Detection and modeling of long memory in biases of daily forecasts of air pressure

Yulia Gel, Bovas Abraham

Department of Statistics and Actuarial Science, University of Waterloo

OUTLINE

1. PROBLEM STATEMENT AND DATA DESCRIPTION.

2. MODEL CANDIDATES: DECAYING AVERAGING, HOLT-WINTERS, ARMA, ARFIMA

3. TIME SERIES PROPERTIES OF BIASES

4. MODELS PREDICTIVE POWER

5. CONCLUSION

PROBLEM STATEMENT AND DATA DESCRIPTION

$$BIAS = FCST - "OBS",$$

where $"OBS"$ is a $REANALYSIS$.

GOAL is to predict future biases, $B\hat{I}AS$, from historical data. Then

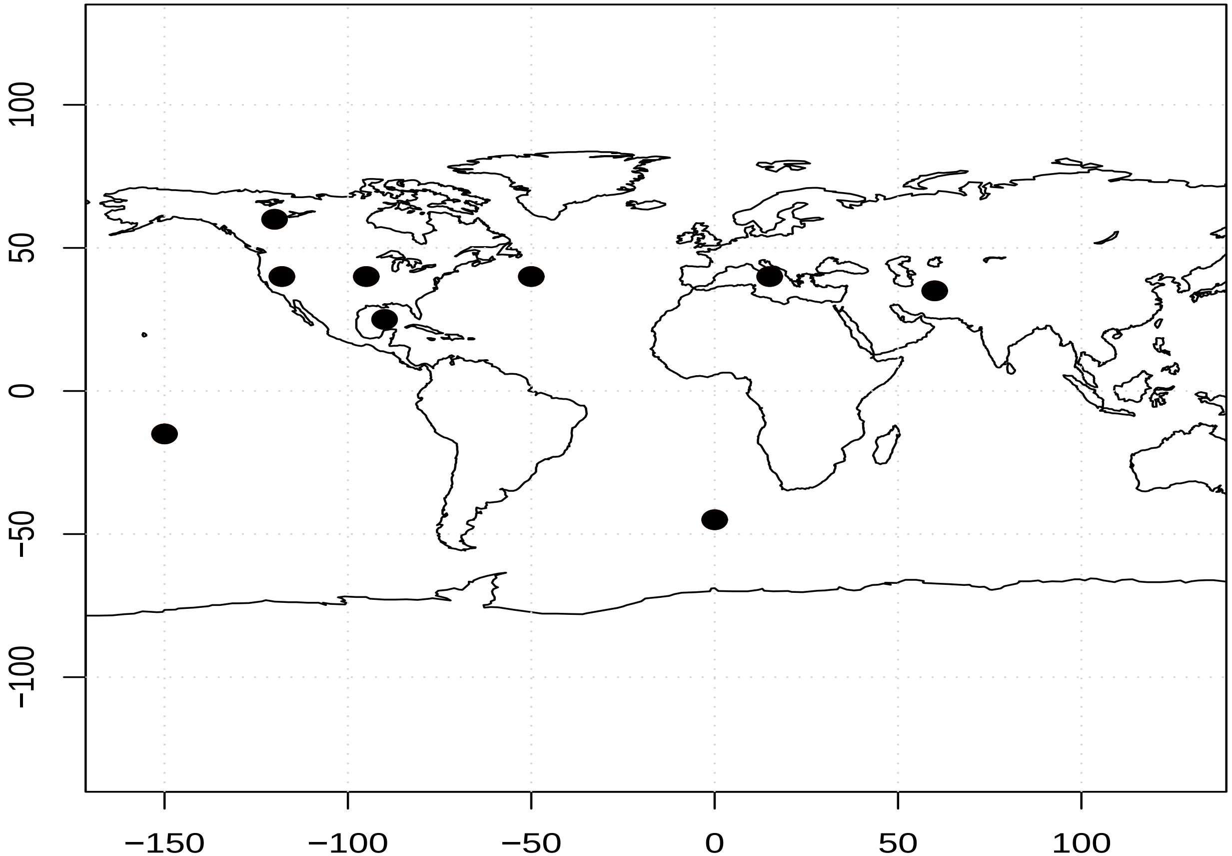
$$FCST_{calibrated} = FCST_{OPR} - B\hat{I}AS.$$

APPLICATION TO NCEP Operational Ensemble

DATA: NCEP T00Z 10 Ensemble Forecast of 500 MB HEIGHT, lead time $12, 24, \ldots, 384$ hours (from 1 to 16 days ahead), 2001-2004.

LOCATIONS: 9 points worldwide

Figure 1: Data from 9 locations are analyzed.



MODEL CANDIDATES

- **Decaying Averaging (DA)** is the NCEP Adaptive (Kalman Filter Type) Bias-Correction Algorithm – *variation of the Single Exponential Smoothing (SES)*

$$\hat{\text{BIAS}} = (1-\omega)\times\text{prior t.m.e.}+\omega\times\text{the most recent BIAS},$$

  - for each lead time separately;
  - t.m.e.=time mean error which is calculated using historical biases within the sliding window of 30 most recent days;
  - $\omega = 2\%$ is selected subjectively (for some not included locations ω is 5% or 10%)

- **Holt-Winters (HW) Smoothing** – *extension of DA and SES* which
  - takes into account linear trends and additive or multiplicative seasonality (if any is detected).
  - is known to provide poor long term predictions
- **ARMA(p, q)** – Autoregressive Moving Average Model

$$a(B)y_t = b(B)v_t,$$

  - $a(\cdot)$ and $b(\cdot)$ are polynomials of degree p and q, $v_t \sim WN(0,1)$.
  - the autocorrelation function (acf) decays exponentially, i.e. a short memory process.

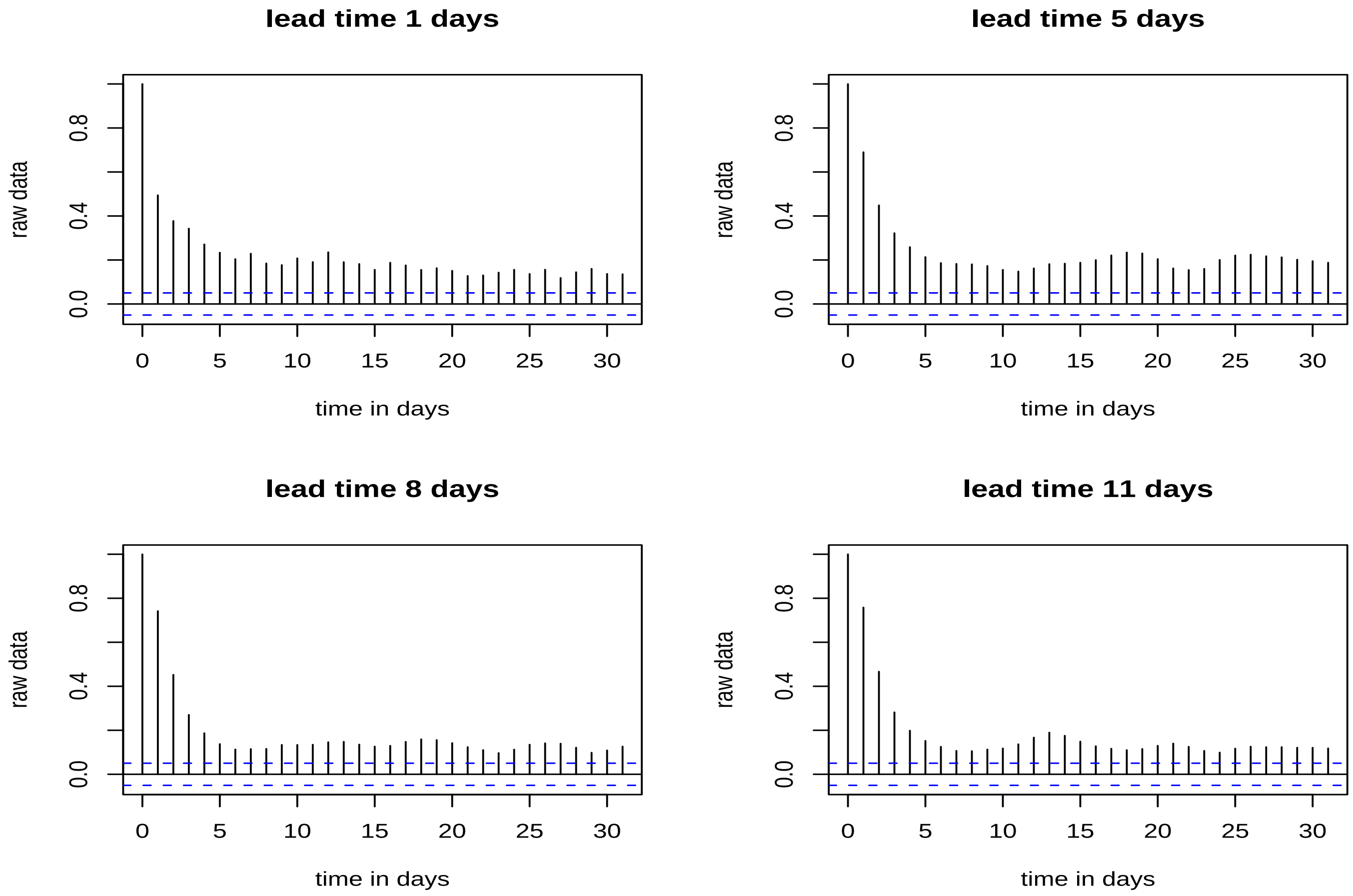
- **ARFIMA(p, d, q)** – Fractionally Integrated Autoregressive Moving Average (*long memory*) model

$$a(B)(1-B)^d y_t = b(B)v_t, \quad d \in (-0.5, 0.5).$$

  - $a(\cdot)$ and $b(\cdot)$ are polynomials of degree p and q, $v_t \sim WN(0,1)$.
  - the autocorrelation function (acf) decays geometrically, i.e. a long memory process.

- Two Stage Procedure – Sliding Window Demeaning + Linear Models:
  - estimate parameters of a linear time series model, e.g. an ARMA or ARFIMA model, whose coefficients are obtained using the most recent available biases (sliding *large* window parameter estimation);
  - removes the mean calculated on the most recent available raw biases (sliding *short* window demeaning).

Figure 2: Typical Autocorrelation Plots of Biases (location (-15, -150)).



LONG MEMORY TESTS: DETECTION OF d

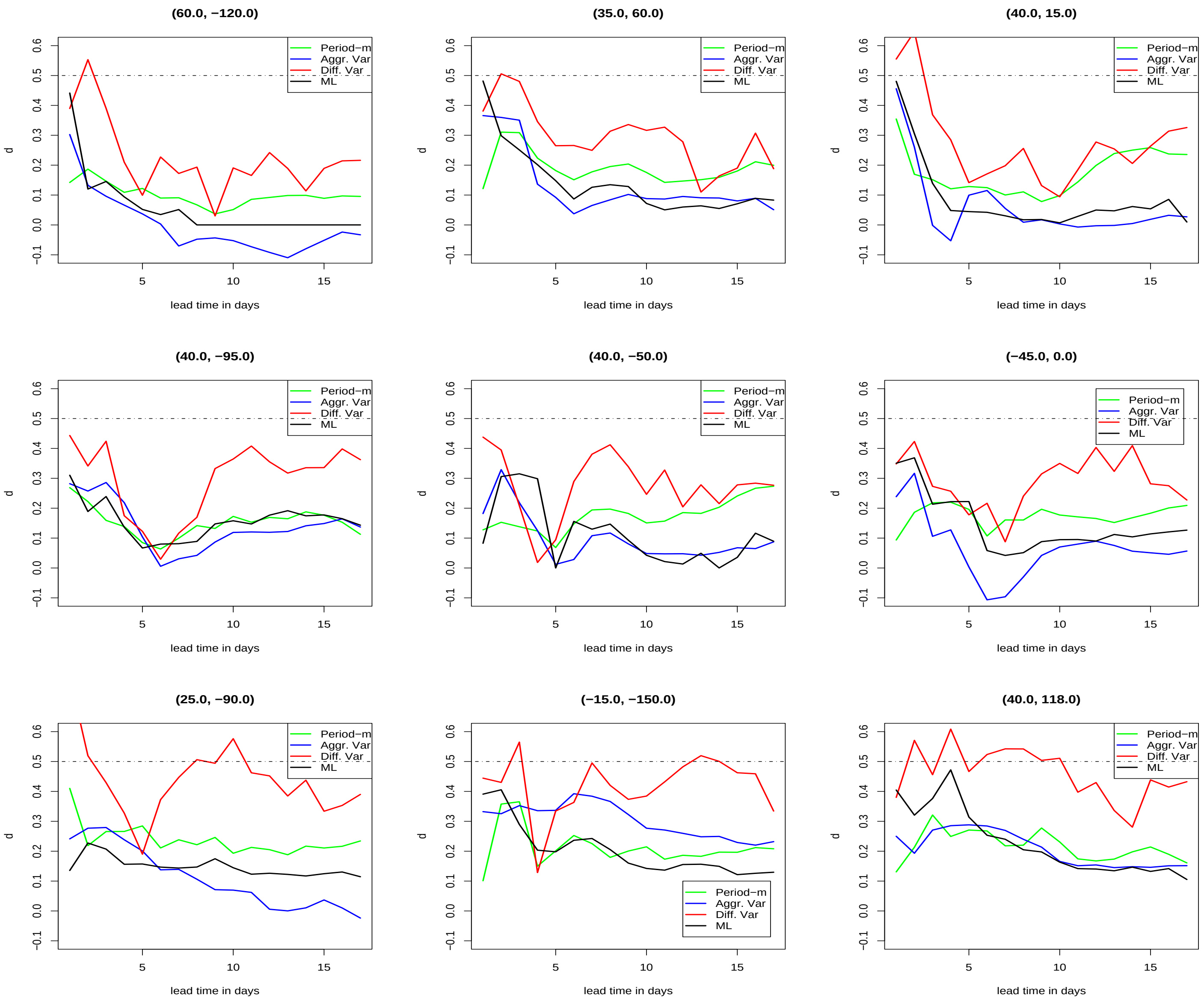


Figure 3: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

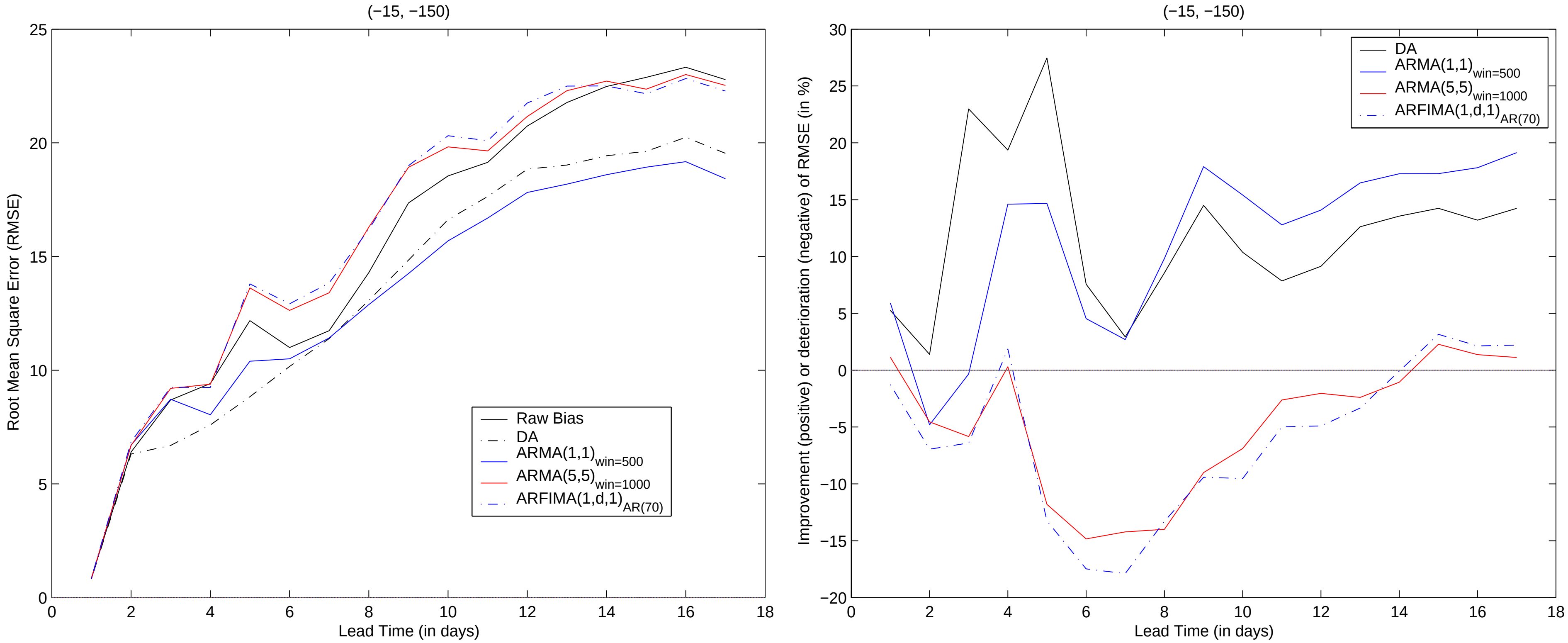


Figure 4: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\mathrm{CorrectedRMSE}}{\mathrm{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

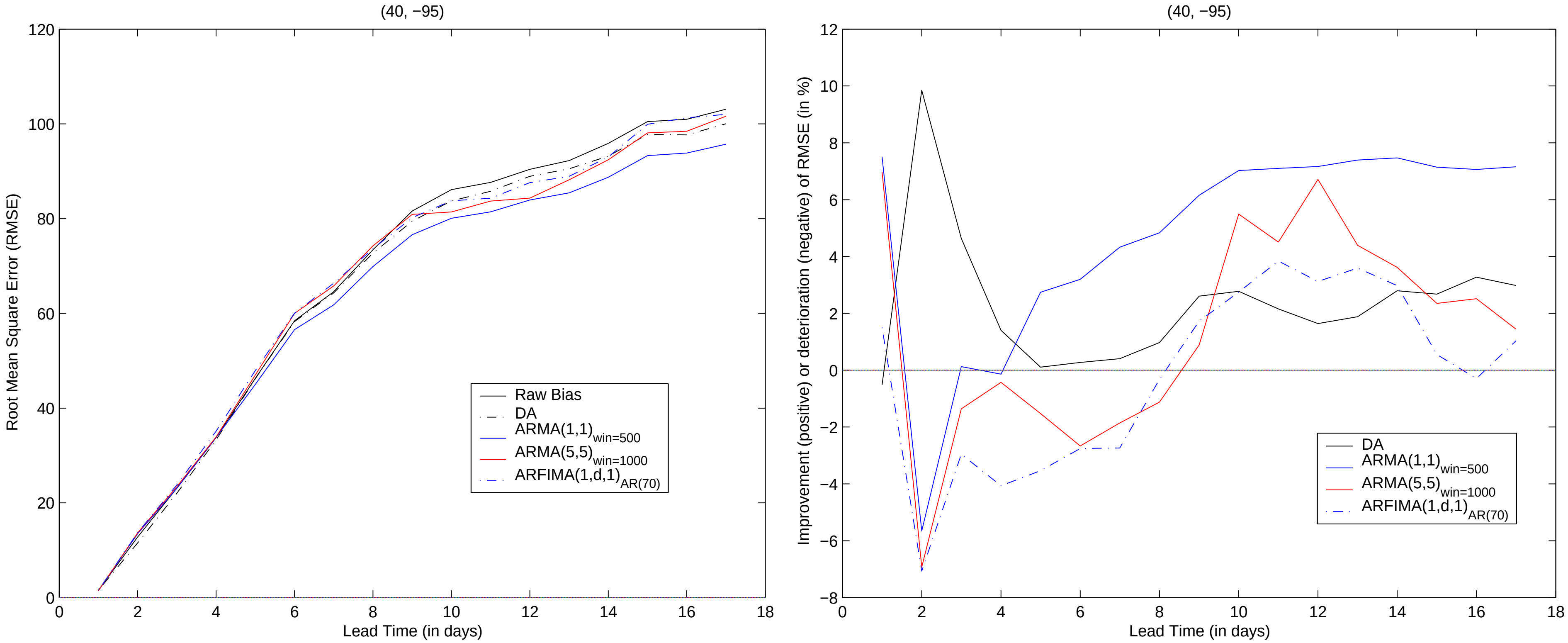


Figure 5: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

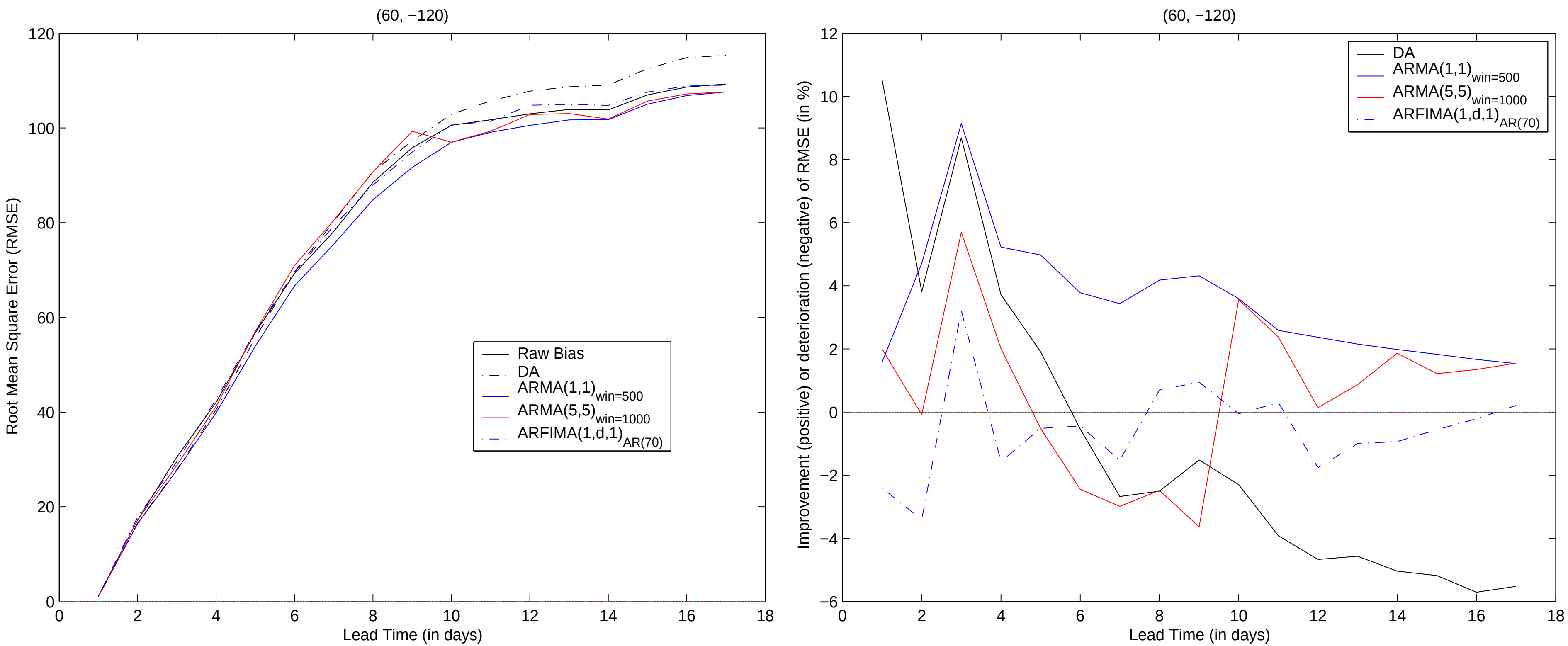


Figure 6: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\mathrm{CorrectedRMSE}}{\mathrm{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

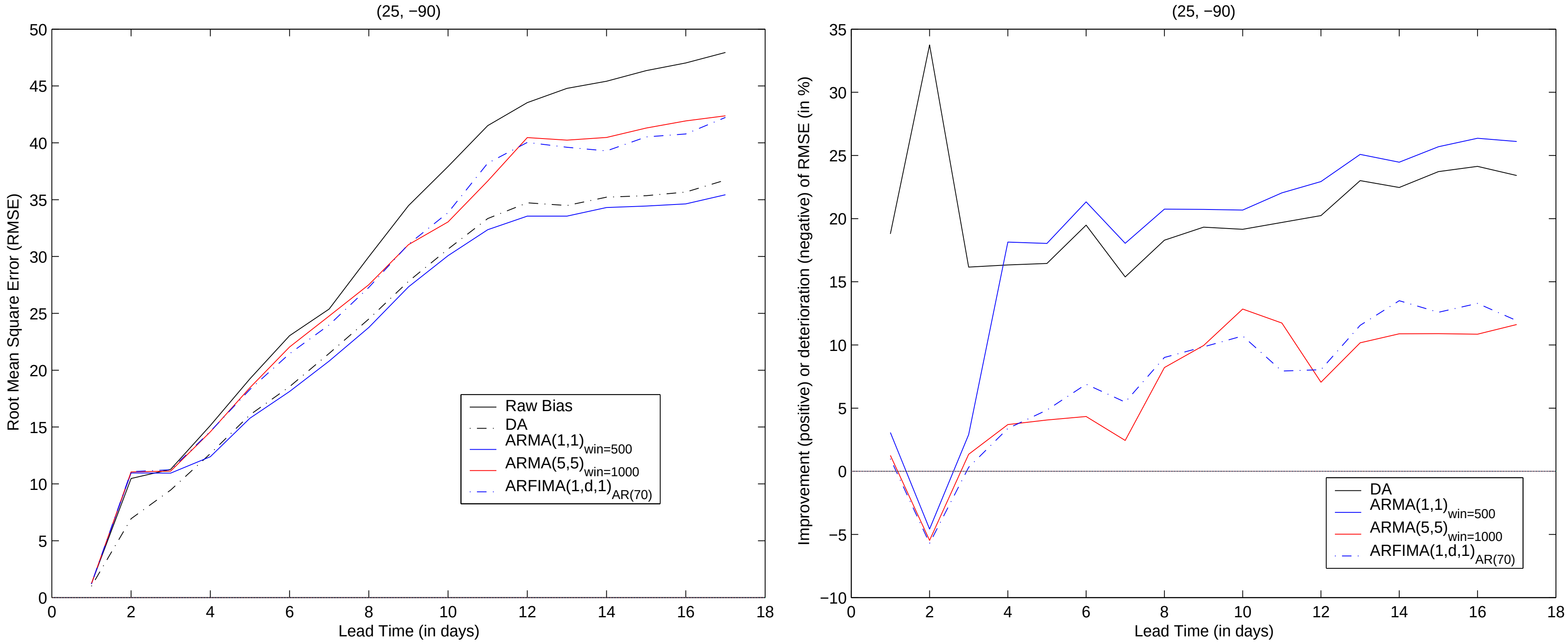


Figure 7: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\mathrm{CorrectedRMSE}}{\mathrm{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration).DA is the Decaying Averaging method.

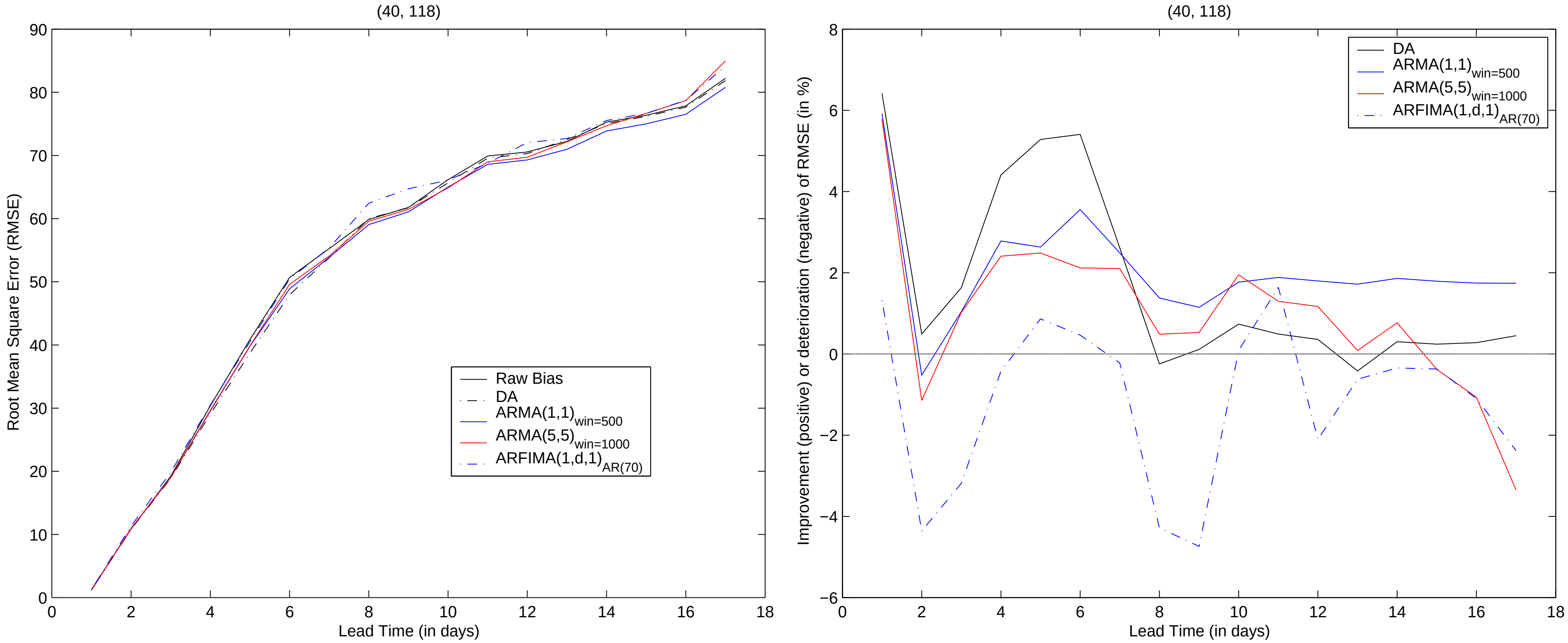
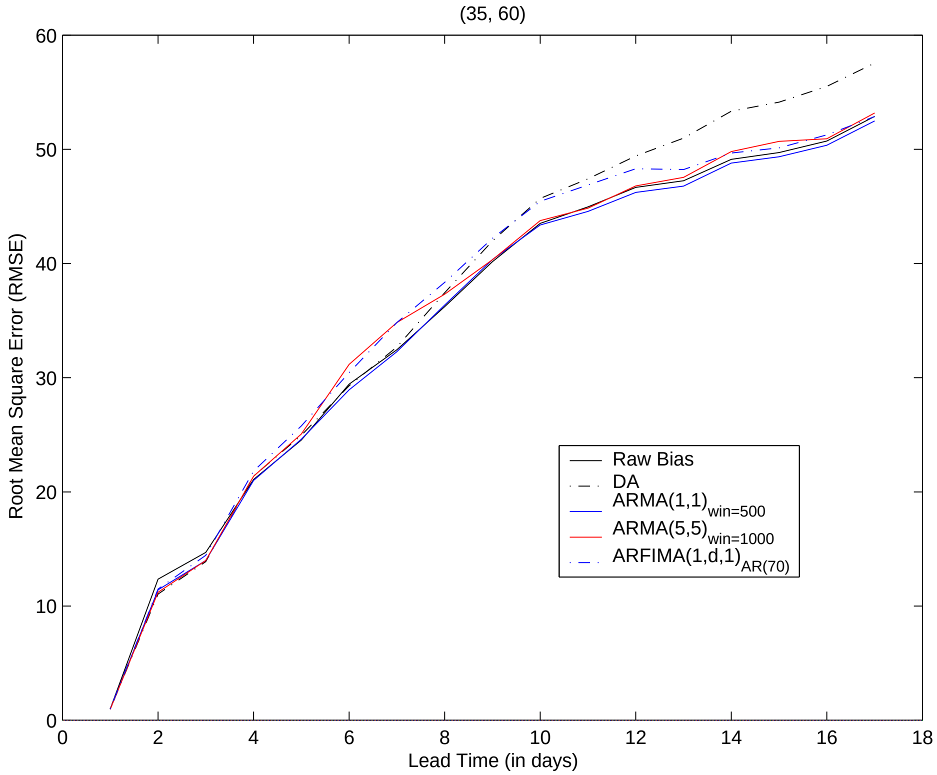


Figure 8: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.



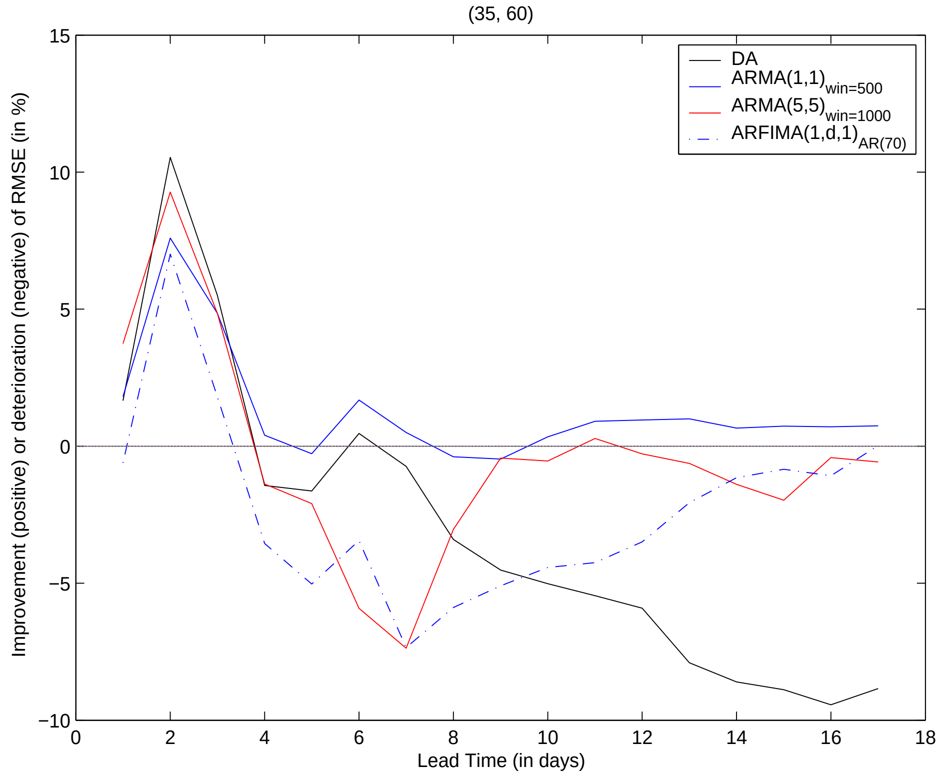


Figure 9: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

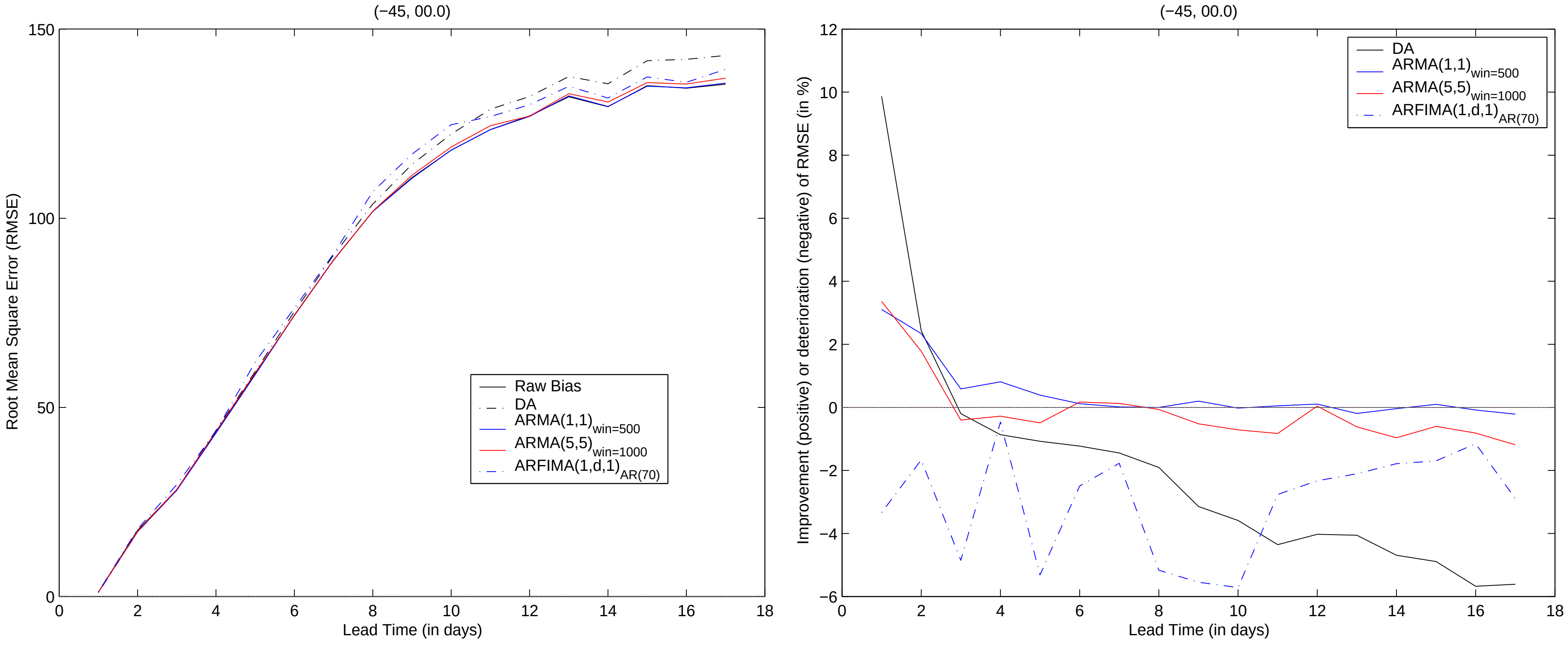


Figure 10: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.

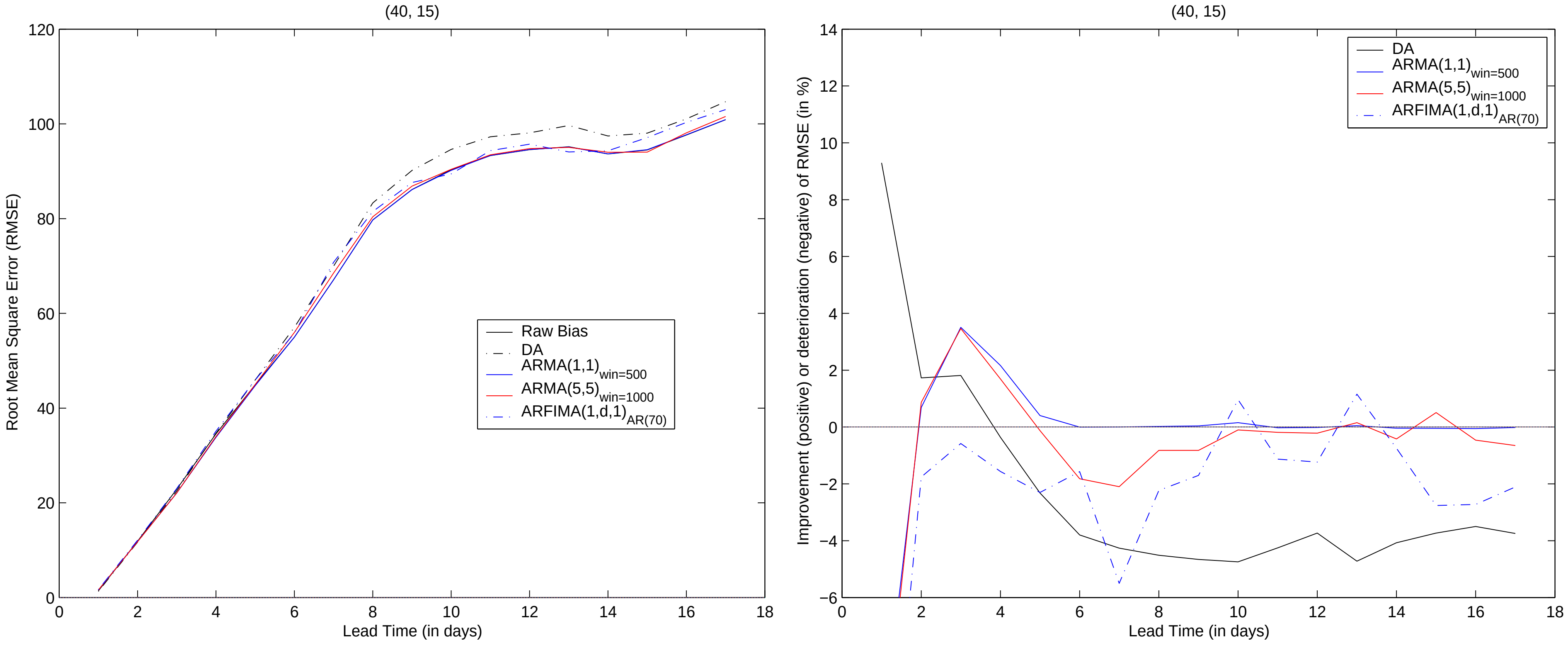
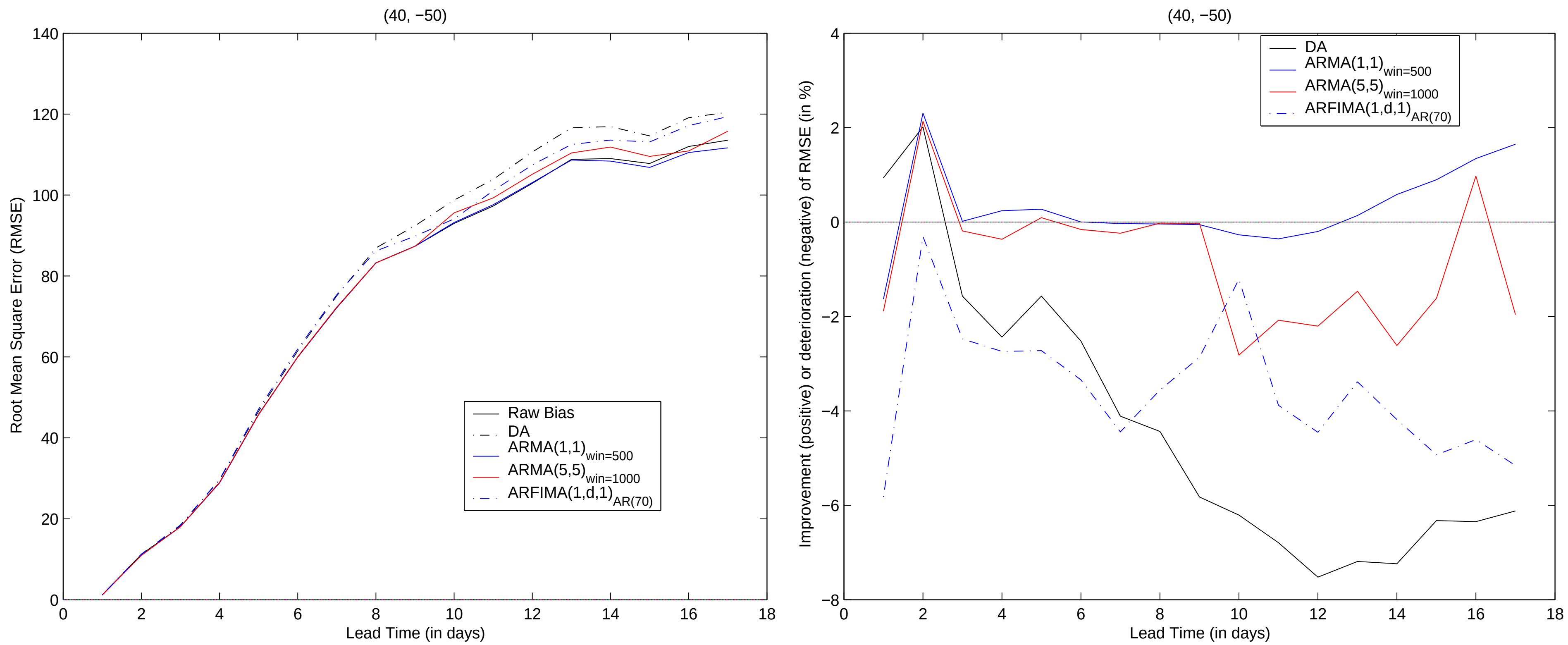
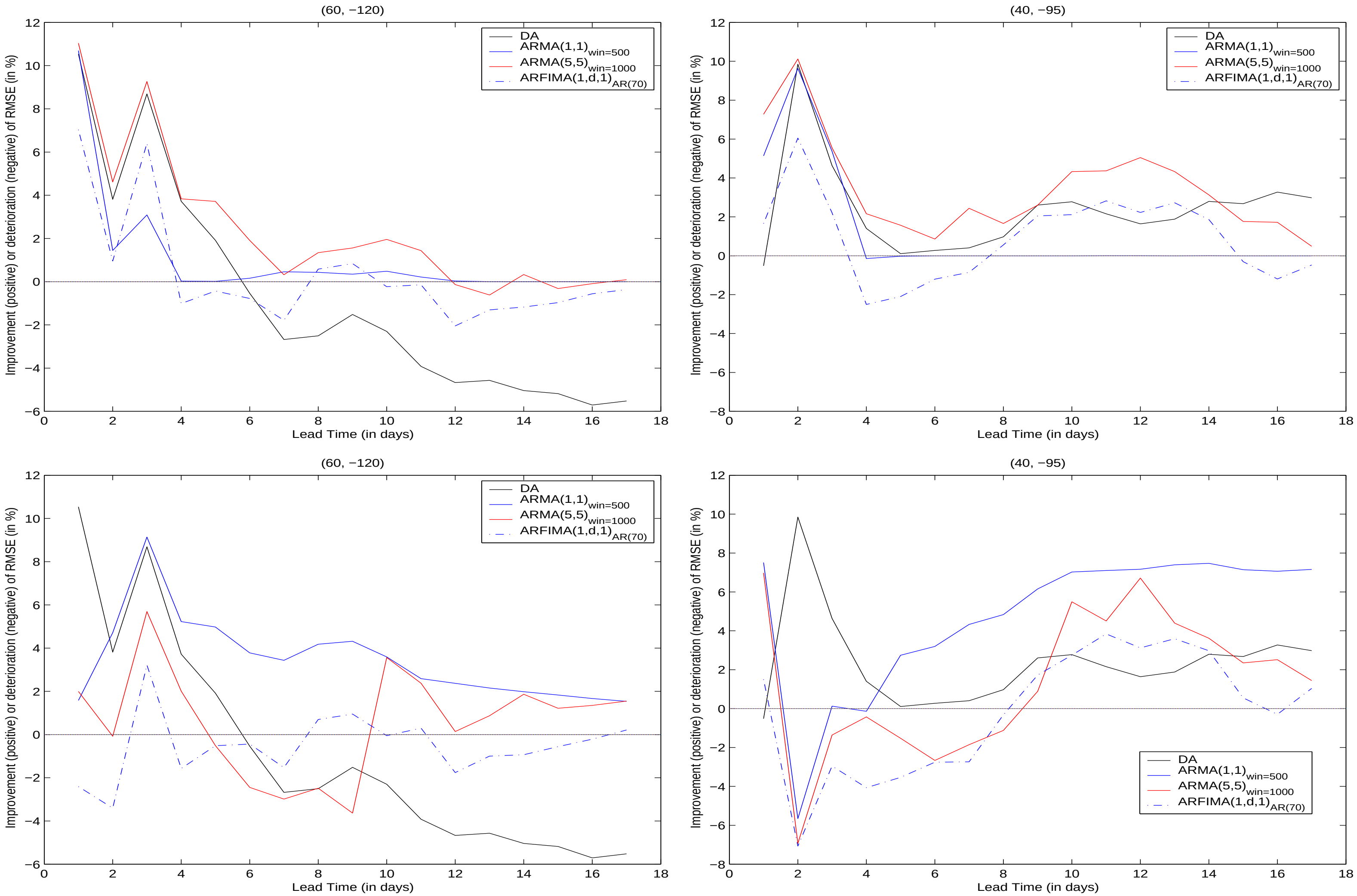


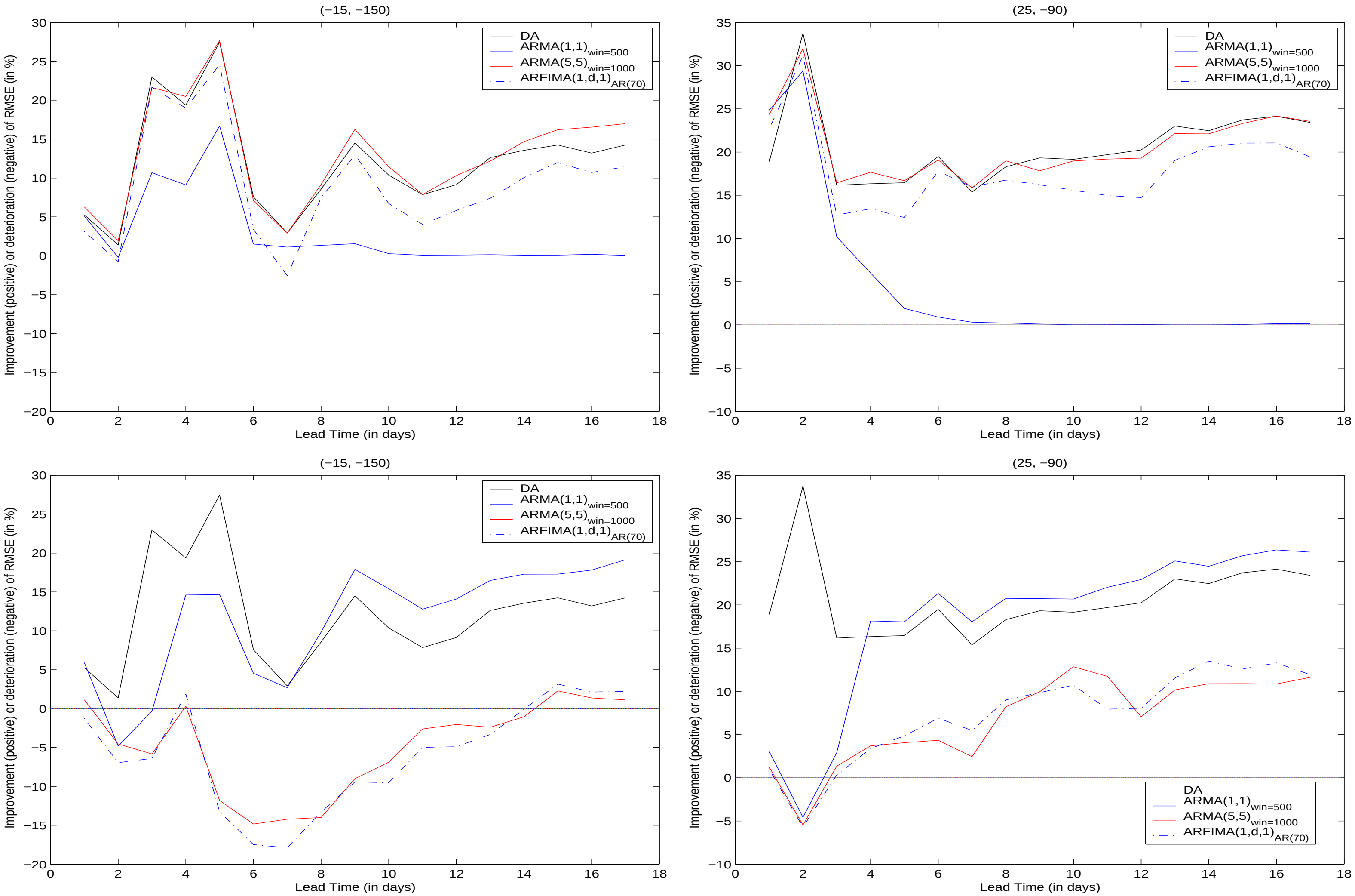
Figure 11: Predictive Power of the Two Stage Procedure: Sliding Window Demeaning + Linear Time Series Models. The Plots of the Raw and Corrected Root Mean Square Errors (RMSE) (left) and their ratios $1 - \frac{\text{CorrectedRMSE}}{\text{RawRMSE}}$ (right) (positive values imply improvement, negative ratios imply deterioration). DA is the Decaying Averaging method.



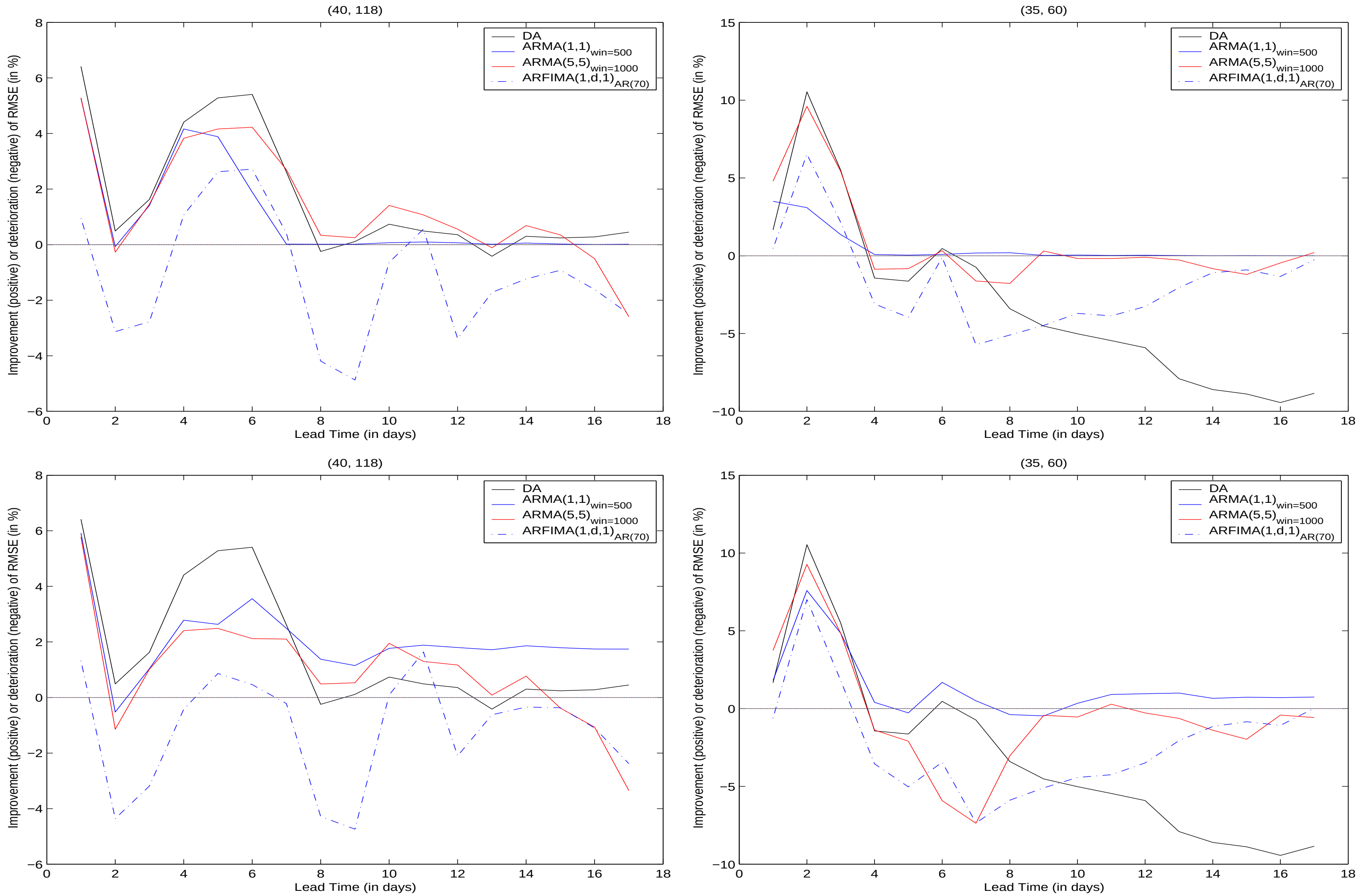
Models Predictive Power: the plots of the ratios of the Raw and Corrected Root Mean Square Errors (RMSE) $1 - \frac{\mathrm{Corrected RMSE}}{\mathrm{Raw RMSE}}$ with (top) and without (bottom) the demeaning step. Positive values imply improvement, negative ratios imply deterioration.



Models Predictive Power: the plots of the ratios of the Raw and Corrected Root Mean Square Errors (RMSE) $1 - \frac{\mathrm{Corrected RMSE}}{\mathrm{Raw RMSE}}$ with (bottom) and without (top) the demeaning step. Positive values imply improvement, negative ratios imply deterioration.



Models Predictive Power: the plots of the ratios of the Raw and Corrected Root Mean Square Errors (RMSE) $1 - \frac{\mathrm{Corrected RMSE}}{\mathrm{Raw RMSE}}$ with (bottom) and without (top) the demeaning step. Positive values imply improvement, negative ratios imply deterioration.



Models Predictive Power: the plots of the ratios of the Raw and Corrected Root Mean Square Errors (RMSE) $1 - \frac{\mathrm{CorrectedRMSE}}{\mathrm{RawRMSE}}$ with (bottom) and without (top) the demeaning step. Positive values imply improvement, negative ratios imply deterioration.

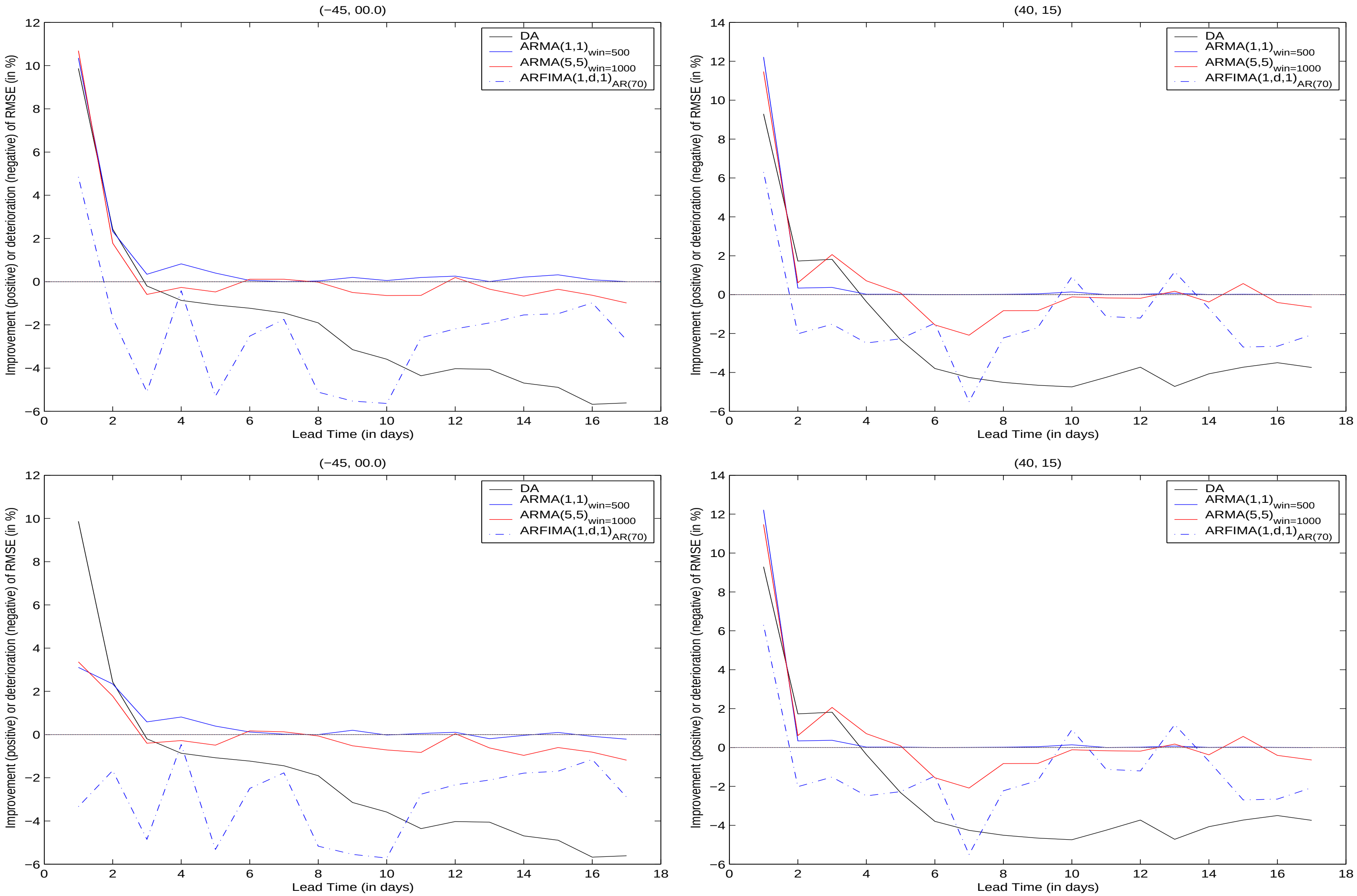
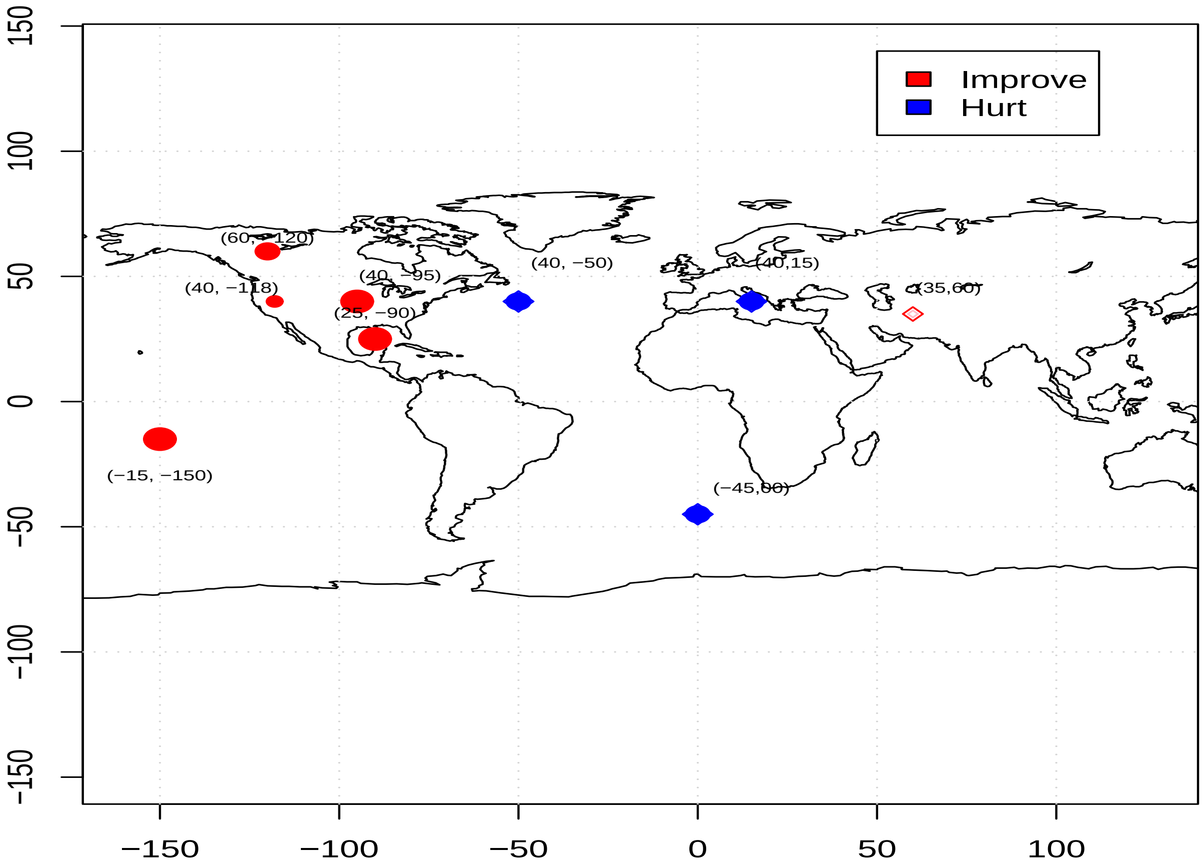


Figure 12: Results on reduction of RMSE for 9 locations. (Red color implies improvement over the raw RMSE and DA. Blue color implies no improvement over the raw RMSE.)



CONCLUSION

- Biases at all 9 location and all lead times exhibit a long memory pattern of various degrees.

- There exists a remarkable change (drop) in the pattern of the estimated d around lead time 5 days for most locations.

- DA method performs well up to around lead time 5 days and then significantly degrades, e.g. even deteriorates RMSE instead of improving it. (*Is there any connection between the change of* d *and degrading of DA around day 5?*)

- ARMA, ARFIMA models without a smoothed demeaning show similar performance for short lead time and are generally close to DA results. For longer lead times they provide no improvement and no harm (with some minor deviations from location to location).

- In contrast, the two stage procedure (demeaning + ARMA) provides a noticeable improvement for high lead times for locations in the North America.