Have Your Cake and Eat it Too?

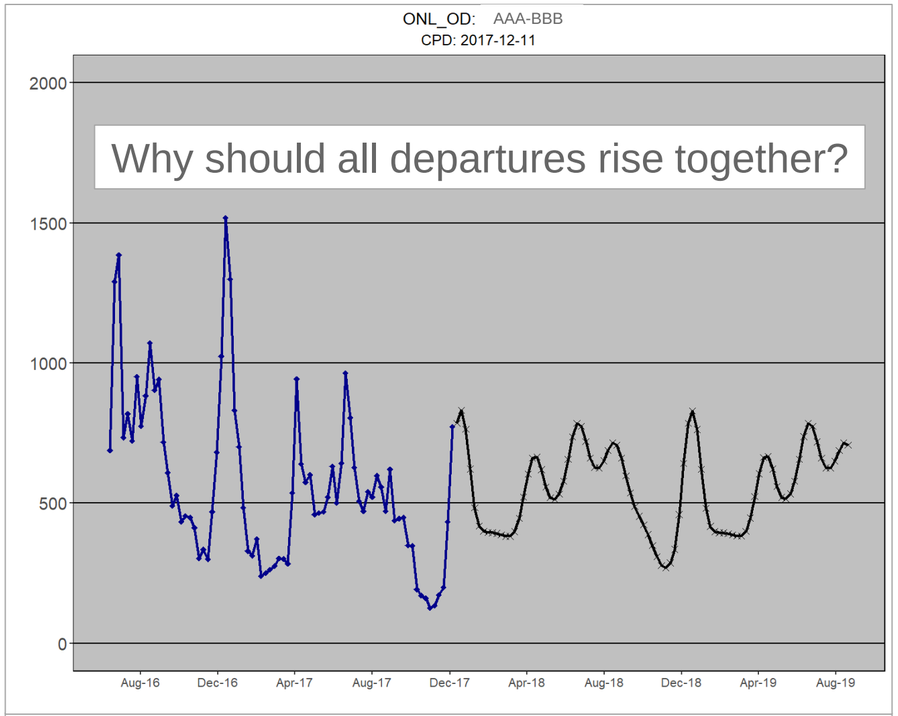
O&D forecasting with short-term responsiveness and long-term stability

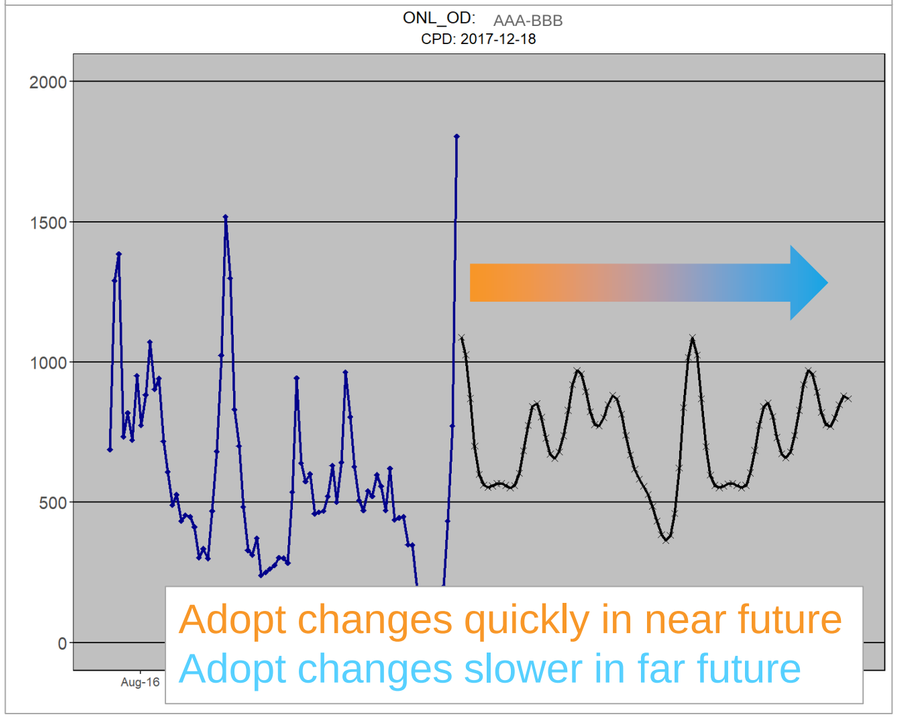
AGIFORS
7-9 May 2019
Panama City, Panama

Inés Garcia Vega – Etihad
Siva Arunachalam – Etihad
Bhaskara Rao Guntreddy – Etihad
Ross Winegar – PROS
Chih-Hsien Wu – PROS

What is Going on Here?

- Setup:
  - Bookings evolve over time
  - Forecasters may use a parameter to capture trending or drift
- Question:
  - Are my forecasts capturing trending well?
  - Looking at forecasts as static snapshots doesn't tell the full story
- Solution:
  - Look at forecasts in motion
  - Observe how they change as new bookings arrive





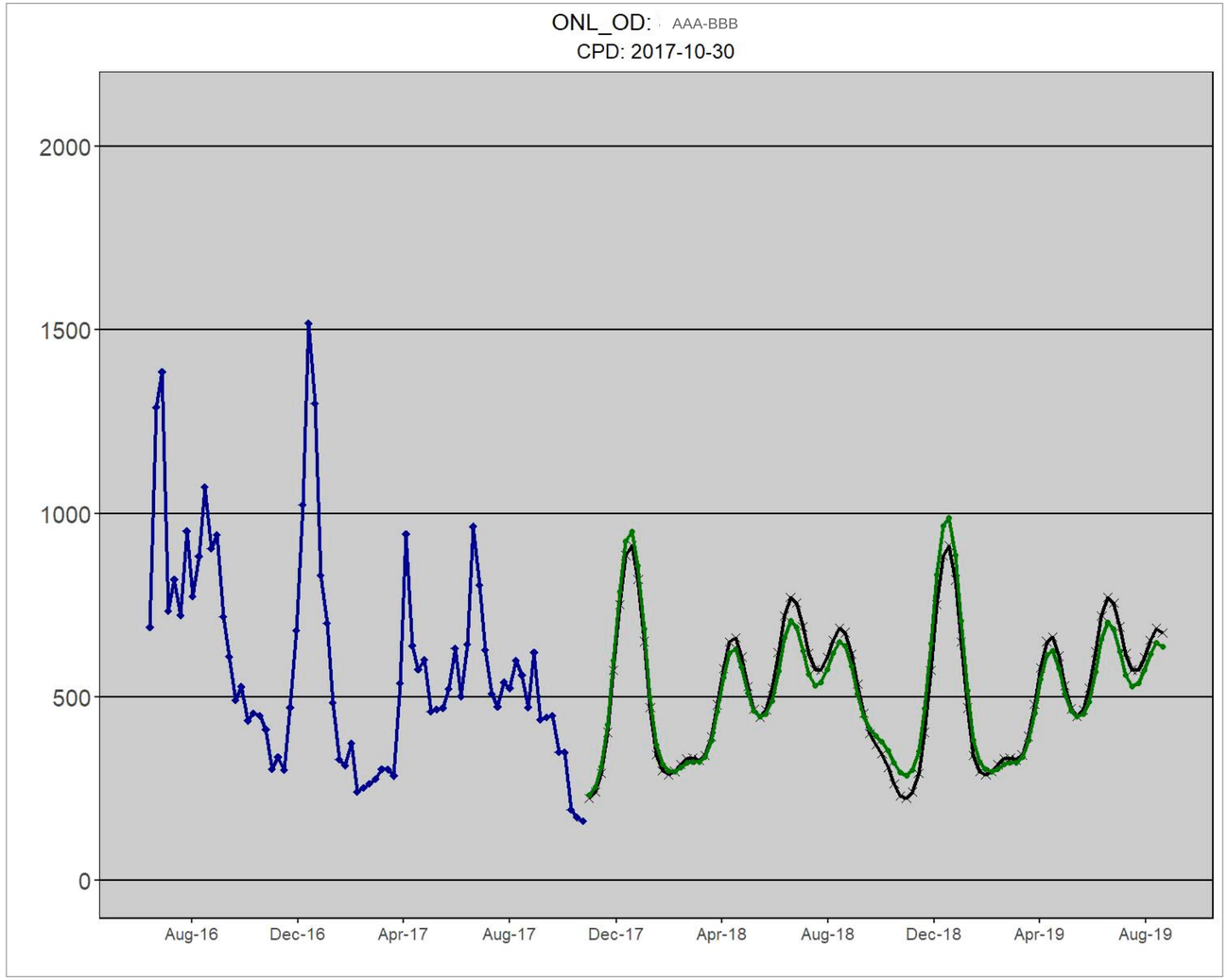
©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Forecasts in Motion

- What if…?
  - Near-term departures captured recent trends
  - Longer-term departures were more stable

- Ensemble forecasts with different drifts

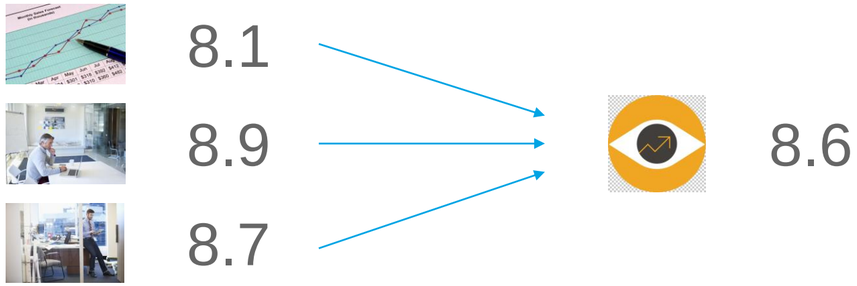
*Alternatives explored include Gradient Boosting and other Machine Learning Models that don't parameterize trending.



©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Ensemble Learning

- Definition
  - Develop a population of base learners
  - Combine base learners into composite predictor to stabilize and improve forecasts



©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

General Framework

- Drift ω is time-weighting parameter
- Convex combination of n forecasts of drifts $\omega \in \Omega = \{\omega_1, \dots, \omega_n\}$
- For Forecast Model m, Departure Date t, DCP d:

$$\hat{f}_{m,t,d}^{Boosted} = \sum_{\omega \in \Omega} \alpha_{\omega,k} \hat{f}_{m,t,d}^{\omega}$$

- Weight $\alpha_{\omega,k}$ depends on weeks from Current Date
  - More weight on low drift for far-away forecasts
  - More weight on high drift for near-future forecasts

©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Setup – Weighting by Distance from Current Date

High drift:
focus on nearby future

High-drift forecast:
$\hat{f}^{\omega_1}$

Slow drift:
focus on far future

Low-drift forecast:
$\hat{f}^{\omega_2}$

Multi-drift:
focus on both

Boosted forecast:
$\alpha \hat{f}^{\omega_1} + (1-\alpha)\hat{f}^{\omega_2}$

History

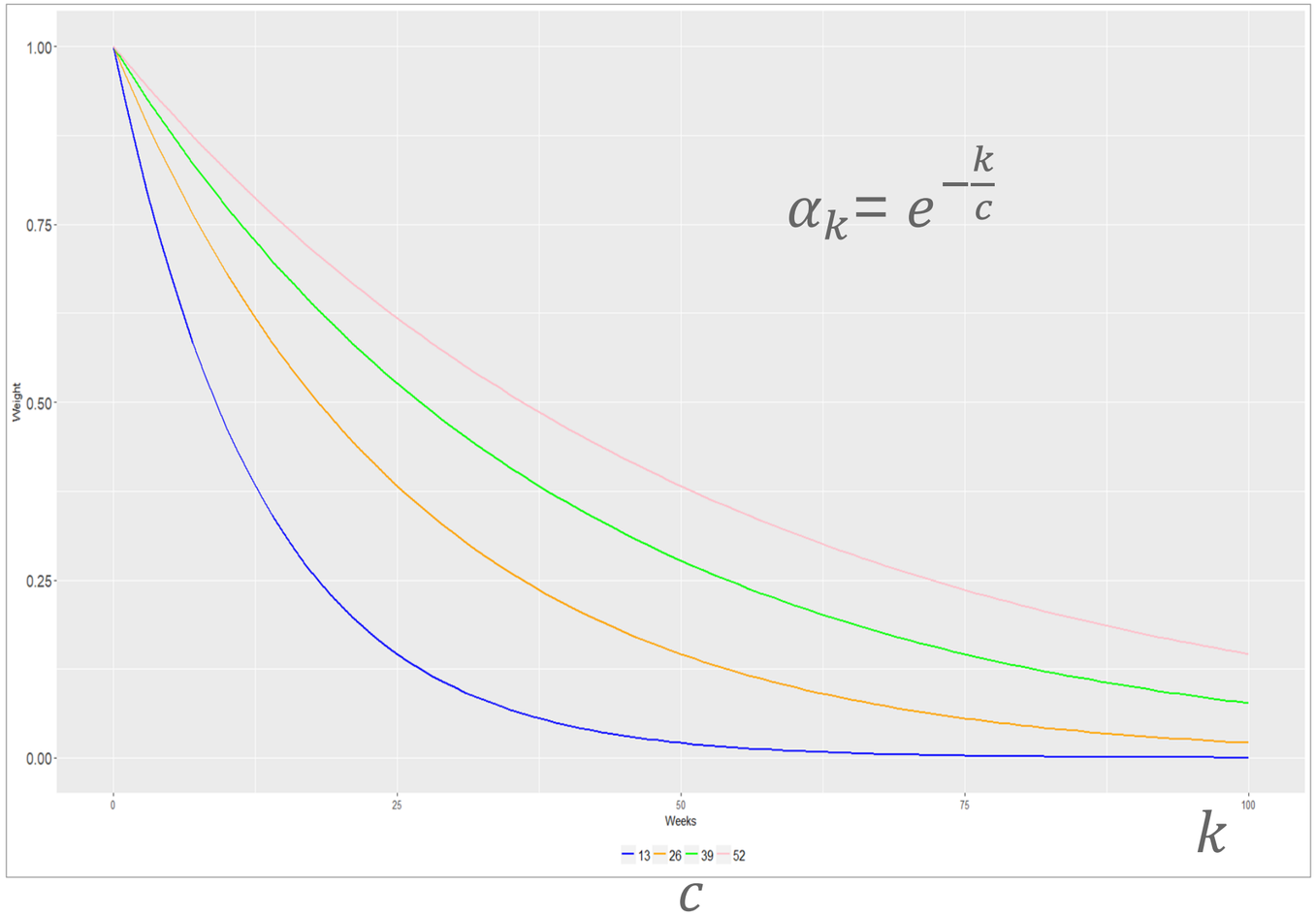
Current Date

Future

©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

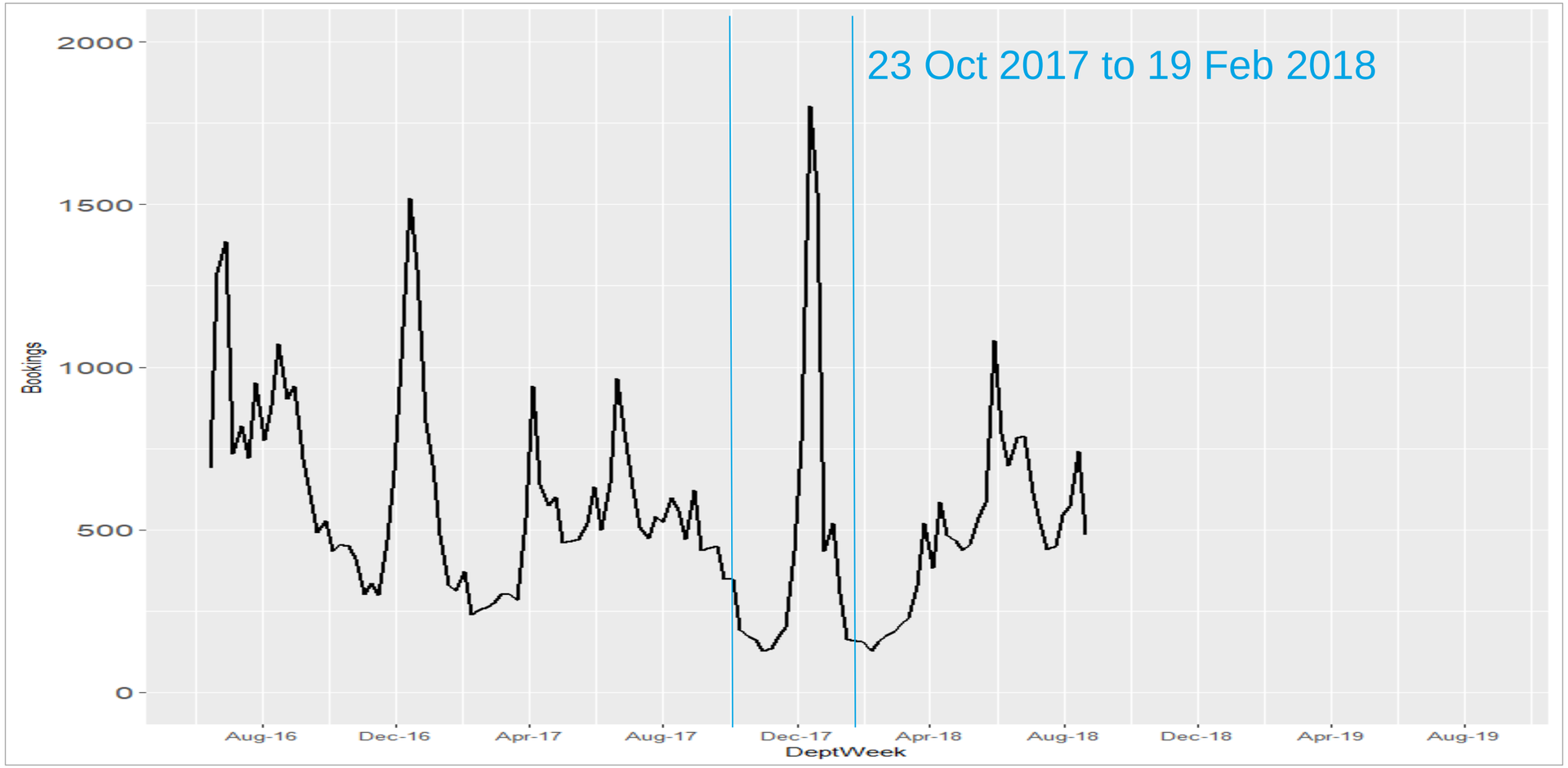
Weight Function

- $\alpha_k = e^{-\frac{k}{c}}$ is weight on high drift forecasts
- k is the departure weeks away from Current Date
- c is the parameter controlling weight
  - $c \to 0$ implies all weight on slow-drift forecasts
  - $c \to \infty$ implies all weight on high-drift forecasts



©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

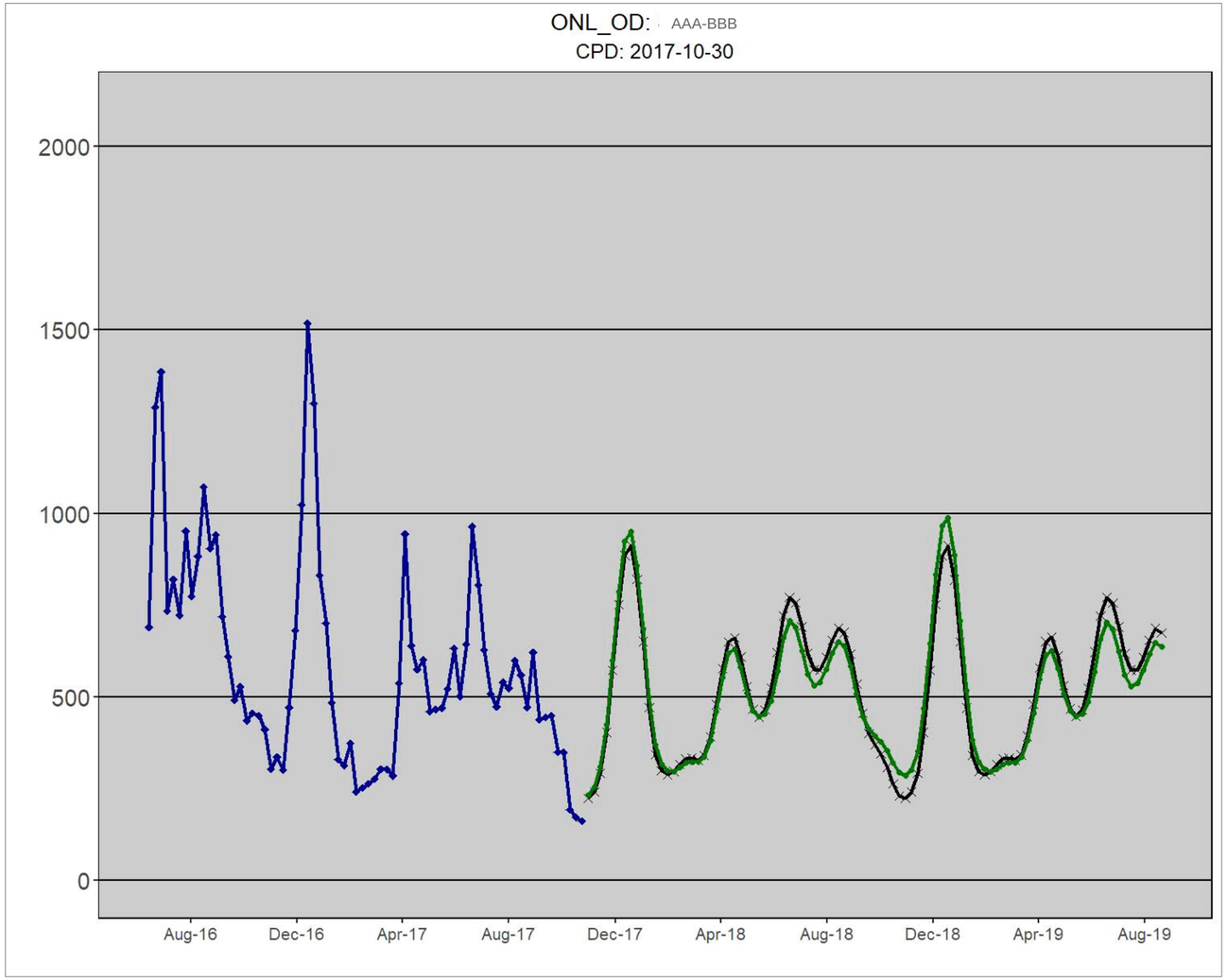
Current Date Moving Through Peak Outlier



©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

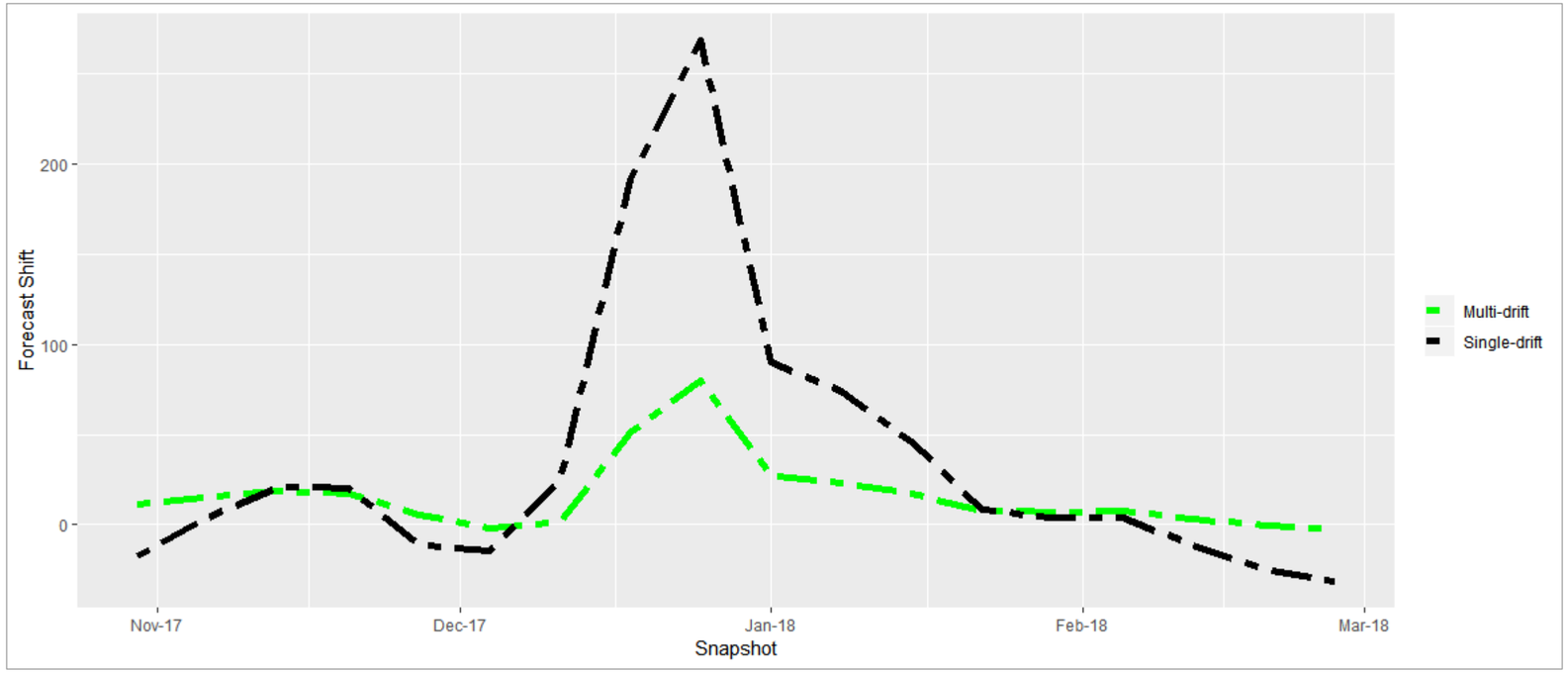
Multi-drift Boosted Forecasts Improve Stability and Accuracy

- Boosted forecasts more stable moving in and out of outlier peak



©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Boosted is More Stable and More Accurate



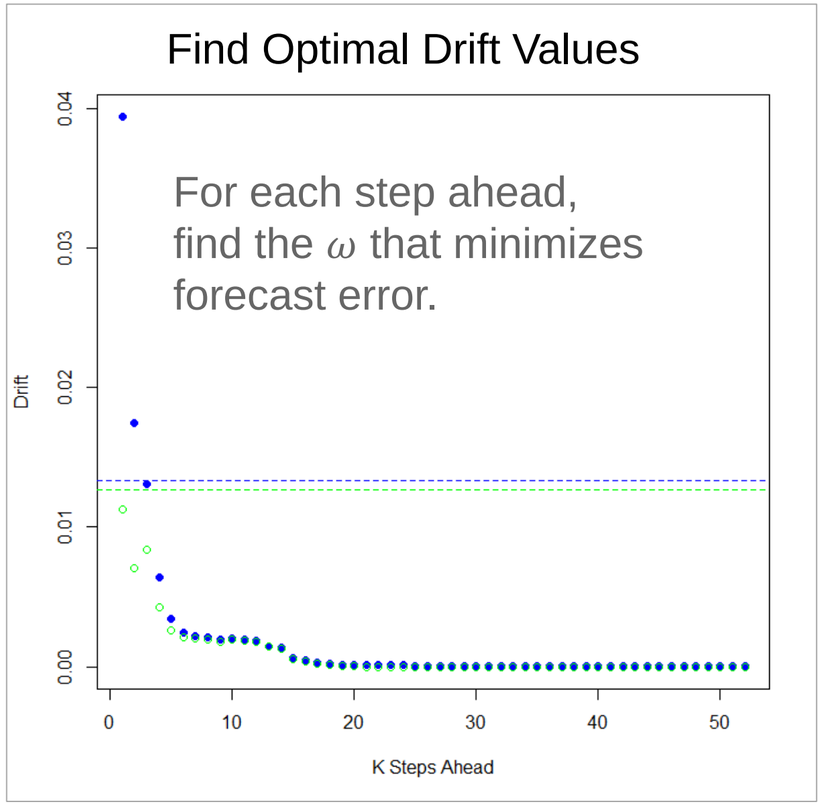
Average forecast level by snapshot, and closer to middle implies more stable. Boosting has improved stability.

k weeks away from Current Date	RMSE for Single Drift forecasts	RMSE for Boosted forecasts
1	1.13	1.13
2	1.27	1.27
3	1.36	1.35
4	1.40	1.40
5	1.44	1.43
6	1.46	1.45
7	1.45	1.43
8	1.47	1.45
9	1.50	1.48
10	1.51	1.49
11	1.52	1.50
12	1.52	1.50
13	1.50	1.47
14	1.52	1.50
15	1.54	1.51
16	1.53	1.50
17	1.44	1.41
18	1.38	1.36
19	1.39	1.36
20	1.41	1.37
21	1.43	1.39
22	1.46	1.41
23	1.47	1.43
24	1.51	1.46
25	1.55	1.50
26	1.63	1.57
27	1.66	1.60
Overall	**1.46**	**1.43**

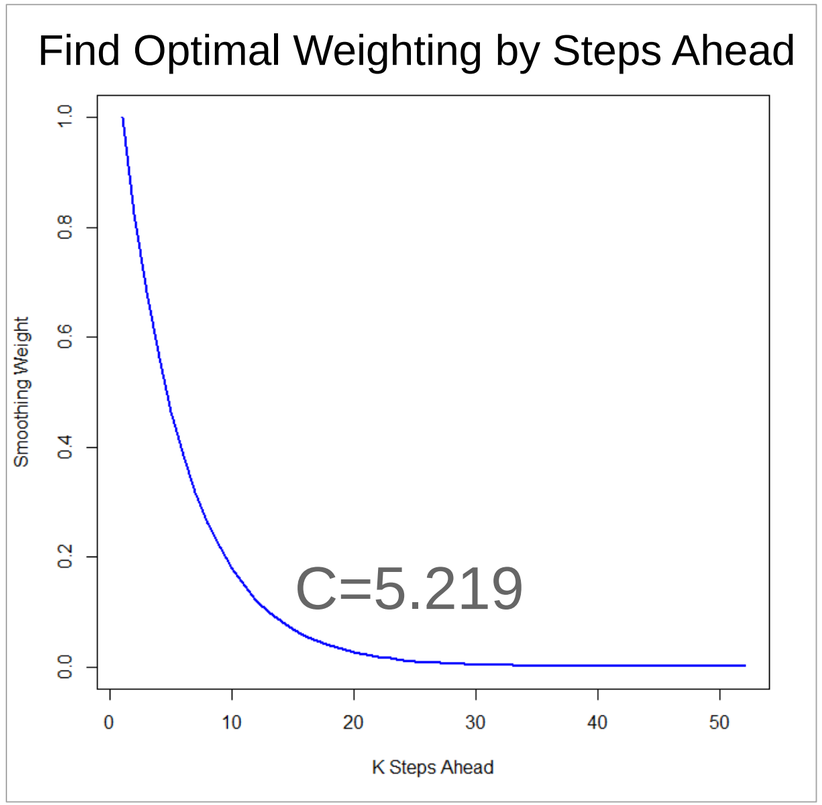
RMSE improves for distant departures.

©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Tuning Multi-Drift



Pick high and low drifts using estimated fit over these ω.



Use smoothing weighting function:

$\alpha_k = e^{-\frac{k}{c}}$, where $c = -\frac{1}{\hat{\beta}}$

Boosted forecast:

$\hat{f}_{boosted} = \alpha \hat{f}^{\omega_{fast}} + (1-\alpha)\hat{f}^{\omega_{slow}}$

Fit regression:

$\ln(\alpha_k) = \beta k + \varepsilon$ to find $\hat{\beta}$

©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Summary – Multi-Drift Boosting

- Boosted method has more stable forecasts – both moving in and out of seasons and for distant departures

- Results stabilize forecasts and provide better RMSE

The forecast is reacting to recent trends (eat the cake) while maintaining stability (and keep it too)

So how does that cake taste for you?

©2019 PROS, Inc. All rights reserved. Confidential and Proprietary

Thank You

PROS Headquarters | Houston, Texas | 3100 Main Street, Suite 900 | Houston, TX 77002, USA | +1-713-335-5105 +1-800-555-3548