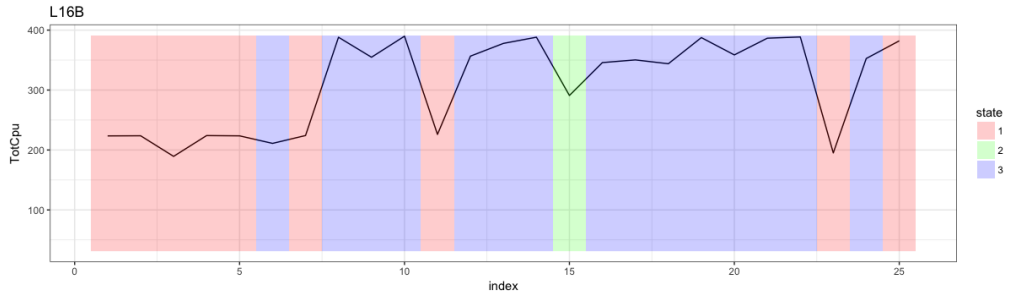

Figure 4.19.: The predicted state of the test set in the software release L16B

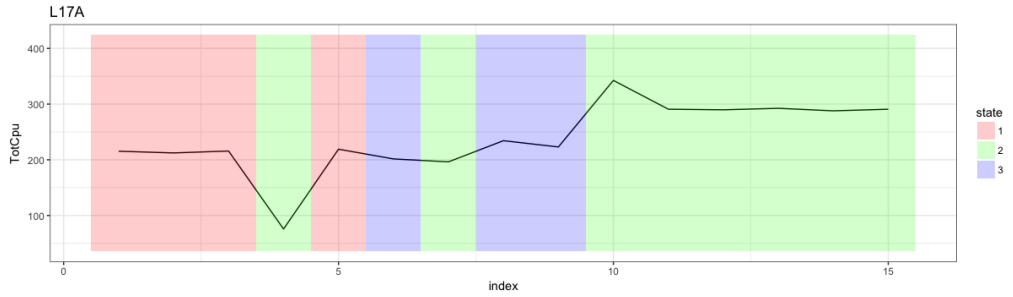

Figure 4.20.: The predicted state of the test set in the software release L17A

4.7. Model evaluation

Eighty percents of the observations from the simulated data were fitted with Markov switching autoregressive model, and the remaining was used as a test set to evaluate a performance of the model.

4.7.1. Simulated Dataset 1

Table 4.6 presents a model output of the Markov switching model of the simulated Dataset 1. Each state has a considerably high r-squared value which is greater than 0.9. It can be seen that State2 has the highest standard error. When inferring back to actual models that used to generate these three states in sec. 3.8, State1, State2, and State3 are a *Bad*, *Normal*, and *Good* state, respectively.

Table 4.6.: Output from the model of the simulated Dataset 1 showing estimated coefficients, residual standard error, and r-squared for each state. A switching coefficient is followed by (S), and a significant coefficient is highlighted in bold.

Estimated coefficient	State 1	State 2	State 3
(Intercept)(S)	**1.3024**	**5.5339**	**-14.3116**
x1(S)	**0.8005**	**0.5900**	**0.7005**
x2(S)	0.0003	**-0.8610**	**0.1941**
y_1(S)	**0.1902**	**0.3843**	**-0.1927**
Residual standard error	1.4900	7.3013	1.8255
r^2	0.9980	0.9467	0.9962

The result of the model performance using Dataset 1 is shown in Table 4.7. There are two observations from a *Bad* state which were incorrectly predicted to be in a Normal state. Moreover, two more observations from a *Good* state were predicted to be in a *Normal* state. The overall accuracy of the model is 0.96, and the misclassification rate is 0.04. One can see that the model was able to perfectly predict the state of the observations that are in a *Normal* state.

Table 4.7.: The confusion matrix of the result from the test set of the simulated Dataset 1

		Predicted state		
		Bad	Normal	Good
Actual state	Bad	58	2	0
	Normal	0	30	0
	Good	0	2	8

4.7.2. Simulated Dataset 2

The Markov switching model of the simulated Dataset 2 are presented in Table 4.8. The r-squared value of each state is less than the obtained result from the simulated Dataset 1, but the standard error in each state is higher. However, it could be said that the model still performs rather well as the r-squared value is fairly high. State1 is said to be a *Bad* state although an intercept is insignificant. State2 and State3 are a *Good* and *Normal* state, respectively.

Table 4.8.: Output from the model of the simulated Dataset 2 showing estimated coefficients, residual standard error, and r-squared for each state. A switching coefficient is followed by (S), and a significant coefficient is highlighted in bold.

Estimated coefficient	State 1	State 2	State 3
(Intercept)(S)	5.0716	**9.0219**	**23.1711**
x1(S)	**0.8672**	**0.5401**	**0.5546**
x2(S)	-0.0195	**0.2599**	**-1.1239**
y_1(S)	**0.1232**	**-0.0882**	**0.2272**
Residual standard error	6.6367	5.3211	7.5079
r^2	0.9089	0.9480	0.9337

Table 4.9 presents a confusion matrix for a test set from a second simulated dataset. The model was able to correctly predict all the observations in a *Bad* state. On the contrary, the model did not perform well in predicting observations which had *Good* and *Normal* state. Nine observations from a *Good* state were predicted to be in a *Bad* state, and another five observations from a *Good* state were inaccurately predicted to be in a *Normal* state. Six observations from a *Normal* state were incorrectly predicted to be in a *Good* state. The overall accuracy of the model and the misclassification rate are 0.8 and 0.2, respectively.

Table 4.9.: The confusion matrix of the result from the test set of the simulated Dataset 2

		Predicted state		
		Bad	Normal	Good
Actual state	Bad	35	0	0
	Normal	0	29	6
	Good	9	5	16