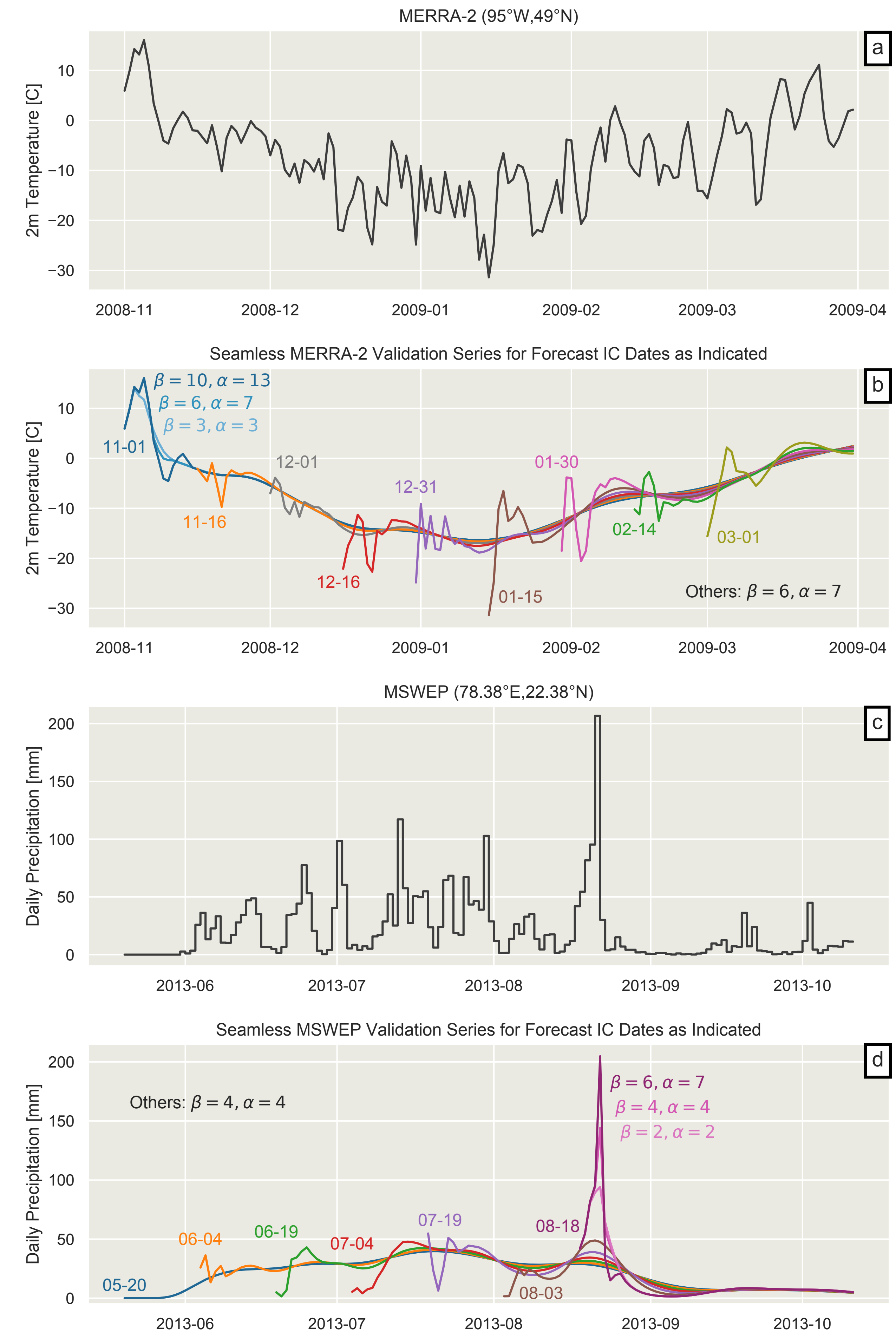
Hill equation

- Low β gives smooth transition with lead time, large β yields a sharp transition.
- α is the lead time at which the Hill equation has a value of 0.5. For any given β, larger α also produces a smoother transition.
- α and β can be chosen to give desired effect.



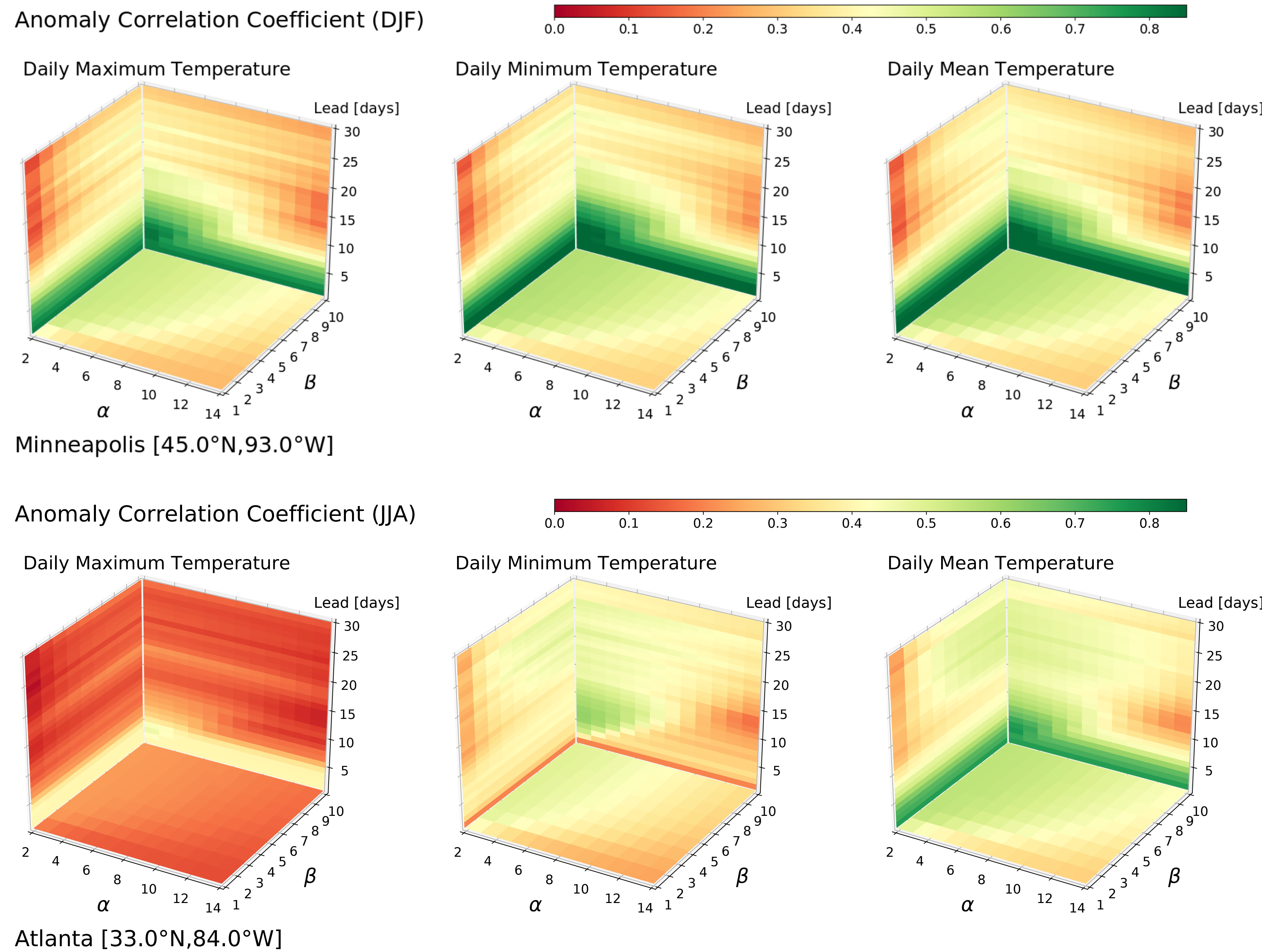
Seamless validation

- Validation becomes more complicated, as there are two time dimensions: validation time and lead (or initialization) time.
  - This should be done anyhow for forecasts with any models that drift (i.e., all dynamical models).
- MERRA-2 temperature (top) and MSWEP precipitation (bottom) examples show how validation time series smooth out with lead time.



Skill(α, β, τ)

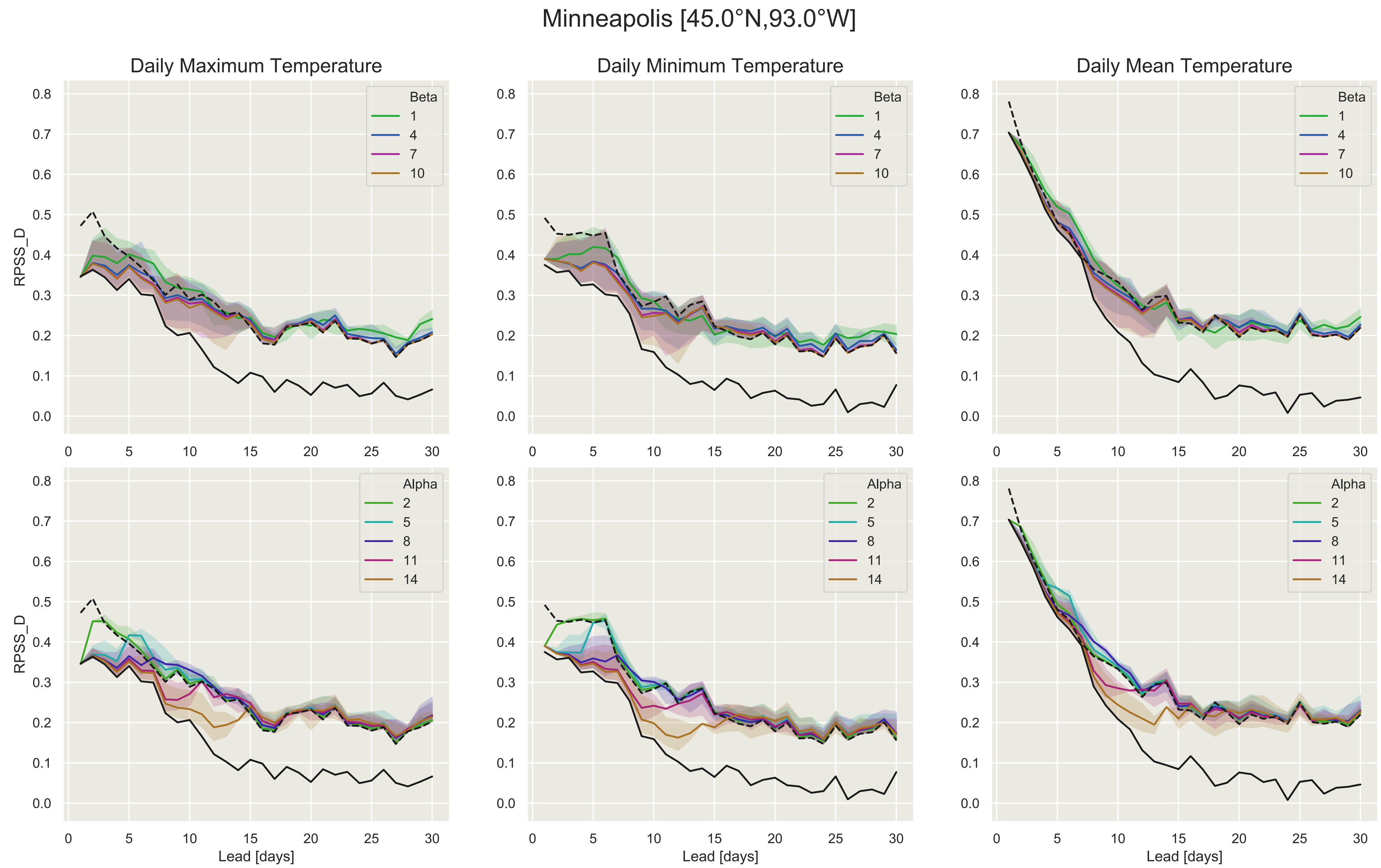
- Obviously skill tends to decrease with lead time, but the transition to time scales means that is not always the case.
- The choice of α and β that gives the best skill scores is not usually the most useful – the goal is not to make models look good.



CMA forecast skill (ACC) [1999-2010], color shows arithmetic mean along the perpendicular dimension.

Ensembles

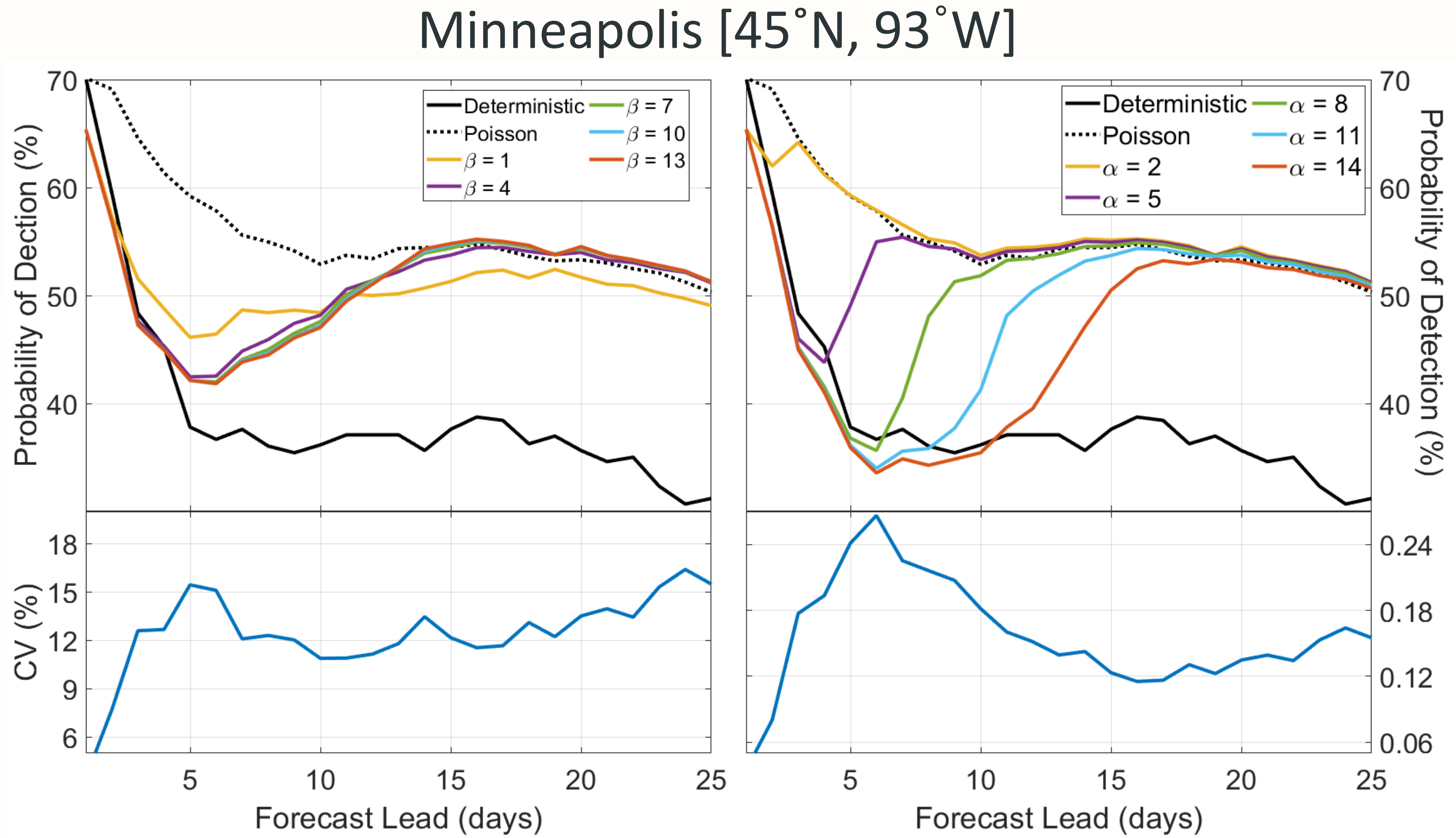
- Approach can be used for ensemble-based skill metrics as well.
- See most of the sensitivity to choices of the Hill equation parameters is in the 1-3 week range.



NCEP CFS $RPSS_D$ [1999-2010]; Black line = purely deterministic forecasts, dashed line = purely Poisson-weighted forecasts. Colored lines and spreads for different α and β are across all values of β and α respectively.

Heat waves

- Event-based statistics require a flexible means of definition for events that can vary with window.
- A forecast 12 days in advance for an event that occurs on day 11 or 13 should not be penalized – it is a useful forecast.



NCEP GEFS probability of detection (POD) [1999-2016]; Black line = purely deterministic forecasts, dashed line = purely Poisson-weighted forecasts.

Summary

- Seamless transition from daily to time-averaged forecasts.
- Applies to deterministic or ensemble forecasts.

Summary

- Seamless transition from daily to time-averaged forecasts.
- Applies to deterministic or ensemble forecasts.
- Highly tailorable for different applications, user interests.

Summary

- Seamless transition from daily to time-averaged forecasts.
- Applies to deterministic or ensemble forecasts.
- Highly tailorable for different applications, user interests.
- Linear combination of Kronecker and Poisson-weighted forecasts can be done a posteriori with various choices of α and β.

Summary

- Seamless transition from daily to time-averaged forecasts.
- Applies to deterministic or ensemble forecasts.
- Highly tailorable for different applications, user interests.
- Linear combination of Kronecker and Poisson-weighted forecasts can be done a posteriori with various choices of α and β.
- Requires a more complicated observed climatology to calculate anomalies: two time dimensions.

Summary

- Seamless transition from daily to time-averaged forecasts.
- Applies to deterministic or ensemble forecasts.
- Highly tailorable for different applications, user interests.
- Linear combination of Kronecker and Poisson-weighted forecasts can be done a posteriori with various choices of α and β.
- Requires a more complicated observed climatology to calculate anomalies: two time dimensions.
- Open question whether there is an objective approach to optimize the choice of the parameters.
- Large values of β reintroduce the "seam".

Dirmeyer & Ford, 2019: *Mon. Wea. Rev.*, submitted.