Data

TIME IS IN HOURS

The data used consists of six cycles. This means that the total data set consists of six simulation runs. However, each of the runs contains many batches in sequence. The following table lists the number of batches in each of the simulations and some basic statistics

Cycle	#batch	μ_{batch}	σ^2_{batch}
A	66	14.776	3.641
B	64	15.644	3.915
C	61	17.714	2.330
D	60	18.069	6.922
E	60	18.088	9.613
F	63	17.227	7.766

Table 1.1: Per cycle batch statistics

Each batch is comprised several states. These include adding materials (IDs 1 through 4), centrifugation (ID 5), product transfer (the precipitate generated from the centrifugation, ID 6), chemical reaction (ID 7), a post operation state (Probably to let it cool down to a point where it is ready for further processing, ID 8), Cooling of the product (ID 9), material transfer (transfer the gained prod-

uct before cleaning of the reaction vessel and/or prepare for the next reaction batch, ID 10).

1.1 Adding of solids and materials

This part of the process corresponds to events tagged with ID 1 through 4

1.2 Centrifugation

ID 5

1.3 Product transfer

ID 6

1.4 Reaction

ID 7

1.5 Post reaction

ID 8

1.6 Cooling

ID 9

1.7   Material transfer

ID 10

1.8   Cleaning operations

Sometimes, the vessel is cleansed. This is however not every time after a batch so might be interesting to investigate further. Initially, per cycle, the cleanings are summarized in the following table with basic statistics. As can be seen, there is quite some differences.

The most notifiable differences per batch are the number of cleanses especially when comparing to Table 1.1. For the first two cycles, the cleanses seem to be in between every batch, which is indeed also the while the later four are only sometimes. Furthermore, although the cleanses are between every batch for cycles A and B, the variances are extremely different. For the last four cycles, they seem to be grouped further, E and F are very alike while cleanses in C and D are generally longer although D has a substantially smaller variance than C.

Cycle	#ops	min	max	μ	σ^2
A	65	1.113	3.067	1.917	0.269
B	63	1.324	1.751	1.566	0.00883
C	9	1.544	3.306	2.153	0.277
D	10	1.474	2.009	1.581	0.0212
E	10	0.827	1.584	1.465	0.0462
F	10	0.748	1.610	1.466	0.0595

Table 1.2: Per cycle cleansing statistics

To verify these observations and potentially discovering more important facts of their probability distributions, histograms are plotted in the following Figure 1.1. We indeed again observe the likeliness between the cycles A and B, C and D, E and F respectively. Also, for the first two cycles and more so cycle B, the cleaning times are somewhat normally distributed although cycle A has a very heavy right tail in that case. The later four cycles only have 10 observations but the mode (i.e. peak) seem to be about the same.

From the above observation of like modes one may want to observe the histogram of the combined set of cleaning times. In particular, under the hypothesis that the durations are actually from the same probability distributions and realized independently within each cycle a histogram of all the observations are of interest and is shown in Figure 1.2 below.
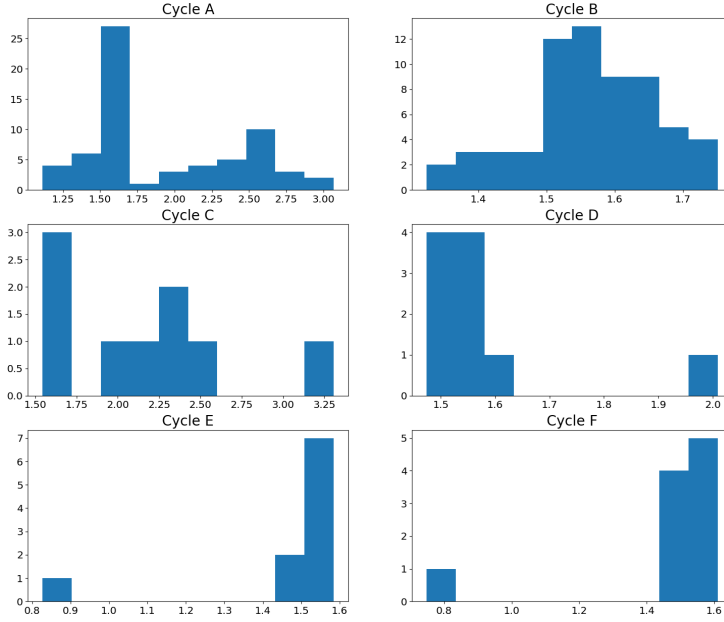
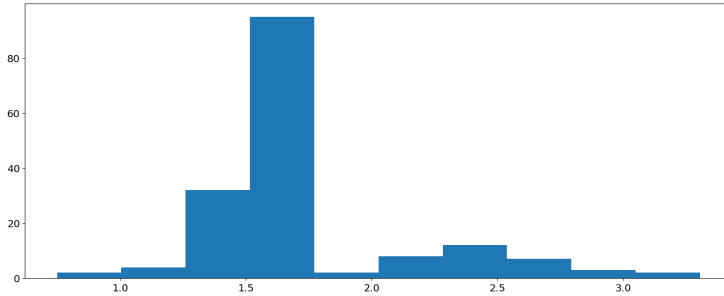**Figure 1.1:** Each of the 6 cycles, cleaning operations histograms.

Figure 1.2: Combined cleaning operations histograms.

Finally, to get a better overview of the irregularities is the number of clean-
ing periods (mostly concerning cycles C through F), each cleaning operation is
shown in the following Figure 1.3. The vertical shaded rectangles signify the
period in which a cleaning operation is taking place. Furthermore, the event
IDs are shown but to get a clearer view on what is going on, a single rectangle
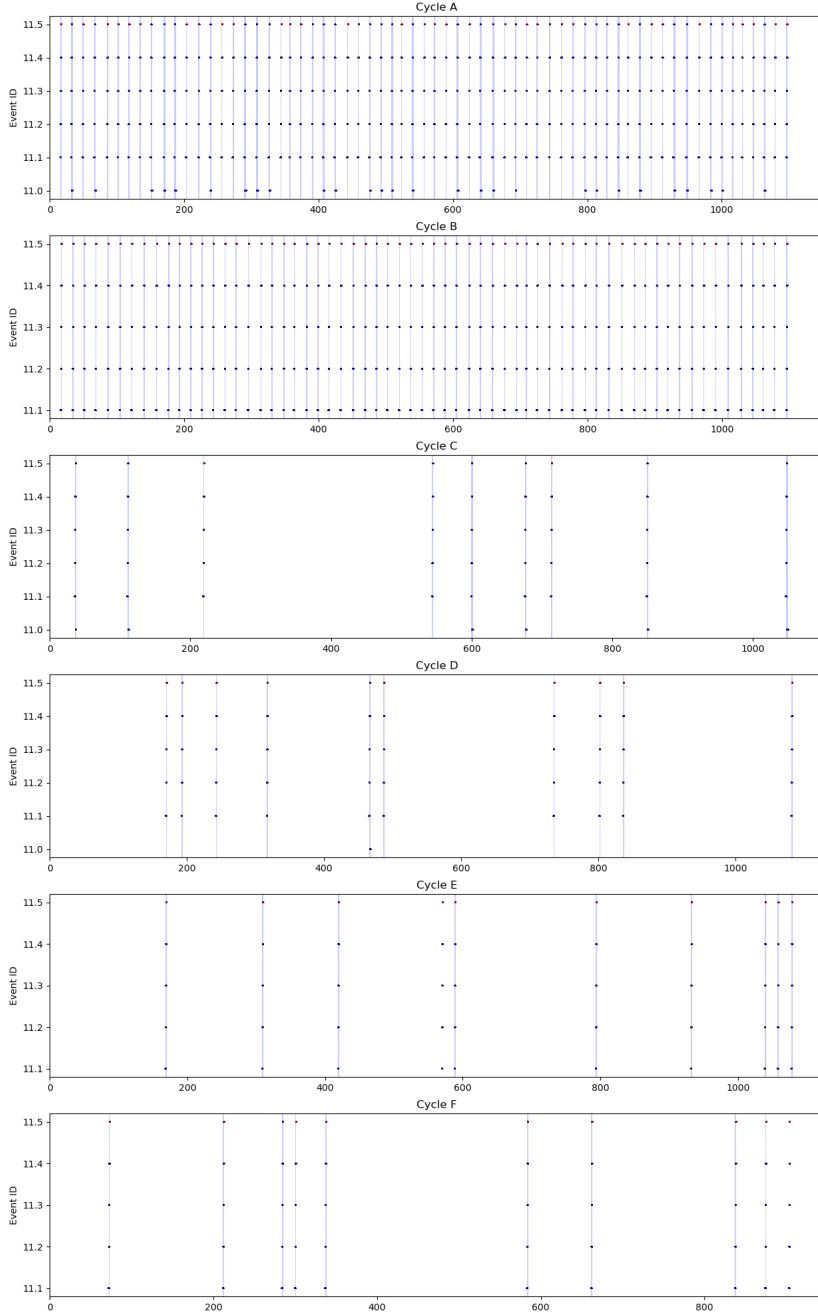(zoomed in) is shown in Figure 1.4.

Figure 1.3: Each of the 6 cycles, cleaning (corresponding to `BatchID = 0`). Each (Cleaning Procedure), CIP, is highlighted with an opaque interval (the blue rectangles). The dots marked with red (only ID 11.5, but not all of these are red), is if the Cleaning ID is 0.
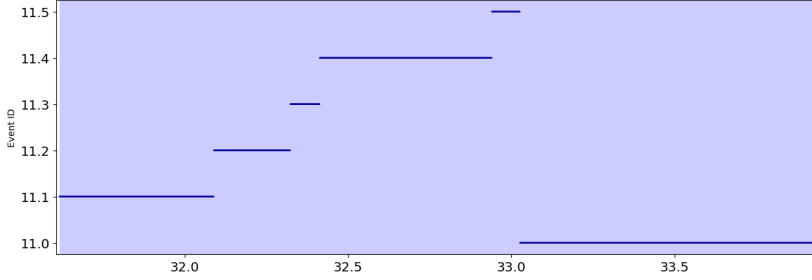
Figure 1.4: A single blue rectangle zoomed in

It is observed that the observations marked with red in figure 1.3 occur exactly when that specific cleaning operation does not go to the state 11.0 after the flush of the tank (event ID 11.5) and vice versa. It is hard to conclude what this may mean, but the cleaning being in state 11.0 may indicate that the system is idle before continuing the next batch like what is observed from the other steps of the process flow. Also, it is noted that while the red dots occur nothing else is happening according to the dataset.

From a modelling point of view, the cycles C through F can be thought of as the cleansing operation having a probability of not happening or equivalently as having a duration of 0. It is thus of interest to observe what the probability of cleaning after an operation is. From Table 1.1 and Table 1.2, we that indeed for cycles A and B, the probability is 100 % when disregarding the possibility of cleaning after the final batch. Hence, we see that for the remaining cycles, the probabilities of cleaning the tank after an operation are as in Table 1.3

Cycle	% cleaning
A	100.00
B	100.00
C	15.00
D	16.95
E	16.95
F	16.13

Table 1.3: Per cycle probability of cleaning

Furthermore, let C_i denote whether the ith batch is followed by a cleaning of the tank or not. It is then of interest if the next batch is followed by a cleaning given whether the current batch is followed by a cleaning. In particular, we count for each of the cycles the transitions which are shown in the following tables. Notice that the number of observations is two less than the total number of batches within each specific cycle. This is due to the last batch is never followed by

a cleaning (nor is the first batch superseded by a cleaning procedure) which results in one less observation and also due to the fact that we are logging transitions and hence lose another observation. To test for randomness, a Chi-squared test is carried out on each of the cycles to check for independence. It is observed all the cycles exhibit independence between the groups i.e. there is no statistical evidence for information is gained about if the next batch is followed by a cleaning operation given whether the current batch is followed by a cleaning operation.

C_i \ C_{i+1}	No	Yes
No	41	9
Yes	9	0

(a) C, $p = 0.3293$

C_i \ C_{i+1}	No	Yes
No	41	8
Yes	7	2

(b) D, $p = 0.6456$

C_i \ C_{i+1}	No	Yes
No	41	7
Yes	7	3

(c) E, $p = 0.3532$

C_i \ C_{i+1}	No	Yes
No	41	9
Yes	9	1

(d) F, $p = 1.0000$

Table 1.4: Contingency table for Cycle C-F

Thus collecting the observations from all the last four cycles, we may want to model the atom of the cleaning procedure independently of the previous batch and with a probability of 0.8375 corresponding to the cleaning procedure only being carried out $16, 25\%$ of cases.

Finally, we show the autocorrelation function for each the four cycles C-F in Figure 1.5 and note that all the ACF stay within the 95% confidence interval.
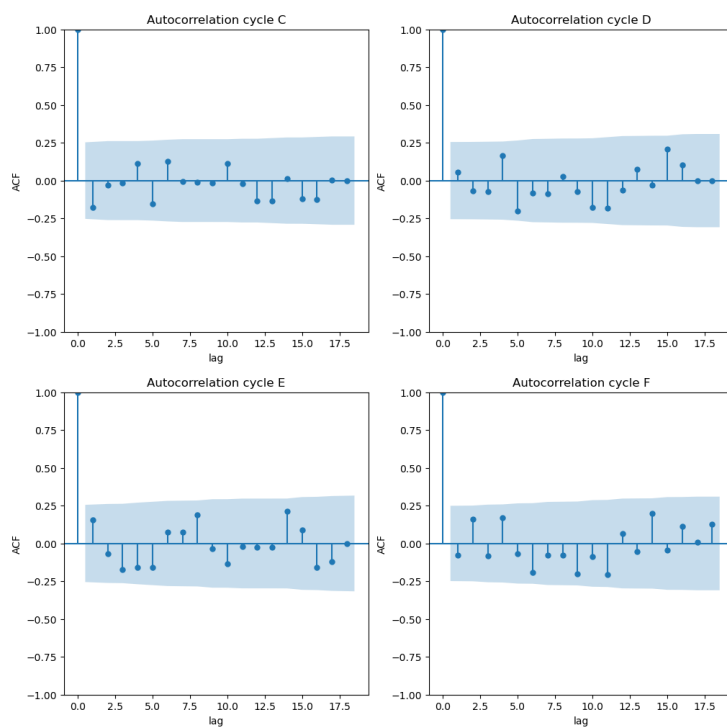
Figure 1.5: Autocorrelation function for each of the final 4 cycles. As can also be seen from this there seem to be no information to be gained of C_i from C_{i-1}.