Experimental results of price prediction methods for Spot instances in Cloud computing

Zhicheng Cai[a,*], Xiaoping Li[b], Rubén Ruiz[c], Qianmu Li[a]

[a]*School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, 210094, China*
[b]*School of Computer Science and Engineering, Southeast University, Nanjing, China*
[c]*Instituto Tecnológico de Informática, Acc. B. Universitat Politècnica de València, València, Spain.*

Abstract

This file consists of the means plots of the Mean Absolute Percentage Error (*MAPE*) with 95% confidence intervals in all virtual machine types, details of the three main hypotheses check, ANOVA results and Tukey multiple comparisons.

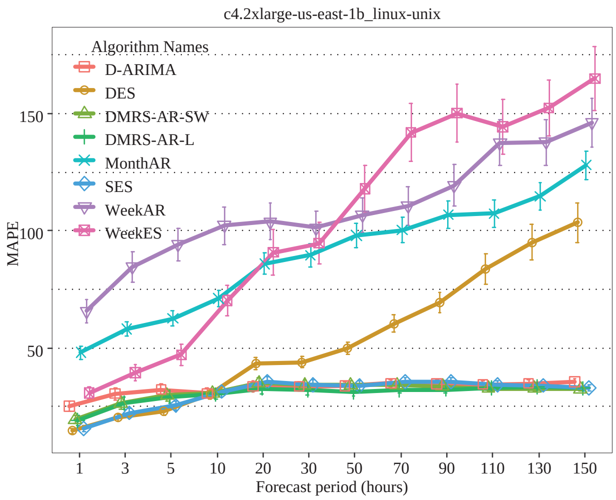
1. Means plot of the Mean Absolute Percentage Error

Spot prices in one hundred VM types have different statistical characteristics, such as means, trends, seasonality and linearity. According to these characteristics, VM types are categorized into five classes. Performances of forecast algorithms are evaluated on five classes of Spot prices. The first class of Spot prices contain two types of linear process and the switching between different processes is gentle. The second class of prices is composed of a type of linear process and different nonlinear processes with similar means. In the third class of Spot prices, there are more than two types of linear process with similar means and different variances. Long linear processes with significantly different means compose the forth class of Spot prices. Spot prices in the last class contain many different short linear processes with different means. Means plots of the Mean Absolute Percentage Error (%) with 95% confidence intervals in five classes of Spot prices are shown in Figures 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 and 11.

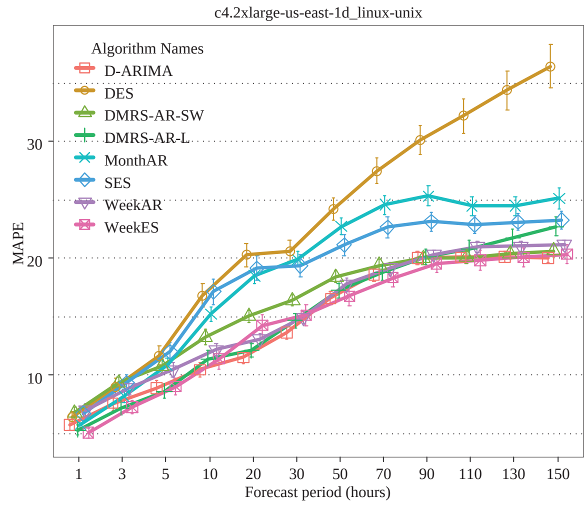
2. Analysis of variance

Experimental results are also analyzed by the analysis of variance (ANOVA) method [1]. The three main hypotheses (normality, homoscedasticity and independence of the residuals) are checked. Numerical tests are usually very strict. For example, numerical tests will normally reject the hypothesis that the data comes from a normal distribution (a numerical test can fail if one generates 100 data with a normal distribution and slightly changes one single data point). Studies support the fact that ANOVA is very robust to the three hypotheses since the sample sizes in each group of this paper are equal and large [2, 3]. McDonald et al. [3] suggested that heteroscedasticity will have a great impact on ANOVA for a balanced design only if one standard deviation is at least three times the size of the other, and the sample size in each group is fewer than 10. Therefore, graphical tests are commonly used to test the three hypotheses in practice. In this paper, normal QQ plots of *MAPE* residuals (Figure 12), *MAPE* residual plots vs. each factor level (Figure 13) and dispersion plots of *MAPE* residuals over run numbers (Figure 14) are used to test the three main hypotheses respectively. According to these graphs, most points of QQ plots are near the straight line, different algorithms have similar variances for most cases and the residuals over run numbers are like white noise. Therefore, the three main hypotheses are acceptable in most cases. If an algorithm has an extremely different variance compared with those of the other algorithms on some Spot prices, the algorithm is excluded when doing ANOVA in these scenarios. For each forecast period, ANOVA is performed to prove whether there are significant differences among different forecast algorithms. For the example of ”c4.2xlarge-us-east-1c-linux-unix” with the forecast period equal to 2 hours, $p = 5.13e-06$ means that there are significant differences among forecast algorithms as shown in the first subfigure of Figure 15. Then, Tukey multiple comparisons of means are used to recognize the difference between each pair of forecast algorithms. The second subfigure of Figure 15 shows differences of means with 95% family-wise Tukey confidence levels which illustrates that DMRS-AR-SW is significantly better than D-ARIMA, SES and so on. Differences of means with 95% family-wise Tukey confidence levels for other Spot prices with different forecast periods are shown in Figures 16, 17, 18 and 19.

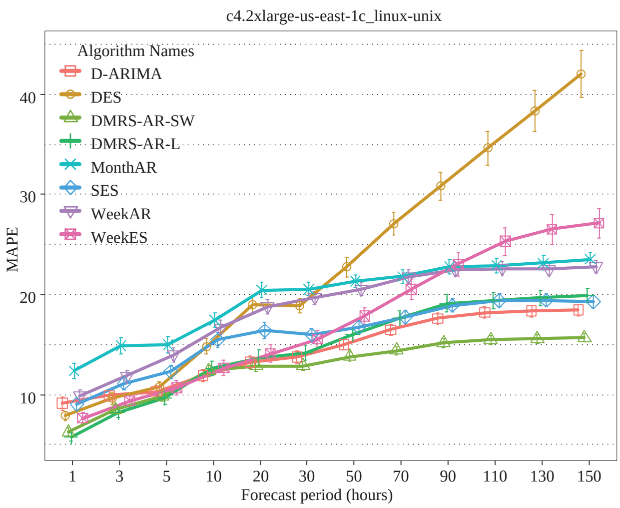
*Corresponding author.
Email addresses: caizhicheng@njust.edu.cn (Zhicheng Cai), xpli@seu.edu.cn (Xiaoping Li), rruiz@eio.upv.es (Rubén Ruiz), qianmu@njust.edu.cn (Qianmu Li)



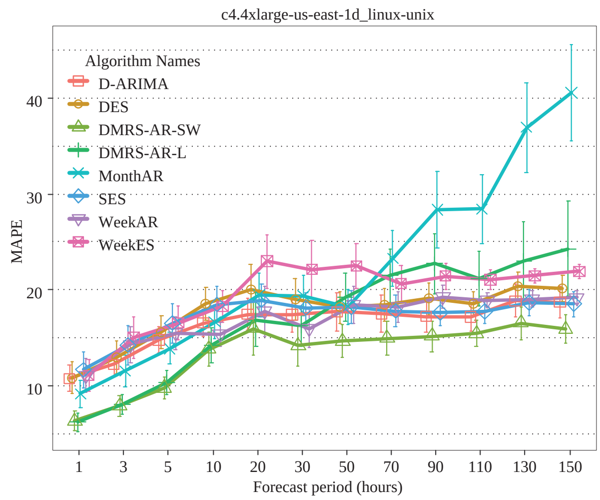
(a)



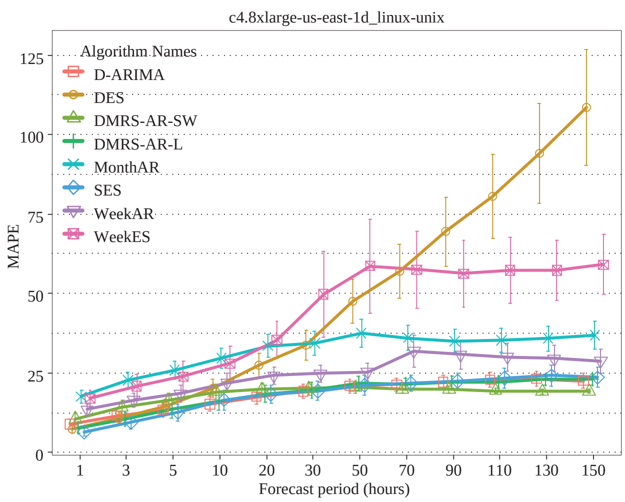
(b)



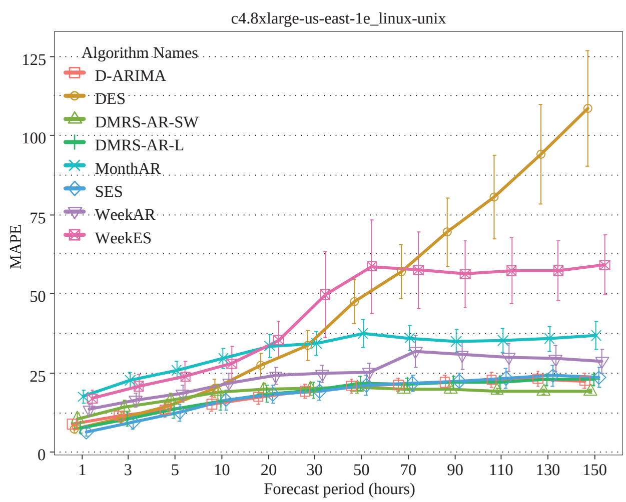
(c)



(d)

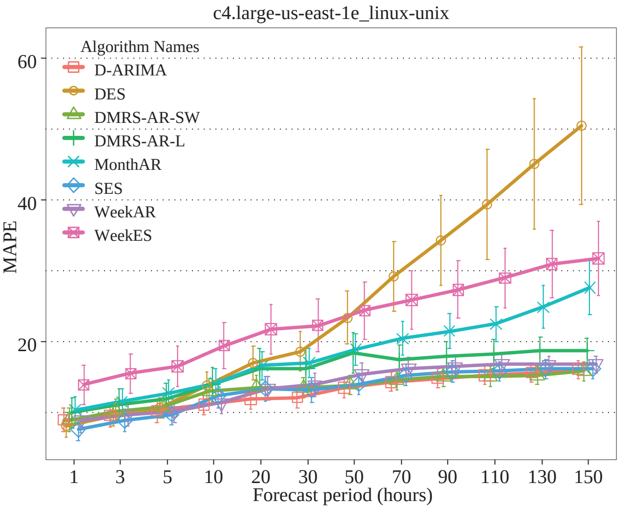


(e)

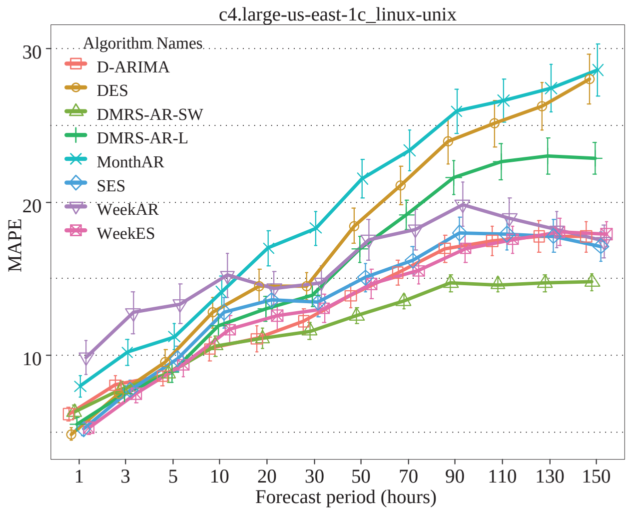


(f)

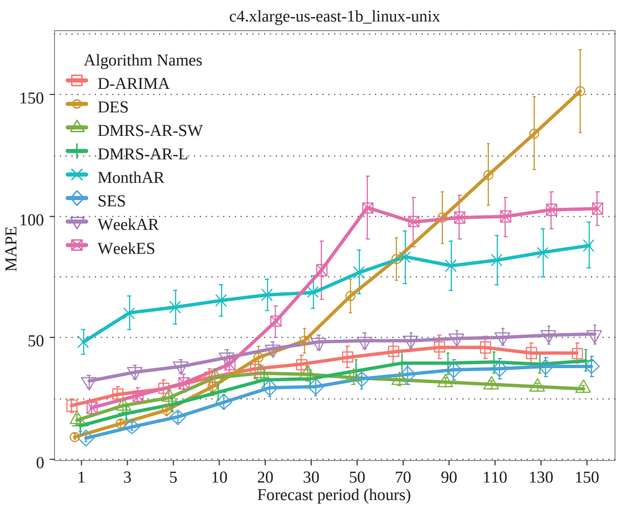
Figure 1: Means plot of the Mean Absolute Percentage Error (MAPE) with 95% confidence intervals on the first class Spot prices.



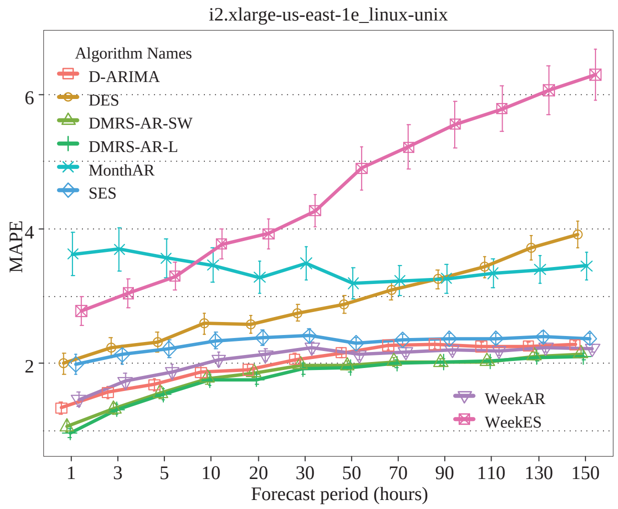
(a)



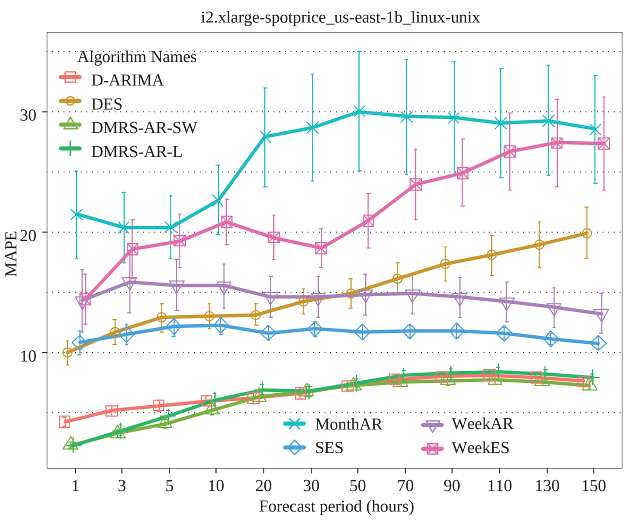
(b)



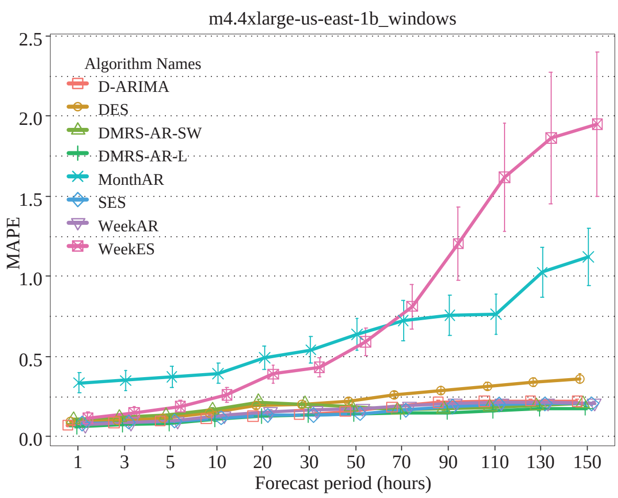
(c)



(d)

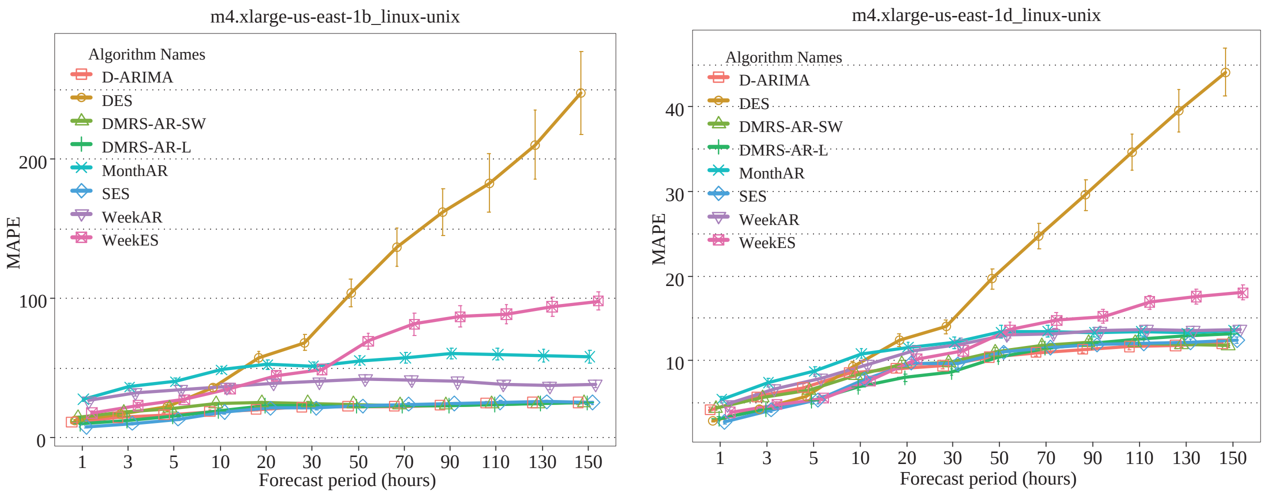


(e)



(f)

Figure 2: Mean Absolute Percentage Error (MAPE) on the first class Spot prices.



(a) (b)

Figure 3: Mean Absolute Percentage Error (MAPE) on the first class Spot prices.

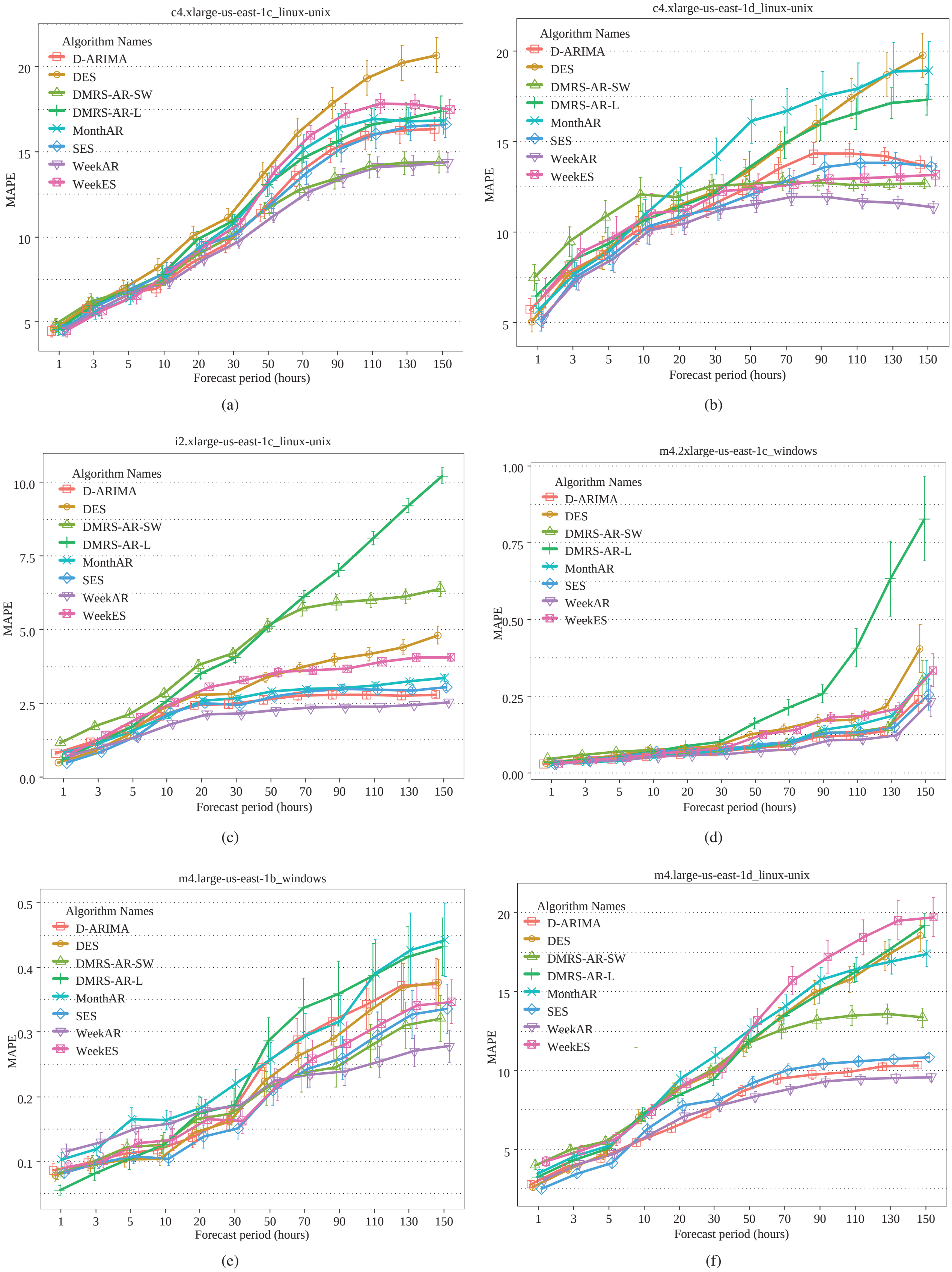
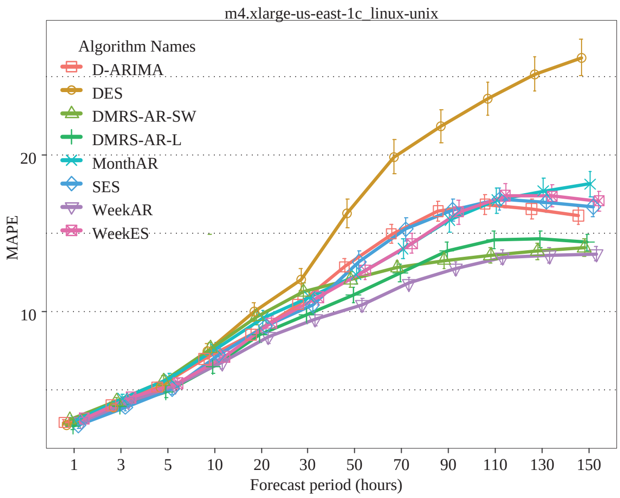
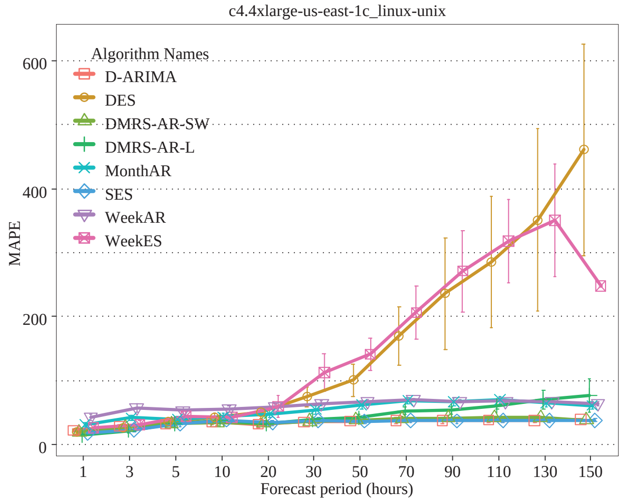


Figure 4: Mean Absolute Percentage Error (MAPE) on the second class Spot prices.

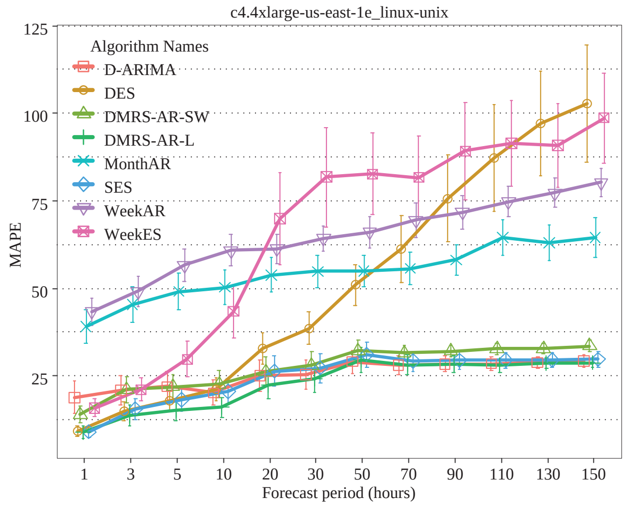


(a)

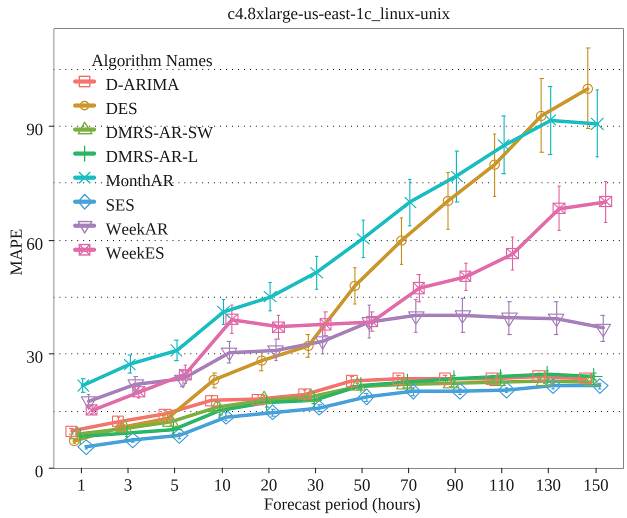
Figure 5: Mean Absolute Percentage Error (MAPE) on the second class Spot prices.



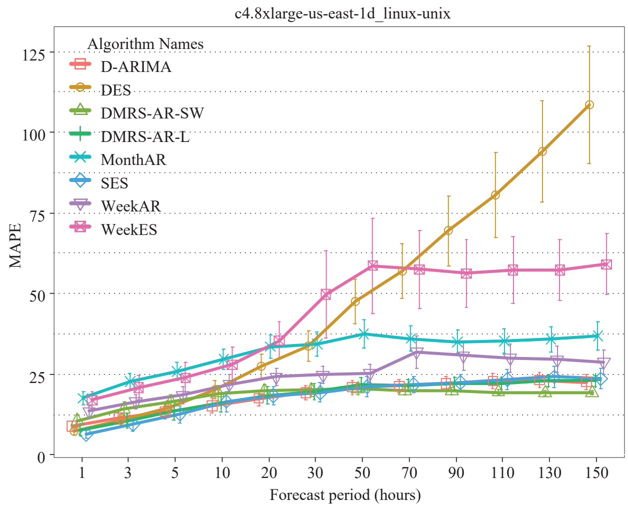
(a)



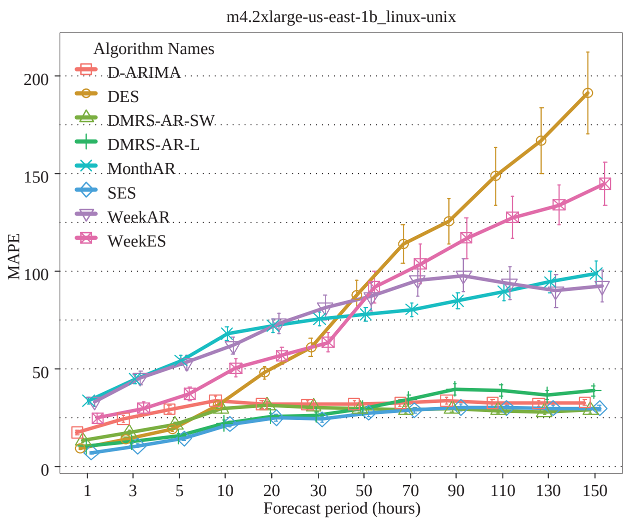
(b)



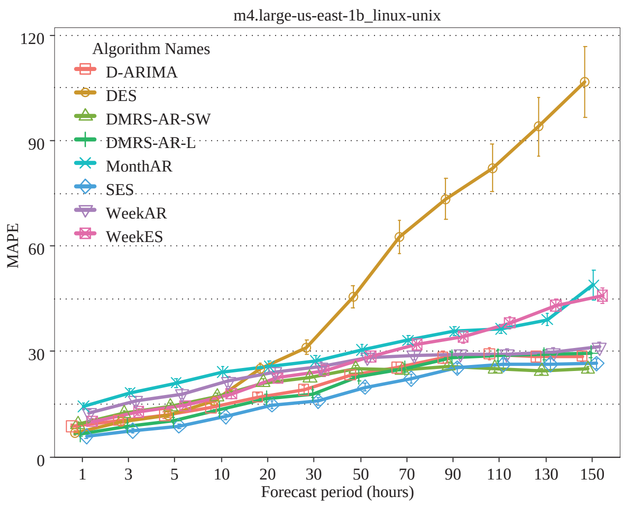
(c)



(d)

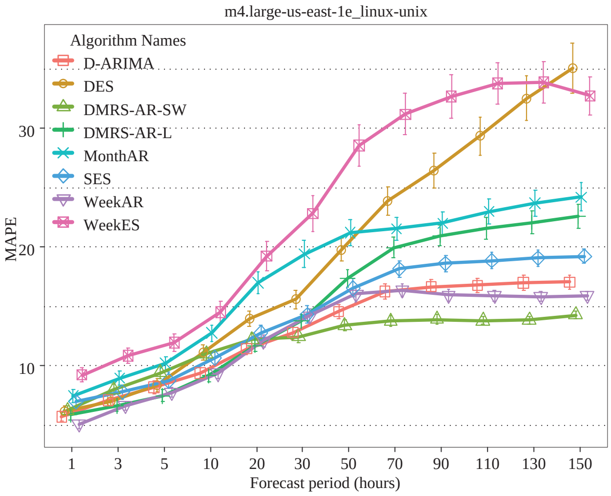


(e)

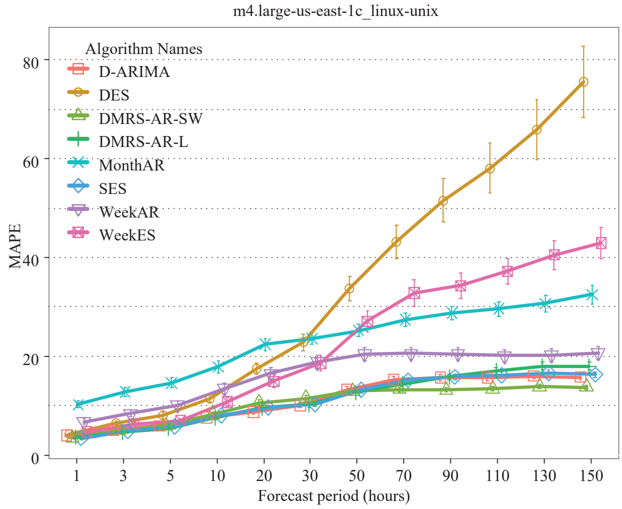


(f)

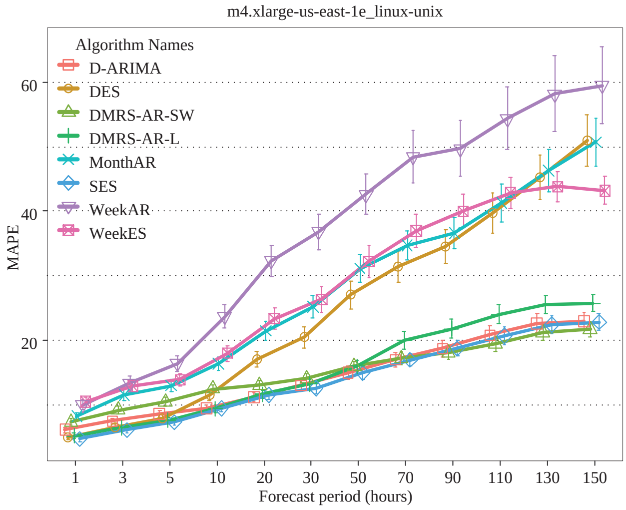
Figure 6: Mean Absolute Percentage Error (MAPE) on the third class Spot prices.



(a)



(b)



(c)

Figure 7: Mean Absolute Percentage Error (MAPE) on the third class Spot prices.

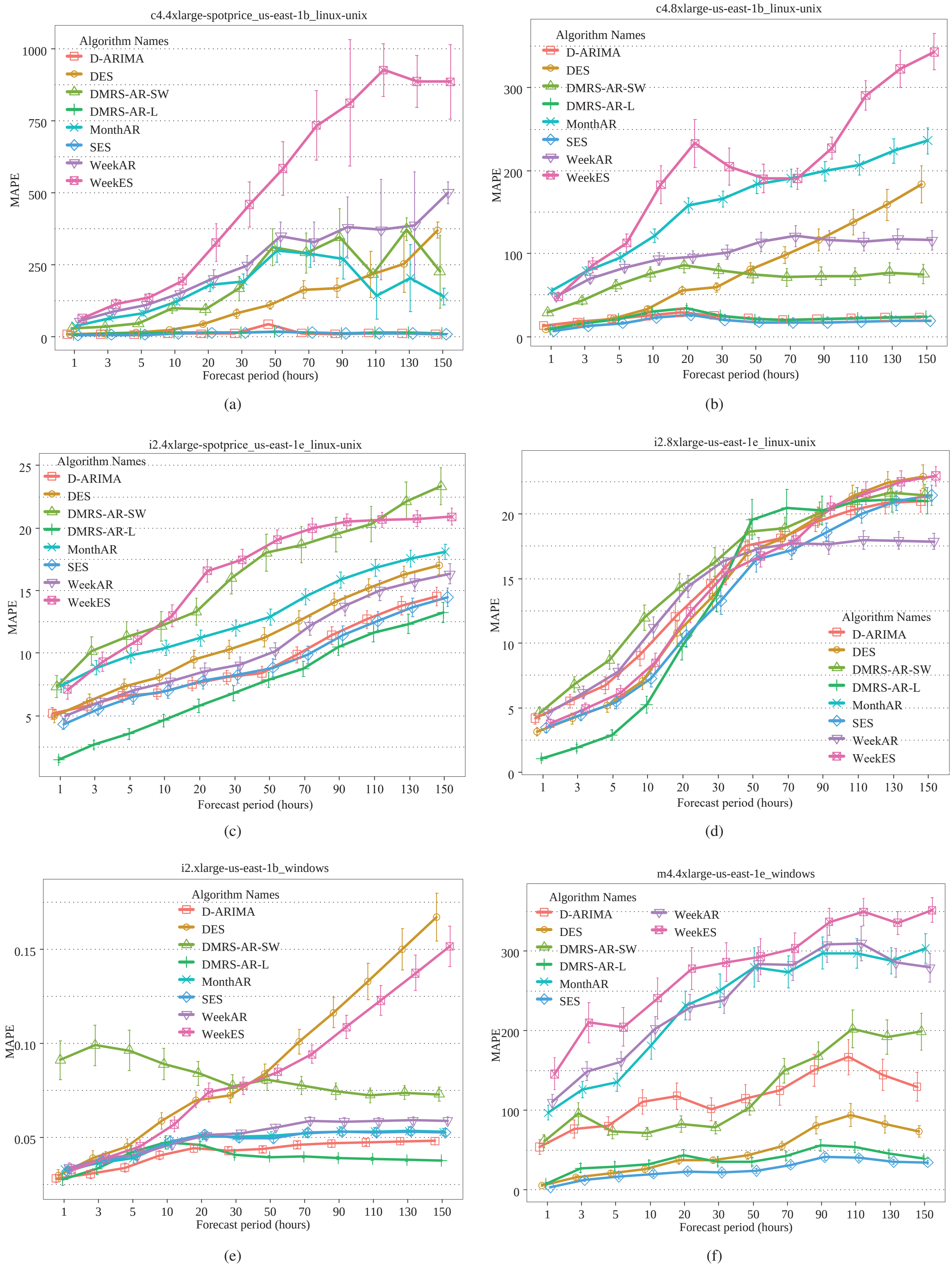
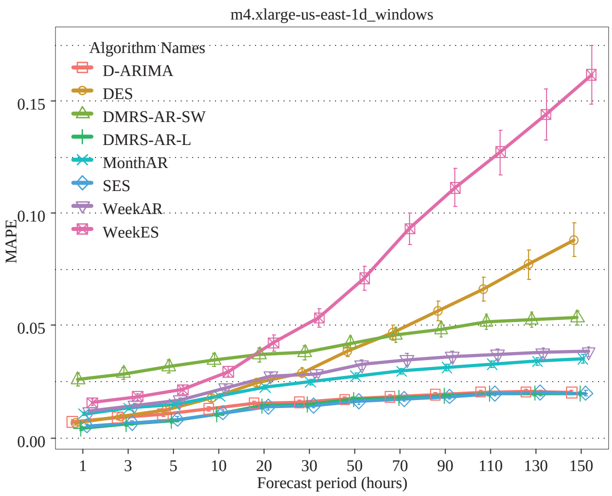


Figure 8: Mean Absolute Percentage Error (MAPE) on the forth class Spot prices.



(a)

Figure 9: Mean Absolute Percentage Error (MAPE) on the forth class Spot prices.

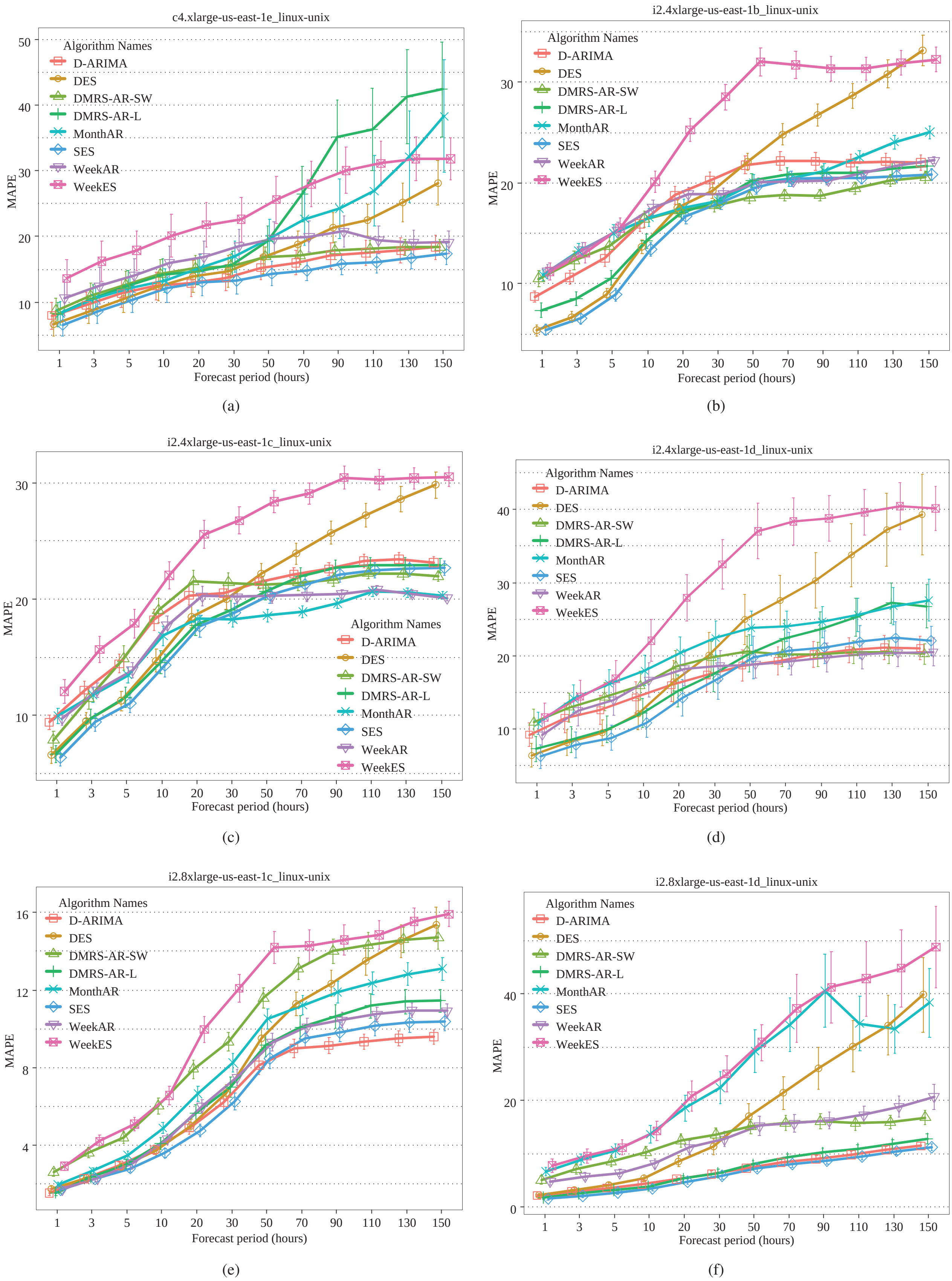
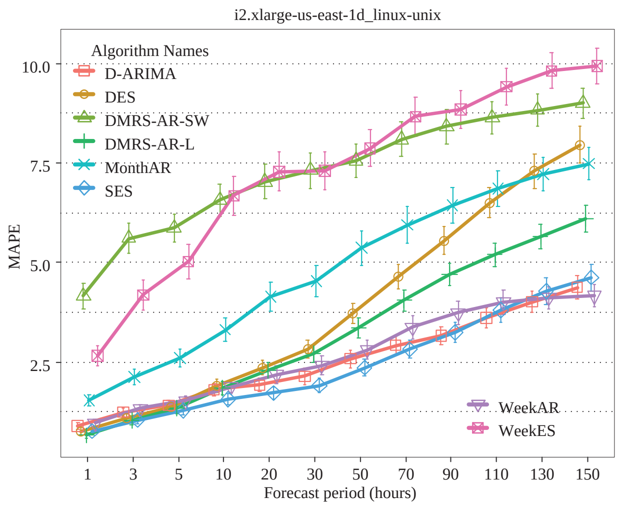
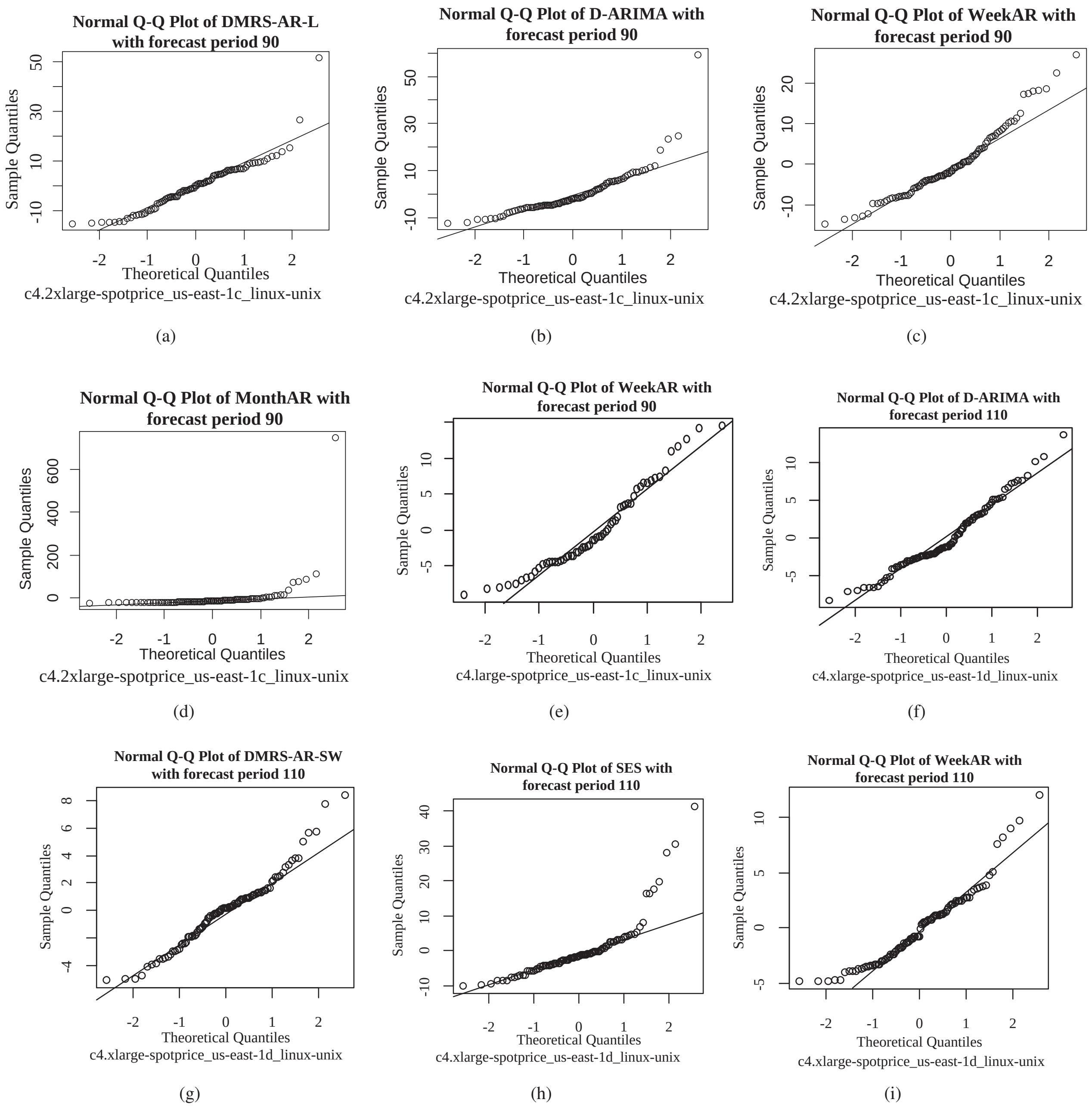


Figure 10: Mean Absolute Percentage Error (MAPE) on the fifth class Spot prices.



(a)

Figure 11: Mean Absolute Percentage Error (MAPE) on the fifth class Spot prices.



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

Figure 12: Normal QQ plots of residuals for the *MAPE* of different algorithms on diverse Spot prices.

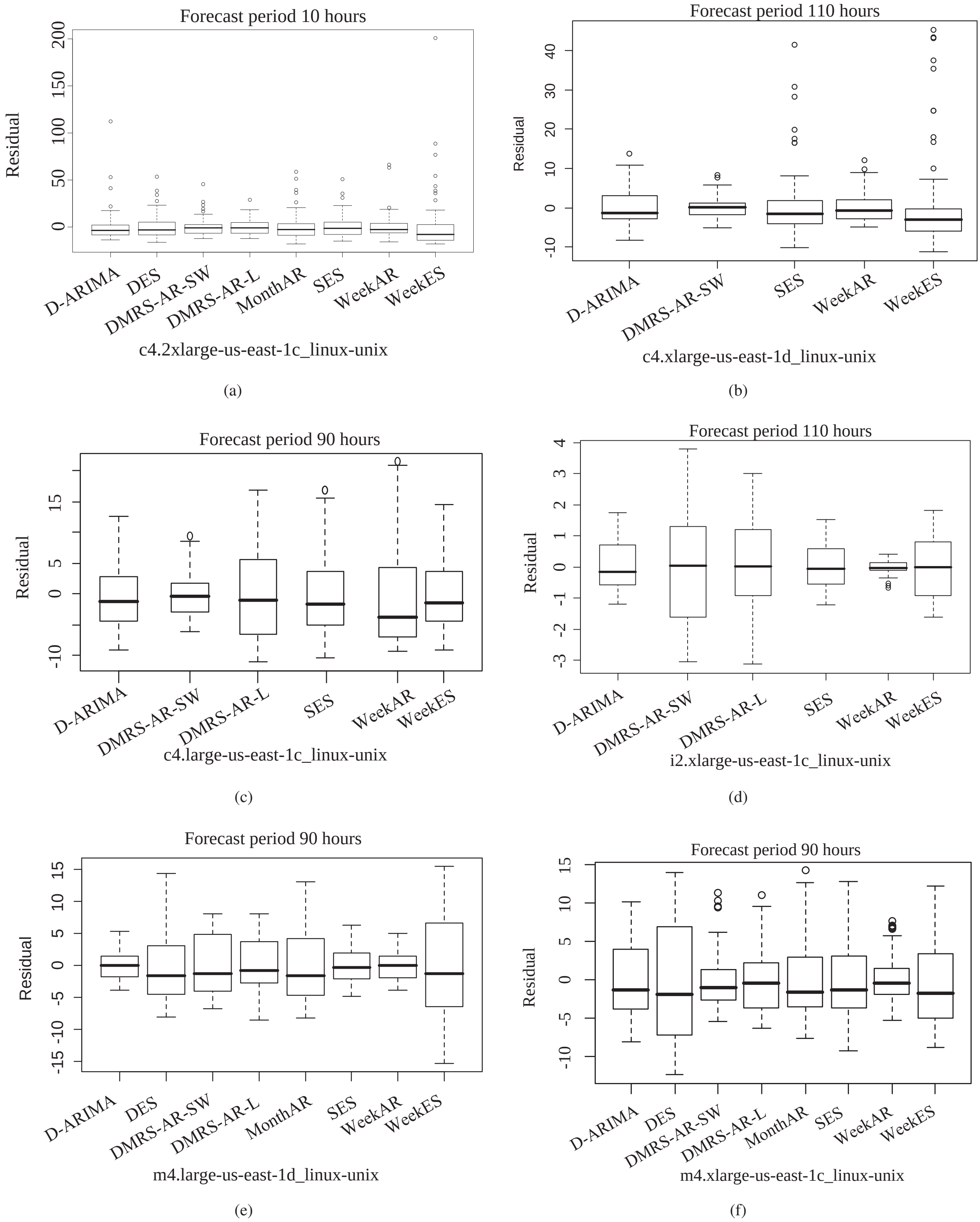


Figure 13: Residual plots of *MAPE* vs. different algorithms on diverse Spot prices.

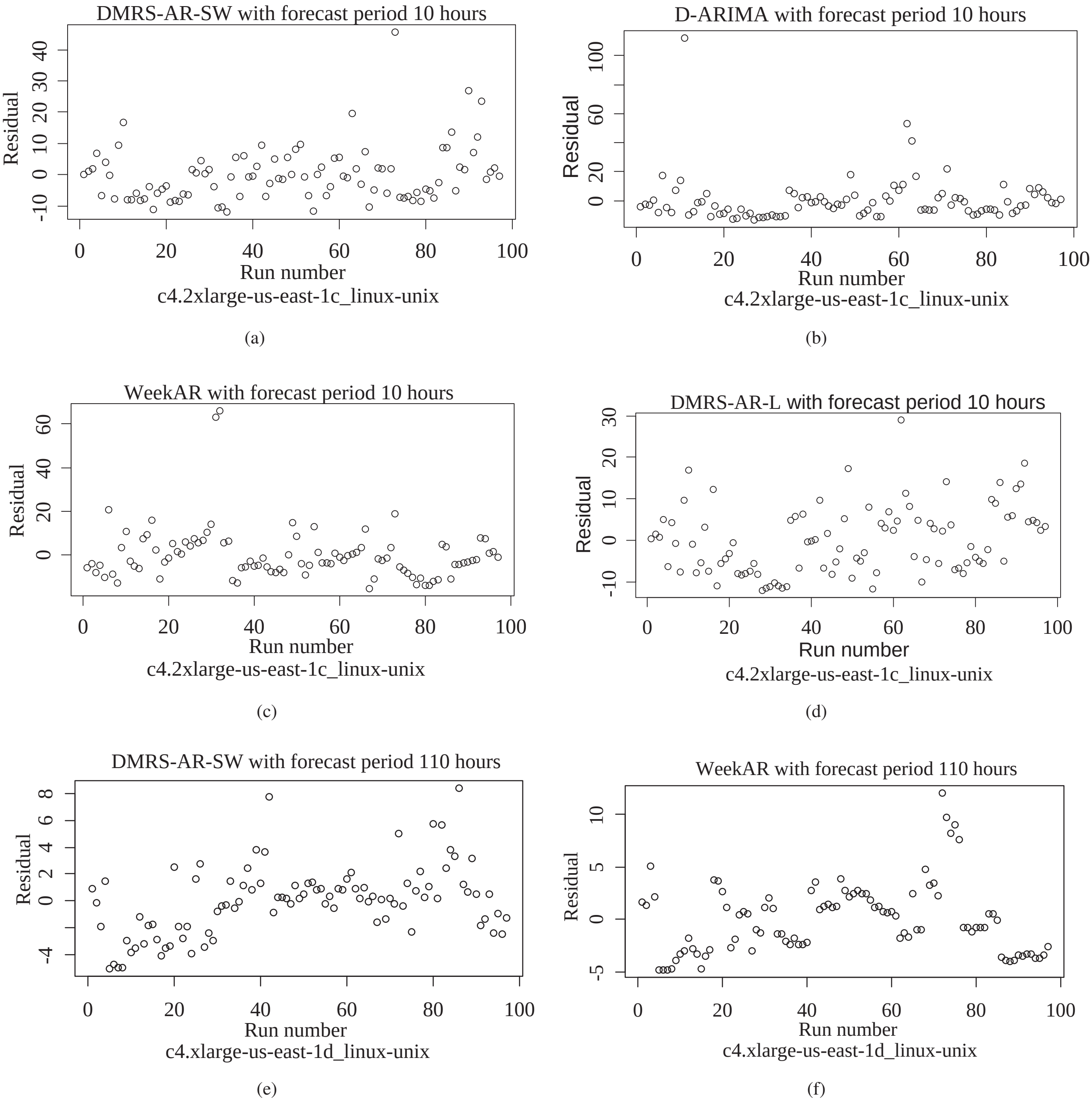


Figure 14: Dispersion plots of residuals over run numbers for *MAPE* on diverse Spot prices.

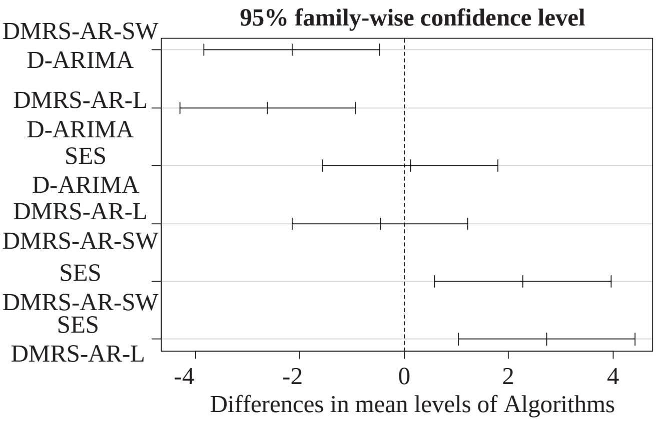
```
summary(fm1 <- aov(MAPE~ Alname, data = subrc))
           Df Sum Sq Mean Sq F value   Pr(>F)
Alname       3    467   155.8     9.5 5.13e-06 ***
Residuals  304   4987    16.4
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> TukeyHSD(fm1, "Alname", ordered = F)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = MAPE ~ Alname, data = subrc)

$Alname
                                diff        lwr        upr     p adj
DMRS-AR-SW – D-ARIMA       -2.1543978 -3.8406218 -0.4681738 0.0059199
DMRS-AR-L – D-ARIMA        -2.6136796 -4.2999036 -0.9274556 0.0004529
SES – D-ARIMA               0.1142568 -1.5719672  1.8004808 0.9980912
DMRS-AR-L – DMRS-AR-SW     -0.4592818 -2.1455058  1.2269422 0.8956095
SES  – DMRS-AR-SW           2.2686546  0.5824306  3.9548786 0.0032583
SES  – DMRS-AR-L            2.7279364  1.0417124  4.4141604 0.0002232
```

c4.2xlarge-us-east-1c_linux-unix
with forecast period of 2 hours

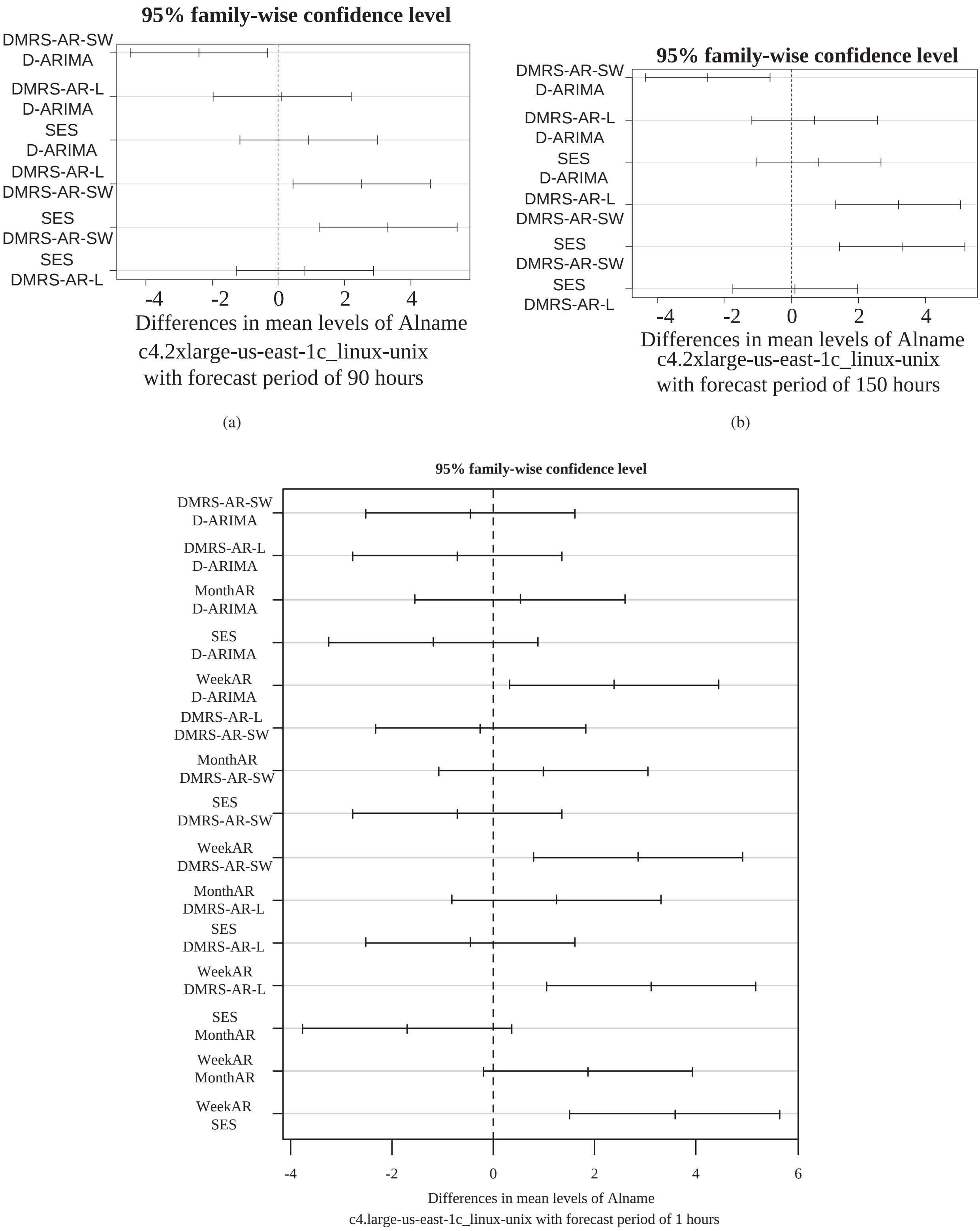
(a)



c4.2xlarge-us-east-1c_linux-unix
with forecast period of 2 hours

(b)

Figure 15: Results of ANOVA and Tukey multiple comparisons of means with 95% family-wise confidence level.

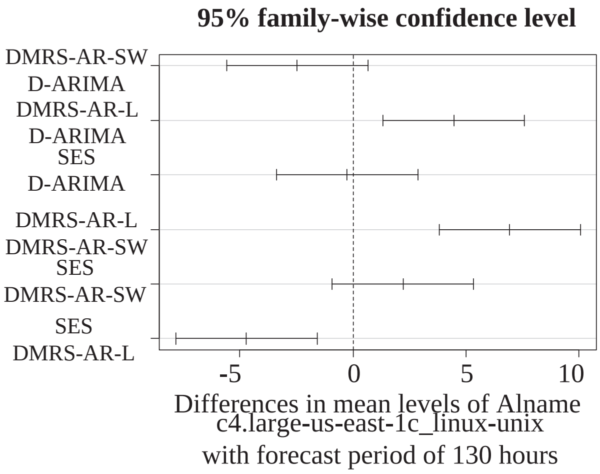


(a)

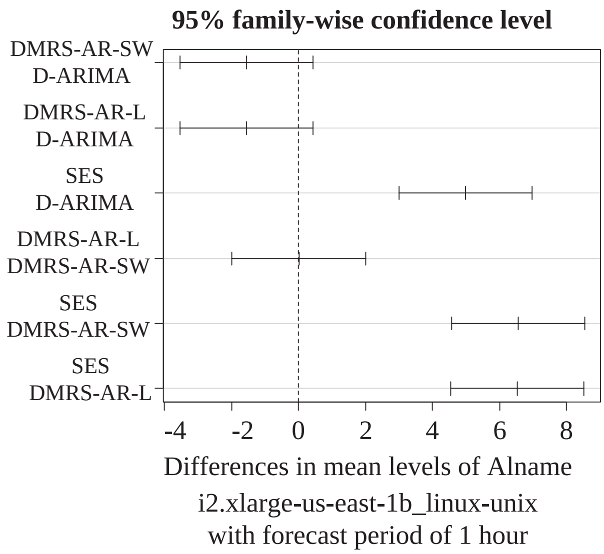
(b)

(c)

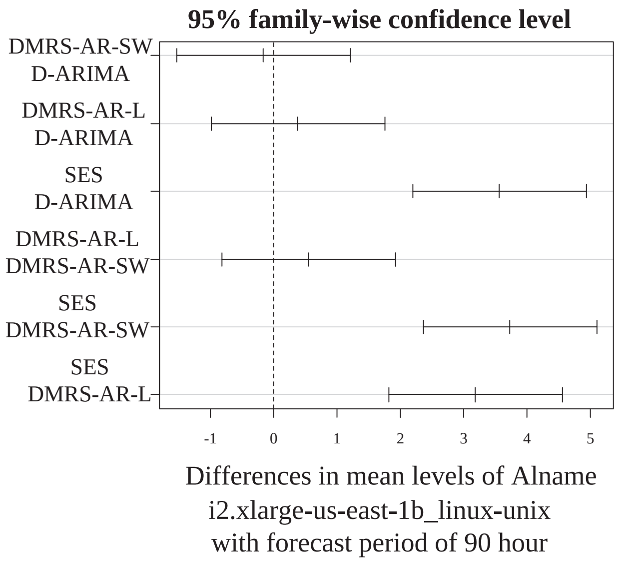
Figure 16: Tukey multiple comparisons of means with 95% family-wise confidence level.



(a)



(b)



(c)

Figure 17: Tukey multiple comparisons of means with 95% family-wise confidence level.

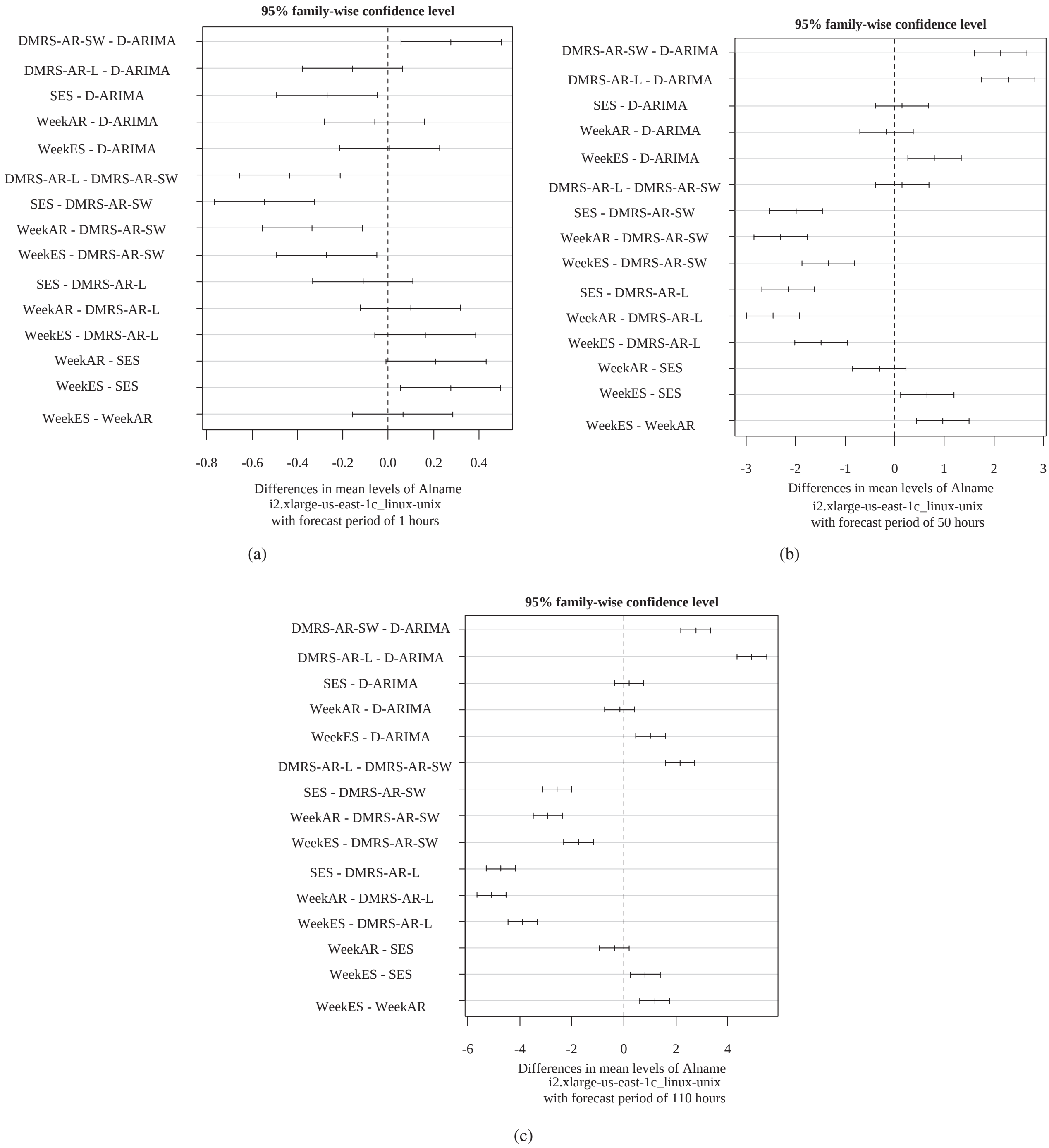
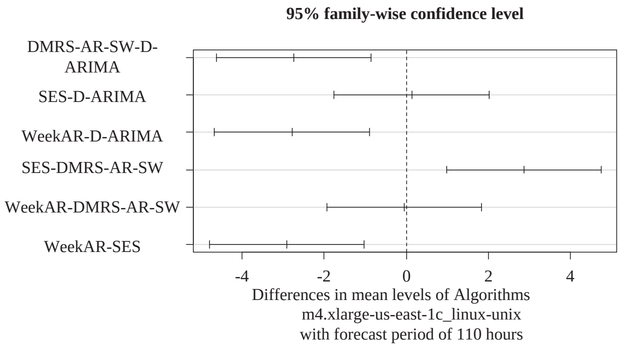
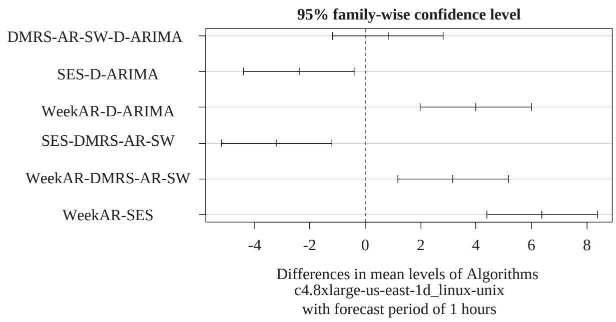


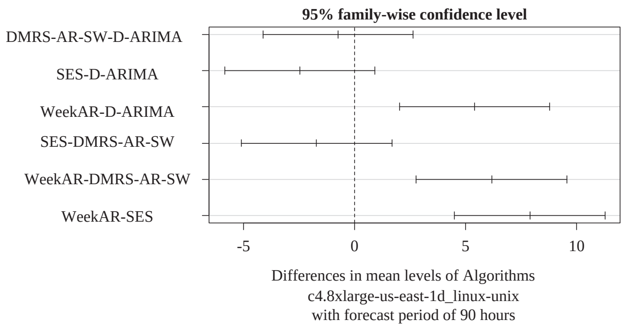
Figure 18: Tukey multiple comparisons of means with 95% family-wise confidence level.



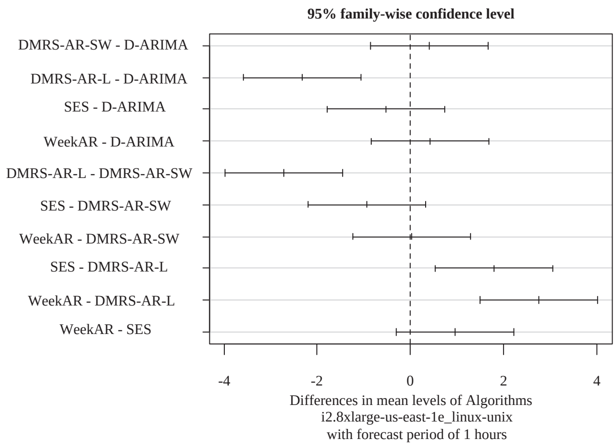
(a)



(b)



(c)



(d)

Figure 19: Tukey multiple comparisons of means with 95% family-wise confidence level.

References

[1] T. Bartz-Beielstein, M. Chiarandini, L. Paquete, M. Preuss, Experimental methods for the analysis of optimization algorithms, Springer, 2010.

[2] G. V. Glass, P. D. Peckham, J. R. Sanders, Consequences of failure to meet assumptions underlying the fixed effects analyses of variance and covariance, Review of Educational Research 42 (3) (1972) 237–288.

[3] J. H. Mcdonald, Handbook of biological statistics, Sparky House Publishing, Baltimore, Maryland 145-156, `http://www.biostathandbook.com/onewayanova.html`, accessed, 2017.9.8.