Online Learning with Optimism and Delay

Genevieve Flaspohler[1,2], Francesco Orabona[3], Judah Cohen[4],
Soukayna Mouatadid[5], Miruna Oprescu[7], Paulo Orenstein[6], and **Lester Mackey**[7]

MIT[1], Woods Hole Oceanographic Institution[2], Boston University[3],
Atmospheric and Environmental Research[4], University of Toronto[5], IMPA[6], Microsoft Research[7]

Optimistic Online Learning

Paradigm for sequential decision making

Each day $t = 1, \ldots, T$

1. Observe hint about future loss function (e.g., an estimate of ℓ_t)
2. Make decision $\mathbf{w}_t \in \mathbf{W}$
3. Suffer loss $\ell_t(\mathbf{w}_t)$
4. Use loss function ℓ_t to improve future decisions

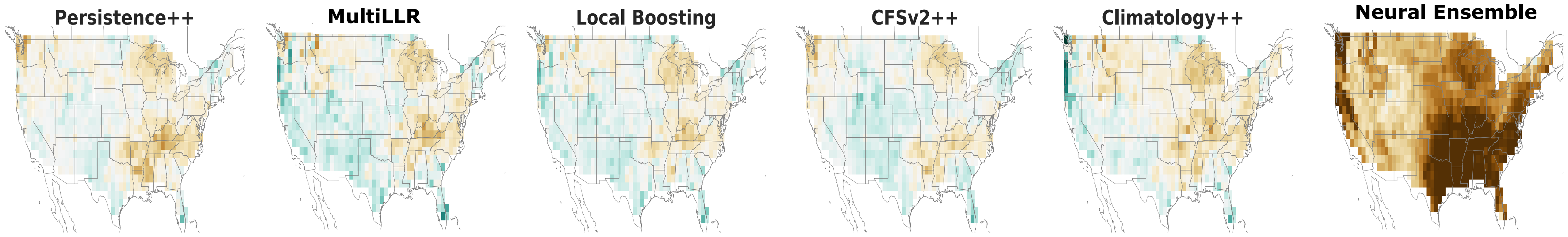
Goal: Perform nearly as well as best constant decision in hindsight

$$\mathrm{Regret}_T = \underbrace{\sum_{t=1}^{T} \ell_t(\mathbf{w}_t)}_{\text{Total loss of online learner}} - \underbrace{\inf_{\mathbf{u} \in \mathbf{W}} \sum_{t=1}^{T} \ell_t(\mathbf{u})}_{\text{Total loss of best constant decision}}$$

Optimistic Online Learning: Why?

Subseasonal climate forecasting

- Predicting temperature and precipitation 2-6 weeks in advance
- Forecasts issued daily, weekly, or semimonthly
- Diverse collection of forecasting models to choose from
- At least one model performs well each year (but unclear which a priori)



Optimistic Online Learning: Why?

Subseasonal climate forecasting

- Predicting temperature and precipitation 2-6 weeks in advance
- Forecasts issued daily, weekly, or semimonthly
- Diverse collection of forecasting models to choose from
- At least one model performs well each year (but unclear which a priori)

Goal: Each year, perform nearly as well as best single model in hindsight

$$\mathrm{Regret}_T = \underbrace{\sum_{t=1}^{T} \ell_t(\mathbf{w}_t)}_{\textit{Real-time forecasting loss}} - \underbrace{\inf_{\mathbf{u}\in\mathbf{W}} \sum_{t=1}^{T} \ell_t(\mathbf{u})}_{\textit{Total loss of best single model}}$$

Perfect fit?

Online Learning for Subseasonal Forecasting

Challenges

❌ Delayed feedback

Must issue multiple forecasts before observing feedback about the first

❌ Short regret horizons

Want small regret after only T=26 semimonthly forecasts

❌ Impractical hyperparameters

Standard settings based on worst-case future losses: challenging to implement / overly conservative

This talk: New algorithms with

✓ Optimal regret under delay

Even variable and unbounded delays

✓ Hints for missed feedback

Mitigate the impact of delay

✓ No hyperparameters!

✓ Learning to hint wrapper

Learn effective hinting strategies

Standard Online Learning Algorithms

Follow the Regularized Leader (FTRL) [Abernethy et al., 2008]

Sum of loss subgradients ($\mathbf{g}_t \in \partial \ell_t(\mathbf{w}_t)$)

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \langle \mathbf{g}_{1:t}, \mathbf{w} \rangle + \lambda \psi(\mathbf{w})$$

- Minimize sum of linearized losses + strongly convex regularizer (w.r.t. norm $\|\cdot\|$)

Online Mirror Descent (OMD) [Warmuth & Jagakota, 1997]

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \langle \mathbf{g}_t, \mathbf{w} \rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$$

Bregman divergence

$$\mathcal{B}_\psi(\mathbf{w}, \mathbf{u}) = \psi(\mathbf{w}) - \psi(\mathbf{u}) - \langle \nabla \psi(\mathbf{u}), \mathbf{w} - \mathbf{u} \rangle$$

- Minimize latest linearized loss while staying close to last decision point $\mathbf{w}_t$

Online Learning with Optimism

Optimistic FTRL (OFTRL) [Rakhlin & Sridharan, 2013]

Sum of loss subgradients ($\mathbf{g}_t \in \partial \ell_t(\mathbf{w}_t)$)

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \, \langle \mathbf{g}_{1:t} + \tilde{\mathbf{g}}_{t+1}, \mathbf{w} \rangle + \lambda \psi(\mathbf{w})$$

Hint vector: Estimate of future feedback $\mathbf{g}_{t+1}$

- **Benefit:** reduced regret whenever $\tilde{\mathbf{g}}_{t+1}$ approximates $\mathbf{g}_{t+1}$ well

Single-step Optimistic OMD (SOOMD) [Joulani et al., 2017]

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \, \langle \mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \mathbf{w} \rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$$

Bregman divergence

$$\mathcal{B}_{\psi}(\mathbf{w}, \mathbf{u}) = \psi(\mathbf{w}) - \psi(\mathbf{u}) - \langle \nabla\psi(\mathbf{u}), \mathbf{w} - \mathbf{u} \rangle$$

Online Learning with Delay

Paradigm for sequential decision making with delayed feedback

Each day $t = 1, \ldots, T$

1. Observe hint about future loss function (e.g., an estimate of ℓ_t)
2. Make decision $\mathbf{w}_t \in \mathbf{W}$
3. Observe delayed loss $\ell_{t-D}(\mathbf{w}_{t-D})$
4. Use delayed loss function ℓ_{t-D} to improve future decisions

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \langle \mathbf{g}_{1:t} + \tilde{\mathbf{g}}_{t+1}, \mathbf{w} \rangle + \lambda \psi(\mathbf{w}) \quad \text{(OFTRL)}$$

Problem:

Unobservable

$$\mathbf{w}_{t+1} = \underset{\mathbf{w} \in \mathbf{W}}{\operatorname{argmin}} \langle \mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \mathbf{w} \rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t) \quad \text{(SOOMD)}$$

Online Learning with Optimism and Delay

Optimistic Delayed FTRL (ODFTRL) [This work]

Sum of subgradients **observed so far**

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_{1:t-D} + \mathbf{h}_{t+1}, \mathbf{w}\rangle + \lambda\psi(\mathbf{w})$$

Hint vector: Estimate of future **and missed** feedback $\mathbf{g}_{t-D+1:t+1}$

Delayed Optimistic OMD (DOOMD) [This work]

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t, \mathbf{w}\rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$$

Last observed subgradient

Delay as Optimism

Learning with delay is a special case of learning with optimism.

Lemma [This work]

DOOMD $\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t, \mathbf{w}\rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$ is

SOOMD $\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_t + \tilde{\mathbf{g}}_{t+1} - \tilde{\mathbf{g}}_t, \mathbf{w}\rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_t)$

with a really bad hint $\tilde{\mathbf{g}}_{t+1} = -\mathbf{g}_{t-D+1:t} + \mathbf{h}_{t+1}$

- Key property: SOOMD depends on $\mathbf{g}_t$ and $\tilde{\mathbf{g}}_{t+1}$ only through $\mathbf{g}_{1:t} + \tilde{\mathbf{g}}_{t+1}$
- Not satisfied by more common (two-step) optimistic OMD algorithms

[Chiang et al., 2012; Rakhlin & Sridharan, 2013a;b; Kamalaruban, 2016]

$$\mathbf{w}_{t+1/2} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \mathbf{g}_t, \mathbf{w}\rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_{t-1/2}) \text{ and } \mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathbf{W}} \langle \tilde{\mathbf{g}}_{t+1}, \mathbf{w}\rangle + \mathcal{B}_{\lambda\psi}(\mathbf{w}, \mathbf{w}_{t+1/2})$$

Unobserved

Delay as Optimism

Any regret bound for optimistic learning immediately implies a regret bound for delayed learning.

New guarantee for optimistic online learning

Theorem 3 (OFTRL regret). *If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the* OFTRL *iterates $\mathbf{w}_t$ satisfy*

$$\text{Regret}_T(\mathbf{u}) \leq \lambda\psi(\mathbf{u}) + \tfrac{1}{\lambda}\sum_{t=1}^T \text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*).$$

$$\text{huber}(\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*, \|\mathbf{g}_t\|_*) = \begin{cases} 0 & \text{for perfect hints } \tilde{\mathbf{g}}_t = \mathbf{g}_t \\ \frac{1}{2}\|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*^2 & \text{for small hint errors } \|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* \\ \|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_*\|\mathbf{g}_t\|_* - \frac{1}{2}\|\mathbf{g}_t\|_*^2 & \text{for large hint errors } \|\mathbf{g}_t - \tilde{\mathbf{g}}_t\|_* \end{cases}$$

- Strictly improves past OFTRL guarantees [Rakhlin & Sridharan (2013a); Mohri & Yang (2016); Orabona (2019, Thm. 7.28); Joulani et al. (2017, Sec. 7.2)]
- Demonstrates robustness to inaccurate hints
  - Same holds true for SOOMD (see our write-up)

Delay as Optimism

Any regret bound for optimistic learning immediately implies a regret bound for delayed learning.

First general analysis of delayed FTRL (see Hsieh et al. [2020] for concurrent work)

> **Theorem 5** (ODFTRL regret). *If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the* ODFTRL *iterates $\mathbf{w}_t$ satisfy*
>
> $$\mathrm{Regret}_T(\mathbf{u}) \leq \lambda\psi(\mathbf{u}) + \tfrac{1}{\lambda}\textstyle\sum_{t=1}^{T} \mathbf{b}_{t,F} \quad \textit{for}$$
> $$\mathbf{b}_{t,F} \triangleq \mathrm{huber}(\|\mathbf{h}_t - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*).$$

- Compounding of regret due to delay
  - Best λ yields $\mathcal{O}(\sqrt{(D+1)T})$ regret, rate optimal in worst case [Weinberger & Ordentlich, 2002]
- Heightened value of optimism
  - Can mitigate delay by hinting at both missed and future subgradients $\mathbf{g}_{t-D:t}$
- Strengthens analyses of special cases: $\|\mathbf{h}_t - \sum_{s=t-D}^{t-1} \mathbf{g}_s\|_* \ll \sum_{s=t-D}^{t-1} \|\mathbf{g}_s\|_*$

  [Hsieh et al., 2020; Quanrud & Khashabi 2015; Korotin et al., 2020]
  - McMahan & Streeter [2014]: similar bound for unoptimistic scalar gradient descent

Tuning Regularizers with Optimism and Delay

Theorem 5 (ODFTRL regret). *If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the ODFTRL iterates $\mathbf{w}_t$ satisfy*

$$\mathrm{Regret}_T(\mathbf{u}) \leq \lambda\psi(\mathbf{u}) + \tfrac{1}{\lambda}\textstyle\sum_{t=1}^{T} \mathbf{b}_{t,F} \quad \textit{for}$$

$$\mathbf{b}_{t,F} \triangleq \mathrm{huber}(\|\mathbf{h}_t - \textstyle\sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*).$$

- Issue: How do we pick the regularization parameter λ in practice?
- Ideal: $\lambda = \sqrt{\dfrac{\sum_{t=1}^{T} \mathbf{b}_{t,F}}{\sup_{\mathbf{u}\in\mathbf{U}} \psi(\mathbf{u})}}$ minimizes regret bound but is unobservable
- Two practical strategies
  1. Tuning-free algorithms (DORM and DORM+): independent of λ, optimally tuned!
  2. Self-tuning strategy (AdaHedgeD): adaptively sets λ near-optimally

Regret Matching and Regret Matching+

Regret Matching (RM) [Blackwell, 1956]

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \quad \mathbf{r}_t \triangleq \mathbf{1}\langle \mathbf{g}_t, \mathbf{w}_t\rangle - \mathbf{g}_t,$$
$$\tilde{\mathbf{w}}_{t+1} \triangleq \max(\mathbf{0}, \mathbf{r}_{1:t}/\lambda)$$

Regret Matching+ (RM+) [Tammelin et al., 2015]

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \quad \mathbf{r}_t \triangleq \mathbf{1}\langle \mathbf{g}_t, \mathbf{w}_t\rangle - \mathbf{g}_t,$$
$$\tilde{\mathbf{w}}_{t+1} \triangleq \max\left(\mathbf{0}, \tilde{\mathbf{w}}_t + \mathbf{r}_t/\lambda\right)$$

- RM developed for finding correlated equilibria in two-player games
- RM+ solved Heads-up Limit Texas Hold’em poker [Bowling et al., 2015]
- Each $\mathbf{w}_t \in \mathbf{W} = \triangle_{d-1}$ represents a convex combination of input model forecasts
- Do not account for delay or optimism
- Regret guarantees are suboptimal for large d [Cesa-Bianchi & Lugosi, 2006; Orabona & Pal, 2015]

Delayed Optimistic Regret Matching (+)

Delayed Optimistic Regret Matching (DORM) [This work]

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \quad \mathbf{r}_{t-D} \triangleq \mathbf{1}\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D}\rangle - \mathbf{g}_{t-D},$$

$$\tilde{\mathbf{w}}_{t+1} \triangleq \max(\mathbf{0}, (\mathbf{r}_{1:t-D} + \mathbf{h}_{t+1})/\lambda)^{q-1}$$

Delayed Optimistic Regret Matching+ (DORM+) [This work]

$$\mathbf{w}_{t+1} = \tilde{\mathbf{w}}_{t+1}/\langle \mathbf{1}, \tilde{\mathbf{w}}_{t+1}\rangle \quad \text{for} \quad \mathbf{r}_{t-D} \triangleq \mathbf{1}\langle \mathbf{g}_{t-D}, \mathbf{w}_{t-D}\rangle - \mathbf{g}_{t-D},$$

$$\tilde{\mathbf{w}}_{t+1} \triangleq \max\left(\mathbf{0}, \tilde{\mathbf{w}}_t^{p-1} + (\mathbf{r}_{t-D} + \mathbf{h}_{t+1} - \mathbf{h}_t)/\lambda\right)^{q-1}, \quad \text{and} \quad p = \frac{q}{q-1}$$

- Each $\mathbf{w}_t \in \mathbf{W} = \triangle_{d-1}$ represents a convex combination of input model forecasts
- Will choose $q \geq 2$ to obtain optimal dependence of regret on dimension d
- Generalize
    - RM [Blackwell, 1956] and RM+ [Tammelin et al., 2015]
    - Undelayed optimistic RM with q = 2 independently developed by Farina et al. [2021]

Delayed Optimistic Regret Matching (+)

Lemma 1. *The* DORM *and* DORM+ *iterates* $\mathbf{w}_t$ *are*

1. *Proportional to* ODFTRL *and* DOOMD *iterates with* $\mathbf{W} \triangleq \mathbb{R}^d_+$, $\psi(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_p^2$, *and surrogate loss* $\hat{\ell}_t(\mathbf{w}) = \langle \mathbf{w}, -\mathbf{r}_t \rangle$.
2. *Independent of the choice of* $\lambda \Rightarrow$ *Automatically optimally tuned*

Corollary 1. *For all* $\mathbf{u} \in \triangle_{d-1}$, DORM *satisfies*

$$\text{Regret}_T(\mathbf{u}) \leq \inf_{\lambda>0} \frac{\lambda}{2}\|\mathbf{u}\|_p^2 + \frac{1}{\lambda(p-1)} \sum_{t=1}^T \mathbf{b}_{t,q} = \sqrt{\frac{\|\mathbf{u}\|_p^2}{2(p-1)} \sum_{t=1}^T \mathbf{b}_{t,q}}$$

for $\mathbf{b}_{t,q} = \text{huber}(\|\mathbf{h}_t - \sum_{s=t-D}^t \mathbf{r}_s\|_q, \|\mathbf{r}_t\|_q)$.

Delayed Optimistic Regret Matching (+)

Lemma 1. *The* DORM *and* DORM+ *iterates* $\mathbf{w}_t$ *are*

1. *Proportional to* ODFTRL *and* DOOMD *iterates with* $\mathbf{W} \triangleq \mathbb{R}_+^d$, $\psi(\mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|_p^2$, *and surrogate loss* $\hat{\ell}_t(\mathbf{w}) = \langle \mathbf{w}, -\mathbf{r}_t \rangle$.
2. *Independent of the choice of* $\lambda \Rightarrow$ *Automatically optimally tuned*

Corollary 1. *For all* $\mathbf{u} \in \triangle_{d-1}$, *if* DORM $q = \operatorname{argmin}_{r \geq 2} d^{2/r}(r-1)$ *then* $\mathrm{Regret}_T(\mathbf{u}) \leq \sqrt{(2\log_2(d) - 1)\sum_{t=1}^T \mathbf{b}_{t,\infty}}$.

Optimal dimension dependence!

[Cesa-Bianchi & Lugosi, 2006]

Adaptive Learning with Optimism and Delay

Optimistic Delayed Adaptive FTRL (ODAFTRL) [This work]

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w} \in \mathbf{W}} \langle \mathbf{g}_{1:t-D} + \mathbf{h}_{t+1}, \mathbf{w} \rangle + \lambda_{t+1} \psi(\mathbf{w})$$

Time-varying regularization strength

Theorem 1 (ODAFTRL regret). *If ψ is nonnegative and λ_t is non-decreasing in t, then, $\forall \mathbf{u} \in \mathbf{W}$, the* ODAFTRL *iterates $\mathbf{w}_t$ satisfy*

$$\mathrm{Regret}_T(\mathbf{u}) \leq \lambda_T \psi(\mathbf{u}) + \sum_{t=1}^{T} \min\left(\frac{\mathbf{b}_{t,F}}{\lambda_t}, \mathbf{a}_{t,F}\right) \quad \textit{with}$$

$$\mathbf{b}_{t,F} \triangleq \mathrm{huber}(\|\mathbf{h}_t - \sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*) \quad \textit{and}$$

$$\mathbf{a}_{t,F} \triangleq \mathrm{diam}(\mathbf{W}) \min\left(\|\mathbf{h}_t - \sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*\right).$$

- Delay mitigation from accurate hints + robustness to hinting error
- Improves undelayed bounds of [Rakhlin & Sridharan, 2013a; Mohri & Yang, 2016; Joulani et al., 2017] and concurrent unoptimistic bound of Hsieh et al. [2020]
- Bounded-domain factors $\mathbf{a}_{t,F}$ enable practical λ_t -tuning strategies under delay without any prior knowledge of unobserved subgradients

Adaptive Tuning with Optimism and Delay

Theorem 1 (ODAFTRL regret). *If ψ is nonnegative and λ_t is non-decreasing in t, then, $\forall \mathbf{u} \in \mathbf{W}$, the* ODAFTRL *iterates $\mathbf{w}_t$ satisfy*

$$\mathrm{Regret}_T(\mathbf{u}) \leq \lambda_T \psi(\mathbf{u}) + \sum_{t=1}^{T} \min\left(\frac{\mathbf{b}_{t,F}}{\lambda_t}, \mathbf{a}_{t,F}\right) \quad \textit{with}$$

$$\mathbf{b}_{t,F} \triangleq \mathrm{huber}\left(\|\mathbf{h}_t - \sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*\right) \quad \textit{and}$$

$$\mathbf{a}_{t,F} \triangleq \mathrm{diam}(\mathbf{W}) \min\left(\|\mathbf{h}_t - \sum_{s=t-D}^{t} \mathbf{g}_s\|_*, \|\mathbf{g}_t\|_*\right).$$

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w} \in \mathbf{W}} \langle \mathbf{g}_{1:t-D} + \mathbf{h}_{t+1}, \mathbf{w} \rangle + \lambda_{t+1} \psi(\mathbf{w})$$

- **Issue:** How do we pick the regularization sequence λ_t ?
- **Ideal:** $\lambda_t = \sqrt{\frac{\sum_{s=1}^{t} \mathbf{b}_{s,F}}{\sup_{\mathbf{u} \in \mathbf{U}} \psi(\mathbf{u})}}$ nearly optimizes regret bound but unobservable
- **Standard approach** [Joulani et al., 2016; McMahan & Streeter, 2014; Hsieh et al., 2020]
  - Uniformly upper bound unobserved $\mathbf{b}_{s,F}$ terms
  - Requires bound on any subgradient norm that could arise: impractical or very loose!
- **Our approach:** Set λ_t based on tighter regret bound underlying theorem

AdaHedgeD

Theorem 1 (AdaHedgeD regret). *For $\alpha > 0$, consider the* AdaHedgeD *sequence*

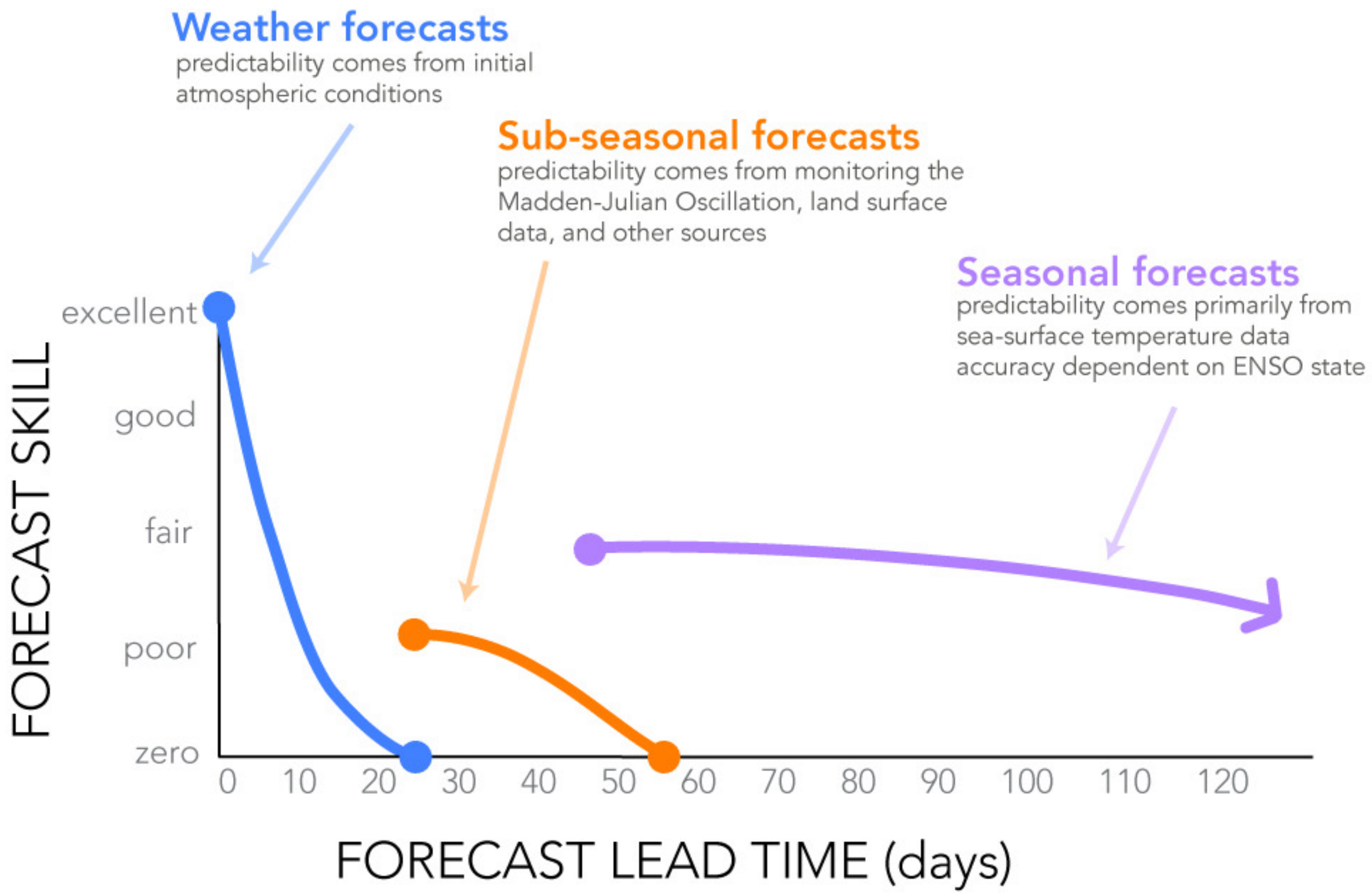
$$\lambda_{t+1} = \frac{1}{\alpha}\sum_{s=1}^{t-D}\delta_s \qquad \textit{for} \qquad \delta_t \triangleq B(\mathbf{w}_t, \lambda_t, \mathbf{g}_{1:t}, \mathbf{h}_t).$$

If ψ is nonnegative, then, for all $\mathbf{u} \in \mathbf{W}$, the ODAFTRL *iterates satisfy*

$$\mathrm{Regret}_T(\mathbf{u}) \le \left(\frac{\psi(\mathbf{u})}{\alpha} + 1\right)\left(2\max_{t\in[T]}\mathbf{a}_{t-D:t-1,F} + \sqrt{\sum_{t=1}^{T}\mathbf{a}_{t,F}^2 + 2\alpha\mathbf{b}_{t,F}}\right).$$

- Rate-optimal $\mathcal{O}(\sqrt{(D+1)T} + D)$ delay dependence in the worst case
- Nearly matches optimally tuned regret bound in hindsight
- No prior knowledge of future subgradients required
- Generalizes popular AdaHedge algorithm [Erven et al., 2011] by incorporating delay, optimism, and tighter regret bounds to mitigate impact of delay

Subseasonal Climate Forecasting



Source: https://iri.columbia.edu/news/qa-subseasonal-prediction-project/

Subseasonal Forecasting: What and Why?

- **What:** Predicting temperature and precipitation 2 – 6 weeks out
- **Why:** (White et al., 2017, Meteorological Applications)
  - Allocating water resources
  - Managing wildfires
  - Preparing for weather extremes
    - e.g., droughts, heavy rainfall, and flooding
  - Crop planting, irrigation scheduling, and fertilizer application
  - Energy pricing

Subseasonal Forecasting: What and Why?

- **What:** Predicting temperature and precipitation 2 – 6 weeks out
- **Why:** (White et al., 2017, Meteorological Applications)
  - Allocating water resources
  - Managing wildfires
  - Preparing for weather extremes
    - e.g., droughts, heavy rainfall, and flooding
  - Crop planting, irrigation scheduling, and fertilizer application
  - Energy pricing

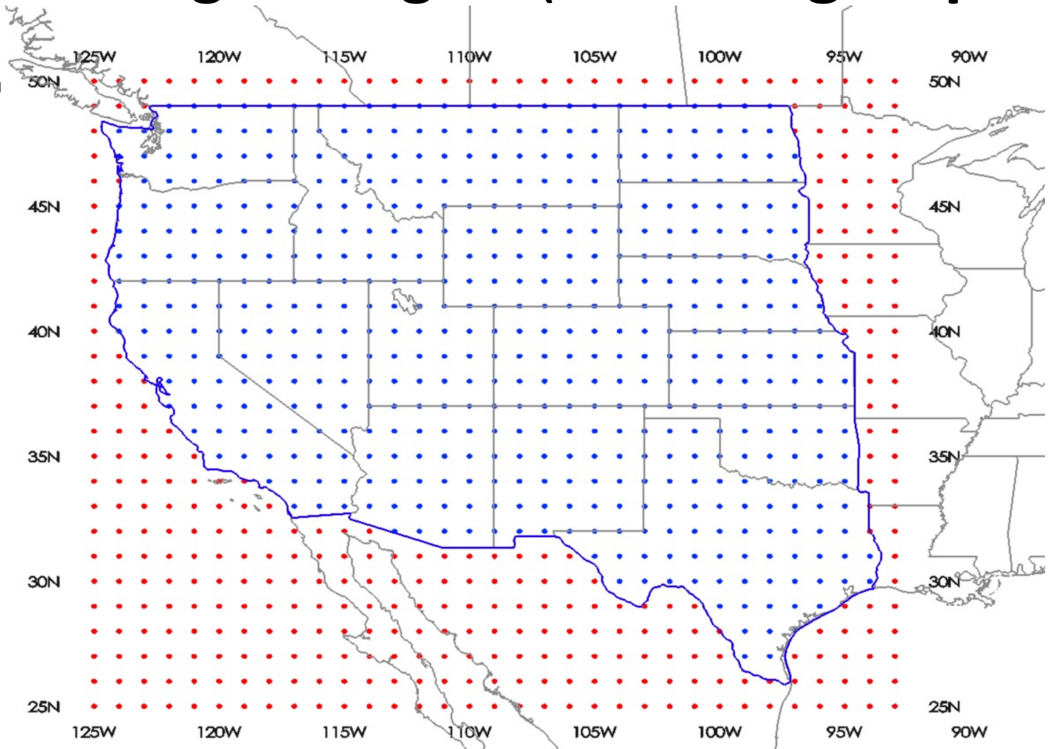
U.S. Bureau of Reclamation

- “The mission of the [USBR] is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.”
- Manages water in 17 western states
  - Provides 1 out of 5 Western farmers with irrigation water for 10 million farmland acres
  - Generates enough electricity to power 3.5M U.S. homes
- “During the past eight years, every state in the Western United States has experienced drought that has affected the economy both locally and nationally through impacts to agricultural production, water supply, and energy.”

Credit: David Raff, USBR

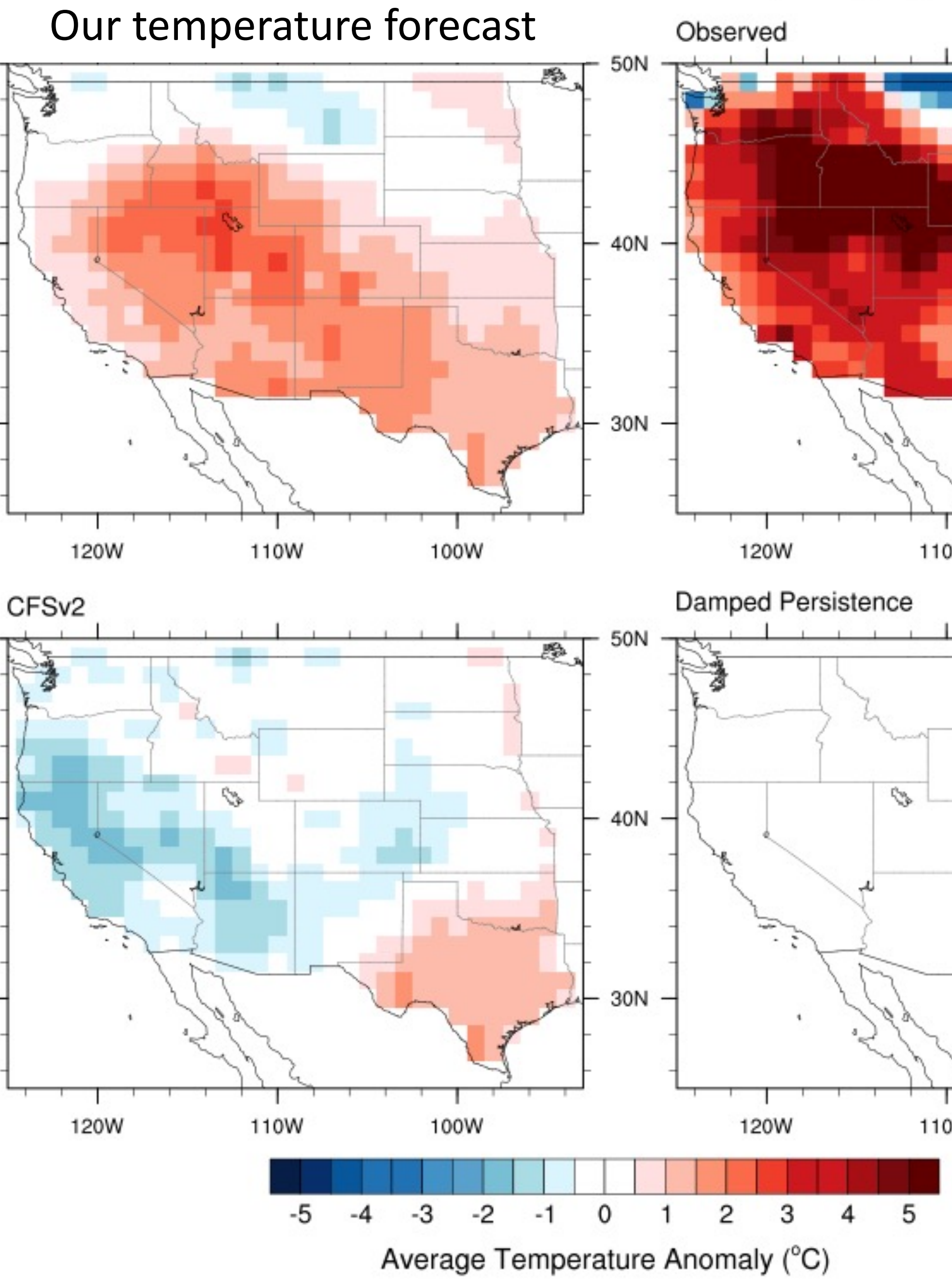
Subseasonal Forecasts

- Four separate forecasting tasks
  - Two variables: **average temperature** (degrees C) and **total precipitation** (mm)
  - Two outlooks: **weeks 3-4** and **weeks 5-6** (forecast is over a 2-week period)
- Issued on a 1°×1° latitude-longitude grid (**G = 514 grid points**)
- Issued **every two weeks** over a year (**T = 26**)

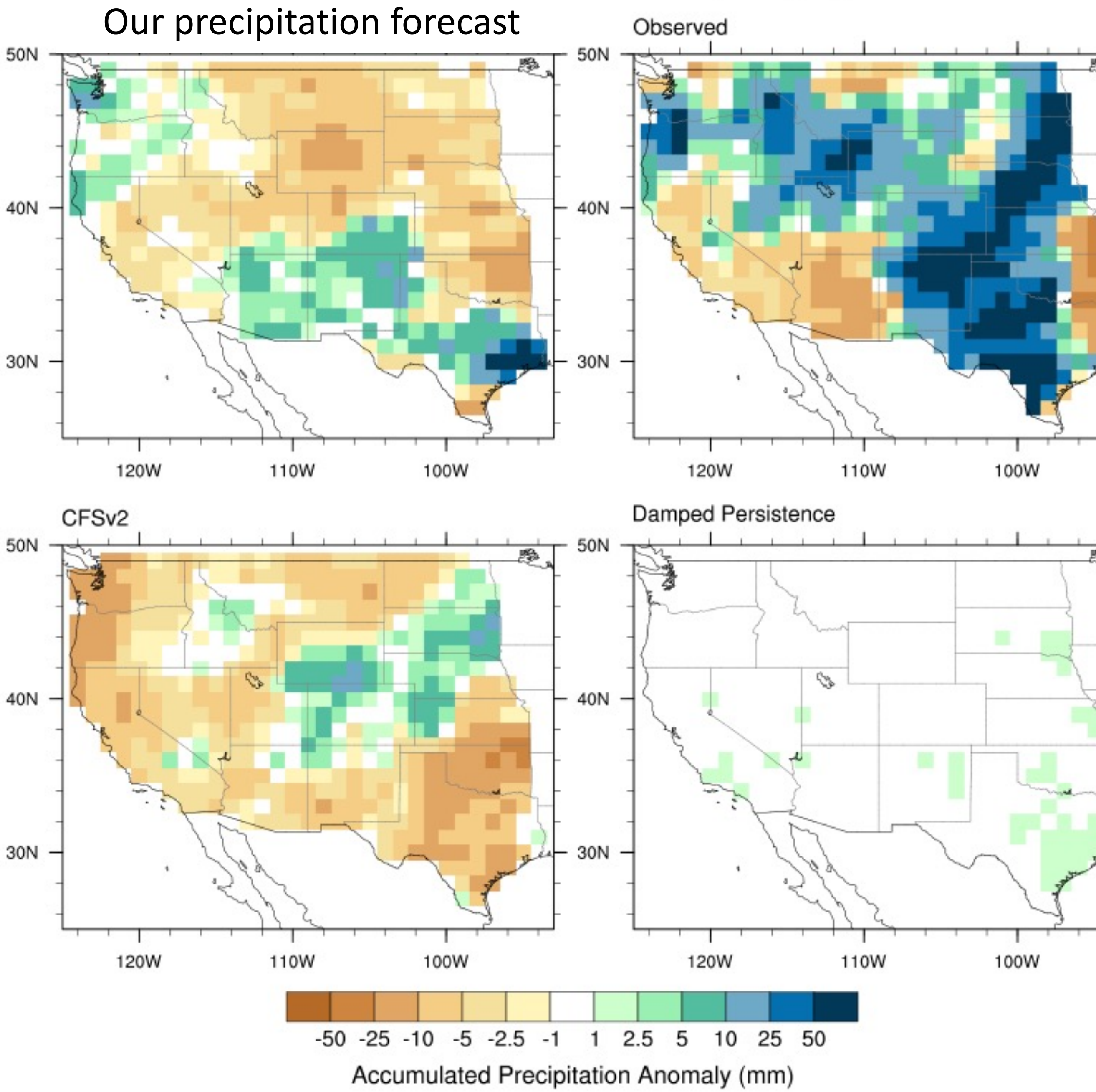


17 State Forecast Rodeo Region

Week 3-4 Forecast submitted 20180109, verifying 20180205



Week 3-4 Forecast submitted 20170905, verifying 20171002



Subseasonal Evaluation

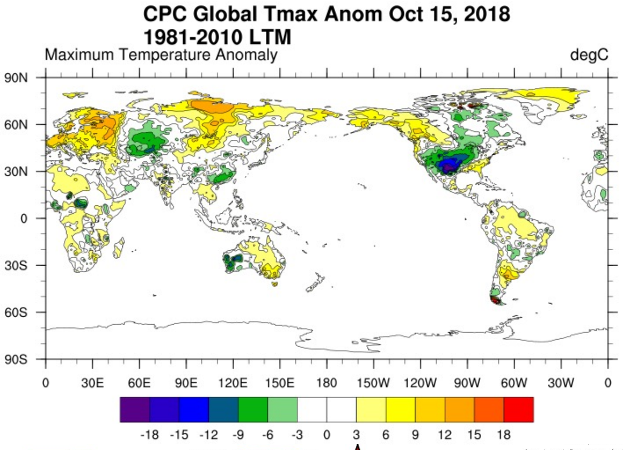
For each 2-week period starting on date t, forecasts judged on geographic **root mean squared error (RMSE)** between observed and predicted vectors of temperature or precipitation $\boldsymbol{y}_t$ and $\hat{\boldsymbol{y}}_t \in \mathbb{R}^G$

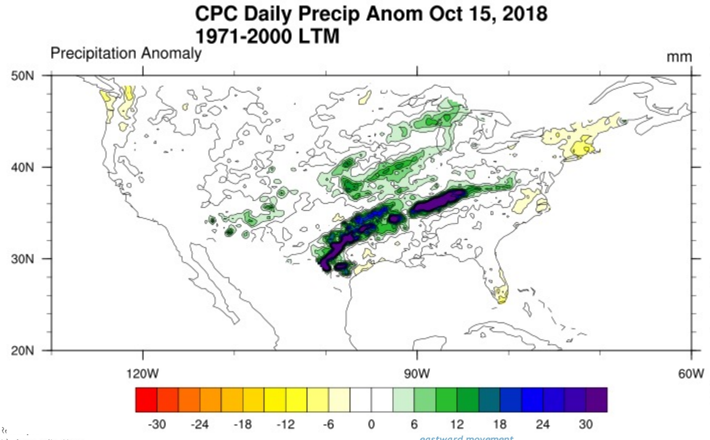
$$\mathrm{rmse}(\hat{\boldsymbol{y}}_t, \boldsymbol{y}_t) = \sqrt{\tfrac{1}{G}\sum_{g=1}^{G}(\hat{y}_{t,g} - y_{t,g})^2}$$

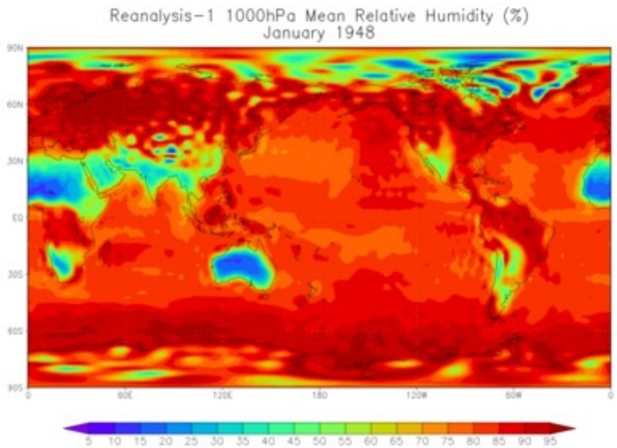
- **Multitask** objective function: couples together the G per-grid point forecasting tasks

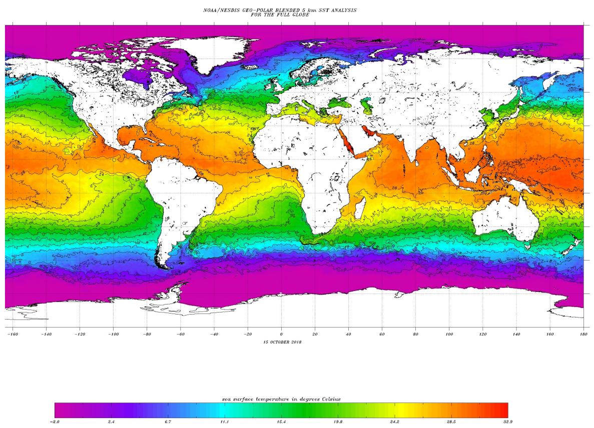
Our SubseasonalRodeo Dataset

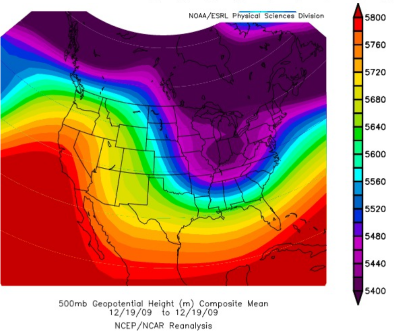
- To train and evaluate our models, we constructed a **SubseasonalRodeo dataset** from diverse data sources
- Released via the Harvard Dataverse https://doi.org/10.7910/DVN/IHBANG

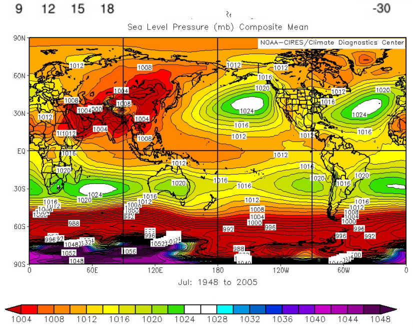


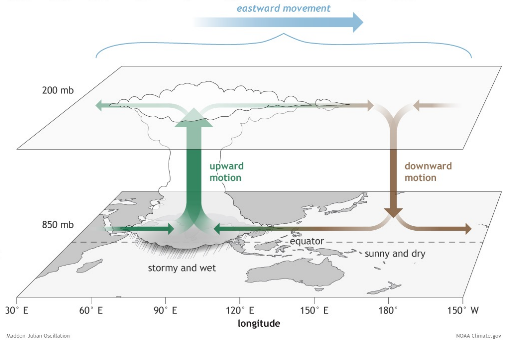


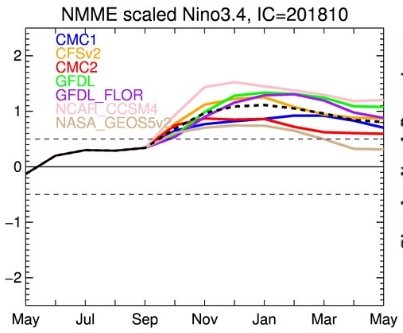


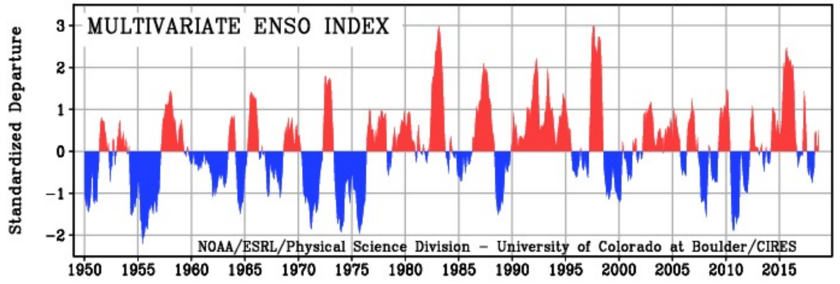












Our SubseasonalRodeo Dataset

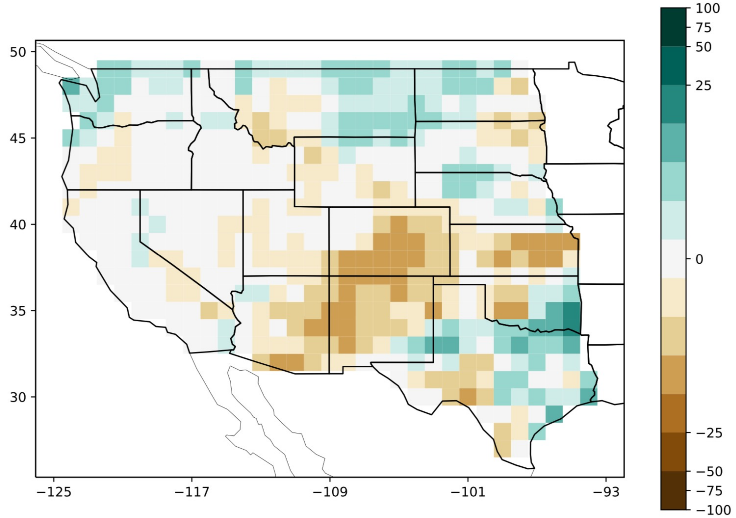
- To train and evaluate our models, we constructed a **SubseasonalRodeo dataset** from diverse data sources
- Released via the Harvard Dataverse https://doi.org/10.7910/DVN/IHBANG
- Organized as a collection of Python Pandas objects in HDF5 format
  - Spatial variables (vary with the target grid point but not the target date)
  - Temporal variables (vary with the target date but not the target grid point)
  - Spatiotemporal variables (vary with both the target grid point and the target date)
- Gridded data interpolated to $1° \times 1°$ grid (using distance-weighted average interpolation) and restricted to contest grid points
- Daily measurements replaced with averages (or, for precipitation, sums) over ensuing 2-week period

A Few of Our Forecasting Models

Climatology++

Predict mean or geometric median observation in adaptively selected window around target date.

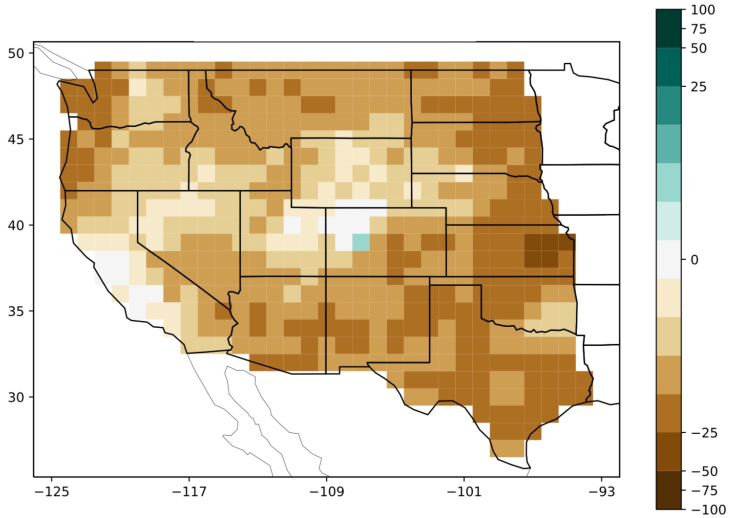
Best in 2016, 2017



Neural Ensemble

Neural network ensemble using sea surface temperature features.

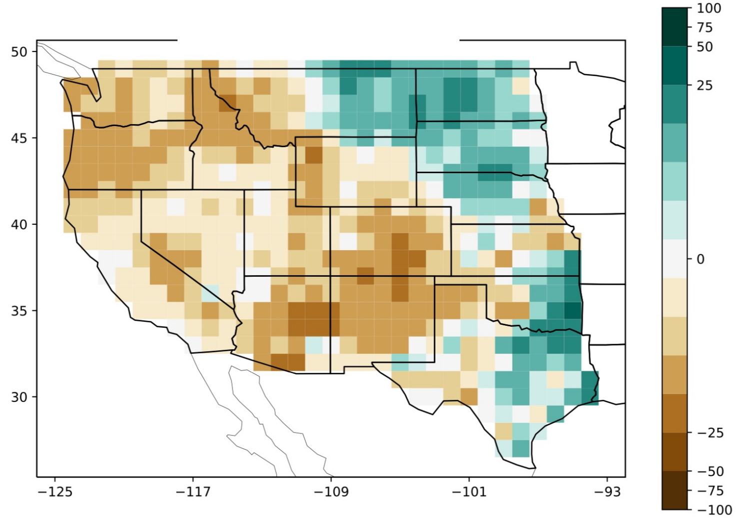
Best in 2012



CFSv2++

Learned correction for operational physics-based numerical weather prediction.

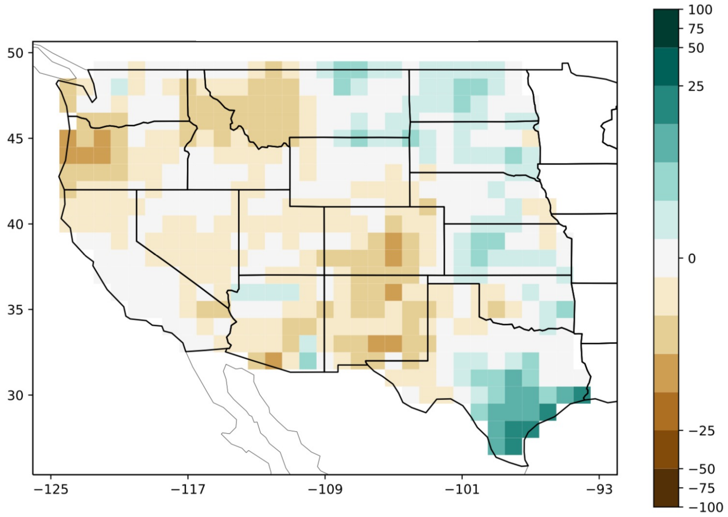
Best in 2011, 2013, 2019, 2020



Persistence++

Learned combination of physics-based forecasts, lagged measurements, and climatology.

Best in 2014, 2015, 2018



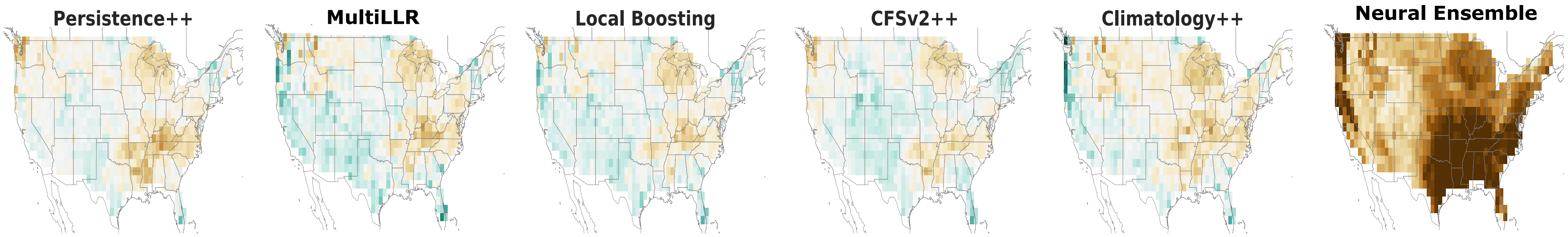
Question: How do we choose a single forecast to issue each day?

Online Learning for Subseasonal Forecasting

Answer: Online learning (with optimism and delay)!

- Select weights $\mathbf{w}_t \in \Delta$ to predict a convex combination of input forecasts

$$\hat{\boldsymbol{Y}}_t =$$



- **Goal:** Perform nearly as well as best model each year in 2011-2020
- **Loss:** $\ell_t(\mathbf{w}) = \mathrm{rmse}(\boldsymbol{y}_t, \hat{\boldsymbol{Y}}_t \mathbf{w})$
- **Algorithms:** DORM, DORM+, AdaHedgeD

Hinting with Delay

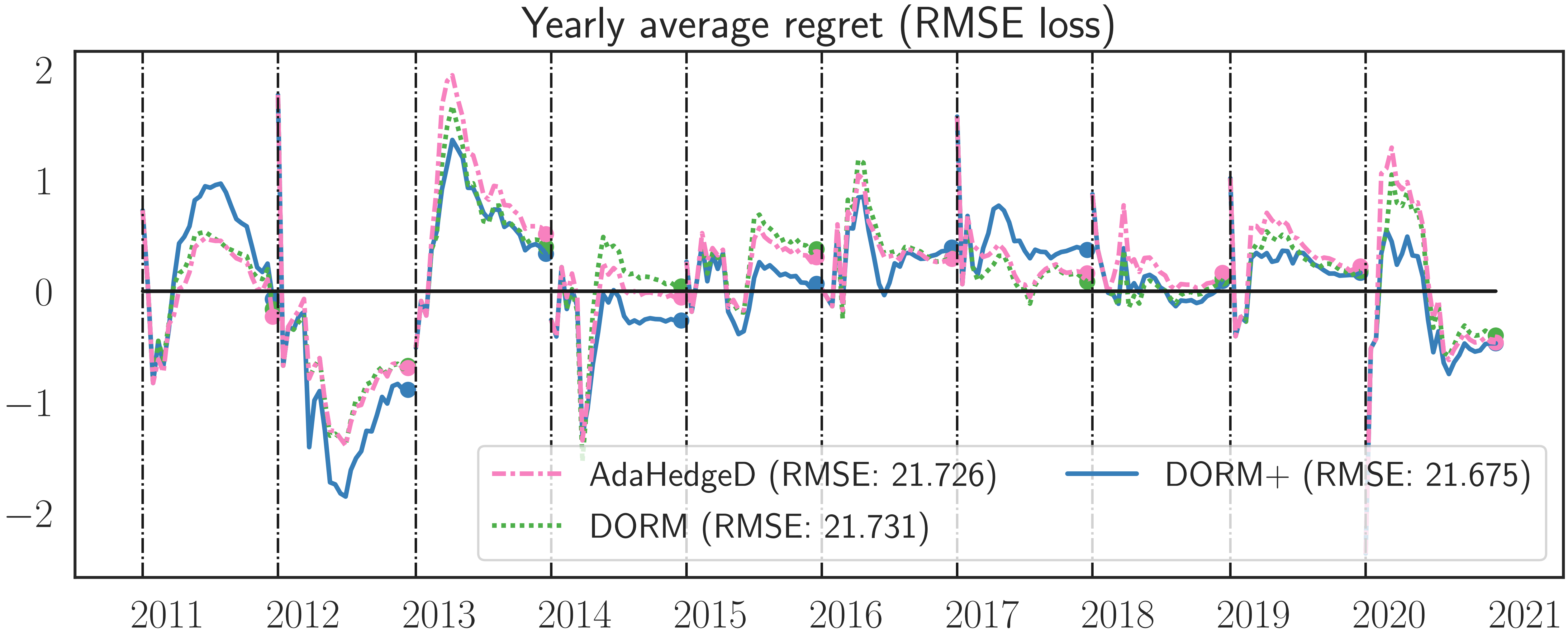
Hinting strategies

- $\texttt{none} : \mathbf{h}_{t+1} = 0$ [don't hint!]
- $\texttt{prev\_g} : \mathbf{h}_{t+1} = \mathbf{g}_{t-2D:t-D}$ [use the last D+1 subgradients]
- $\texttt{mean\_g} : \mathbf{h}_{t+1} = \frac{D+1}{t-D}\mathbf{g}_{1:t-D}$ [replicate mean of all subgradients]
- $\texttt{recent\_g} : \mathbf{h}_{t+1} = (D+1)\mathbf{g}_{t-D}$ [replicate most recent subgradient]

Table 1: **Average RMSE of 2011-2020 semimonthly forecasts**: The online learners compare favorably with the **best input models** and learn to downweight lower-performing candidates, like the *worst models*.

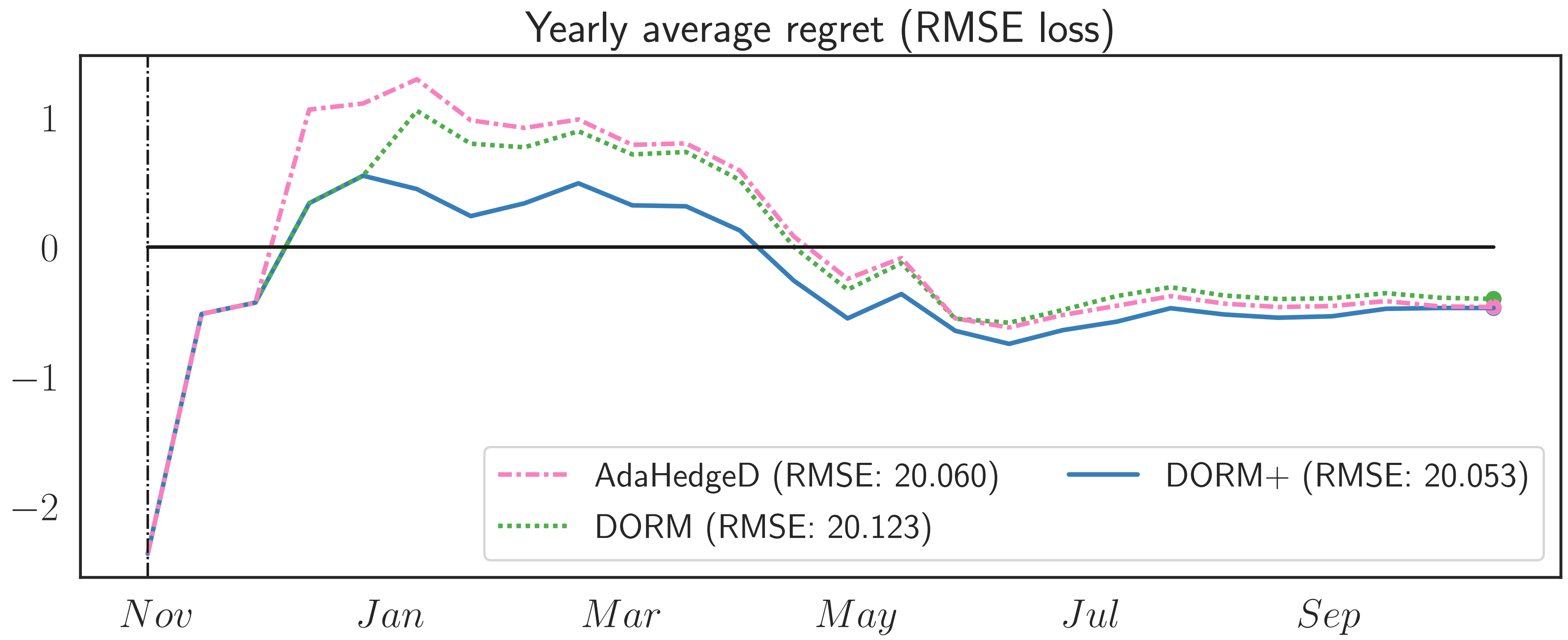
	AdaHedgeD	DORM	DORM+	Model1	Model2	Model3	Model4	Model5	Model6
P3-4	21.726	21.731	**21.675**	**21.973**	22.431	22.357	21.978	21.986	*23.344*
P5-6	21.868	21.957	**21.838**	22.030	22.570	22.383	22.004	**21.993**	*23.257*
T3-4	2.273	2.259	**2.247**	**2.253**	2.352	2.394	2.277	2.319	*2.508*
T5-6	2.316	2.316	**2.303**	**2.270**	2.368	2.459	2.278	2.317	*2.569*

Precip. Weeks 3-4



Takeaway: Small average regret each year despite only T = 26 observations per year

Precip. Weeks 3-4, 2019



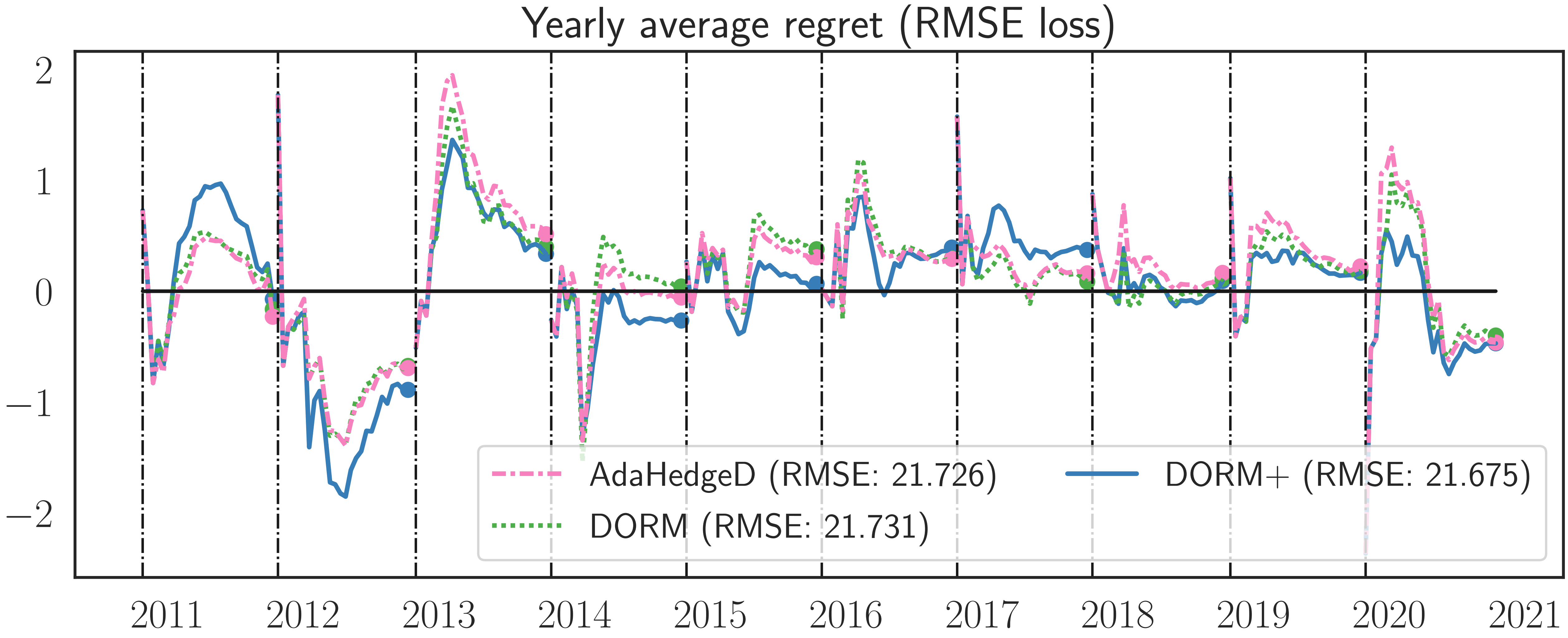
Positive regret

Zero regret = Performance of best input model

Negative regret

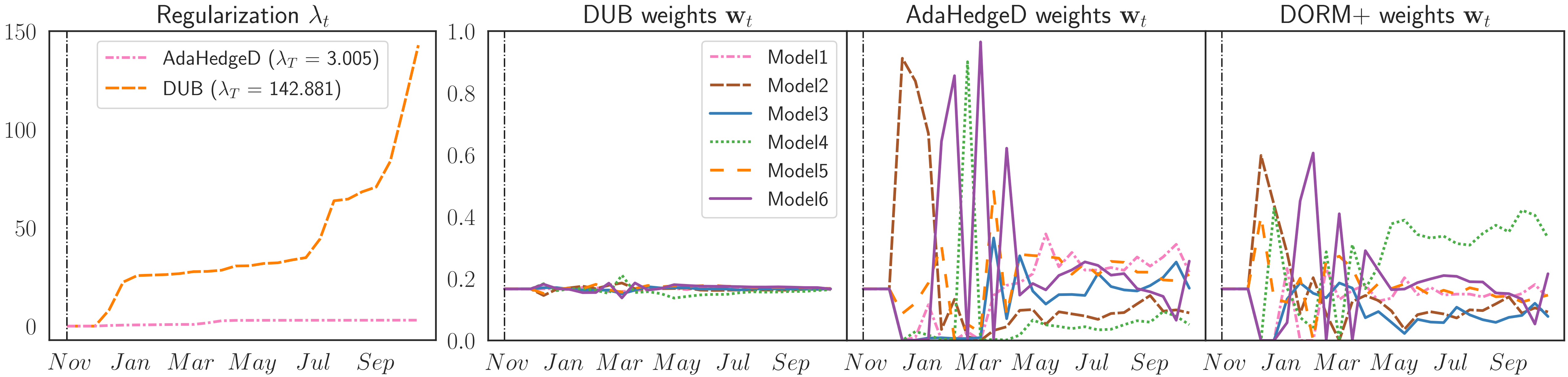
T = 26

Precip. Weeks 3-4



Takeaway: Small average regret each year despite only T = 26 observations per year

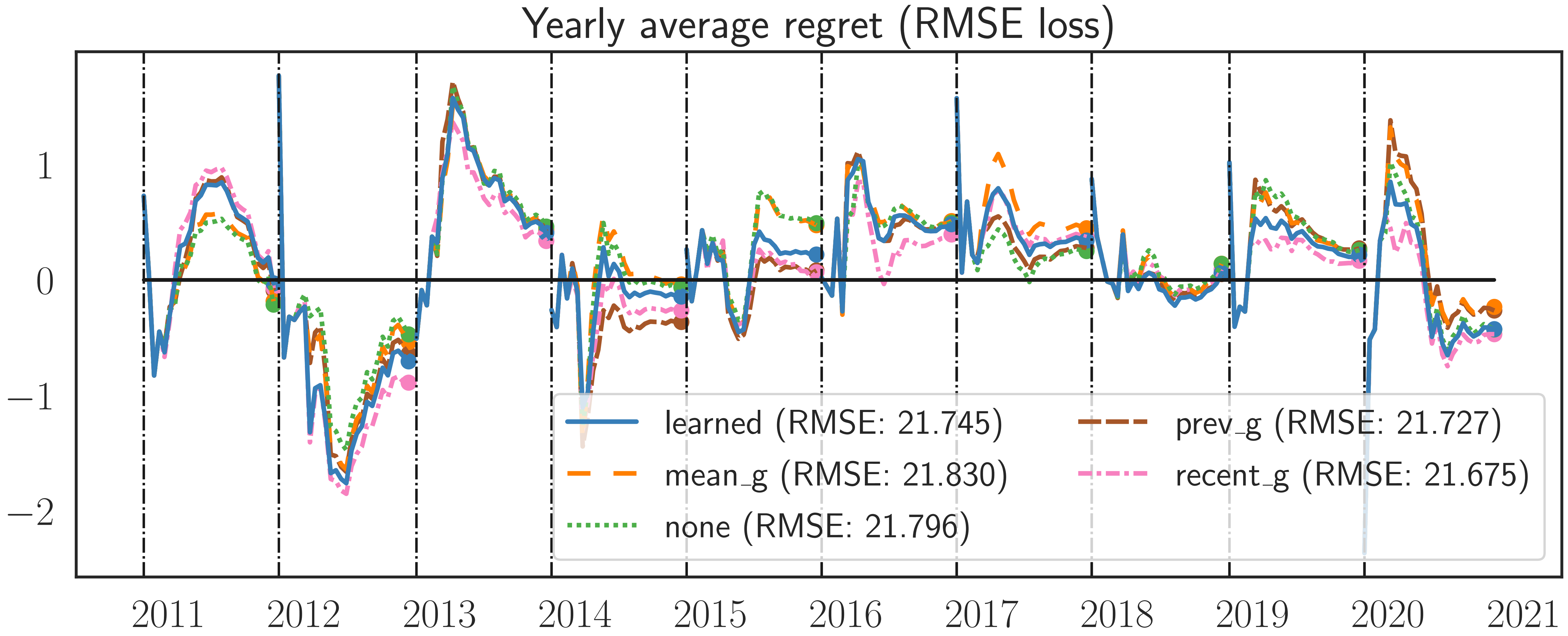
Impact of Regularization: Temp. Weeks 3-4, 2019



Evolution of regularization and model weights for DORM+, AdaHedgeD, and DUB (a more conservative tuning strategy based on a looser regret bound)

Takeaway: AdaHedgeD and DORM+ are much more adaptive to changing model quality

Impact of Optimism: Precip Weeks 3-4, DORM+



- **none** outperformed by all hinting strategies except **mean_g**
- **recent_g** performs best on all four tasks

Learning to Hint with Delay

Observation: DORM, DORM+, & AdaHedgeD all admit bounds of the form

$$(*)\ \mathrm{Regret}_T(\mathbf{u}) \le C_0(\mathbf{u}) + C_1(\mathbf{u})\sqrt{\sum_{t=1}^{T} f_t(\mathbf{h}_t)} \quad \text{for} \quad f_t \text{ convex}$$

(e.g., $f_t(\mathbf{h}_t) = \|\mathbf{r}_t\|_q \|\mathbf{h}_t - \mathbf{r}_{t-D:t}\|_q$ for DORM)

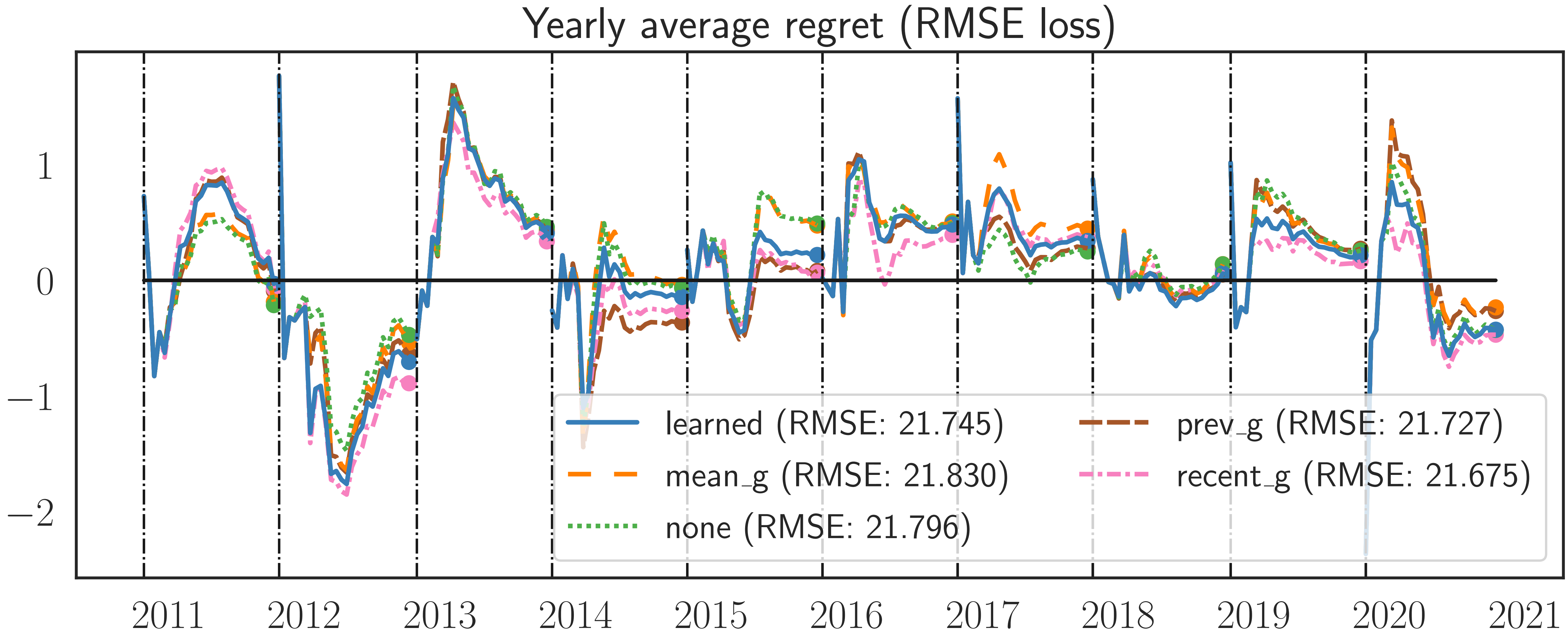
Idea: Combine m different hinting strategies using delayed online learning!

- **Combo hint** $\mathbf{h}_t(\omega) = \mathbf{H}_t\omega$ for hint matrix $\mathbf{H}_t$ and $\omega \in \triangle_{m-1}$

Theorem 1. *If the base online learner satisfies (*) then learning to hint with DORM+ satisfies*

$$\mathrm{Regret}_T(\mathbf{u}) \le C_0(\mathbf{u}) + C_1(\mathbf{u})\sqrt{\inf_{\omega\in\Omega} \sum_{t=1}^{T} f_t(\mathbf{h}_t(\omega))} + o(\sqrt{(D+1)T}).$$

Impact of Optimism: Precip Weeks 3-4, DORM+



- **none** outperformed by all hinting strategies except **mean_g**
- **recent_g** performs best on all four tasks; **learned** is competitive default

Online Learning for Subseasonal Forecasting

Challenges

❌ Delayed feedback

Must issue multiple forecasts before observing feedback about the first

❌ Short regret horizons

Want small regret after only T=26 biweekly forecasts

❌ Impractical hyperparameters

Standard settings based on worst-case future losses: challenging to implement / overly conservative

This talk: New algorithms with

✓ Optimal regret under delay

Even variable and unbounded delays

✓ Hints for missed feedback

Mitigate the impact of delay

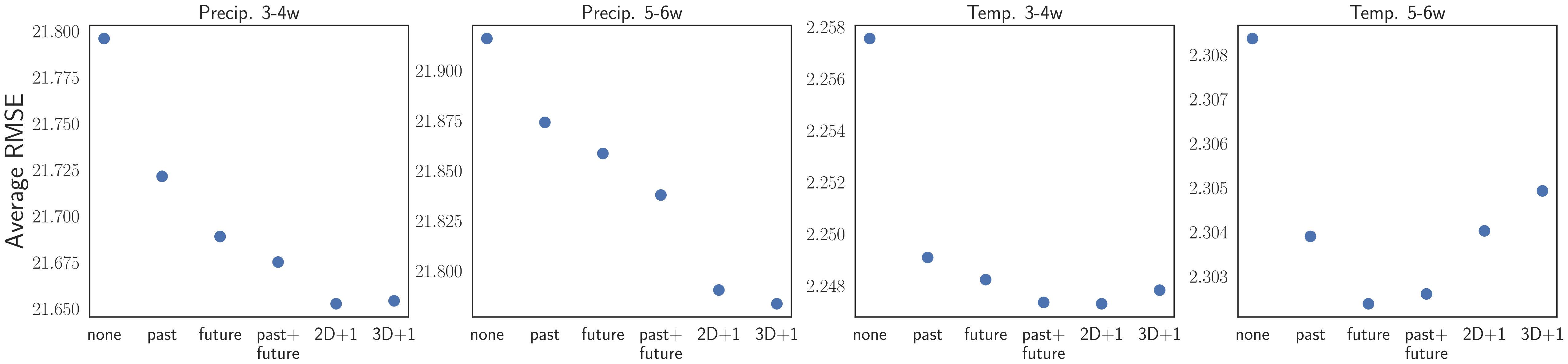
✓ No hyperparameters!

✓ Learning to hint wrapper

Learn effective hinting strategies

Open Questions and Future Work

- Hinting with delay
  - What is the relative impact of hinting at future vs. missed losses?
  - Are there (near-)optimal hinting strategies under delay?



Open Questions and Future Work

- Hinting with delay
  - What is the relative impact of hinting at future vs. missed losses?
  - Are there (near-)optimal hinting strategies under delay?
- Developing tighter convex regret bounds for hint learning
- Domain-specific hinters
  - Use shorter-term forecasters to more accurately predict missed losses

Online Learning with Optimism and Delay
arxiv.org/abs/2106.06885

Python Optimistic Online Learning with Delay (PooID)
github.com/geflaspohler/poold

Code: Paper: