# The University of Auckland Department of Engineering Science Part IV Project

## Developing a Julia/JuMP Optimization Model for Plant Scheduling for an NZ Company

**Denis Plakic** 

Supervisor:

Dr. Andrew Mason

Project Partner:

**Jack Hollings** 

November 2022

#### **Abstract**

An unnamed NZ company has partnered with the University of Auckland to develop a spreadsheet model to evaluate possible schedules for a production plant. The purpose of the model was to generate appropriate schedules for the plant while minimizing some objective. The main objectives considered included minimizing purchased water for the cleaning of machines, and the retention of target volume levels in the silos to maintain a steady working operation at the plant. If implemented successfully, the unnamed NZ company could use the spreadsheet model to assist in scheduling machines at the product plant.

The Excel add-in, SolverStudio, was used to process the data from the production plant spreadsheet and visualize the results of the model. And Julia/JuMP was used to formulate and optimize the model. The optimization model solutions were presented as a brief schedule and the results of the solver in a table format.

The model generates feasible solutions for a week using the rolling horizon algorithm. These spreadsheet solution results regulate flow at the plant, establishing a working prototype for the plant that can be developed further to better represent the goals of the plant.

## Acknowledgements

D.P would like to thank Dr. Andrew Mason for his involvement and counsel throughout the year in helping make the project possible. His knowledge and insight were invaluable to the project.

D.P would also like to thank members of the NZ company, who shall remain unnamed, for their contributions throughout the year. This included being available for scheduled meetings, consultations when required, and regular communication.

## Table of Contents

Abstract	2
Acknowledgements	3
1 Introduction	6
2 Literature Review	7
2.1 Machine Scheduling Models for Dairy Plants	7
2.2 Rolling Horizon Algorithm	7
2.3 Constraint Programming in Scheduling Models	8
2.4 JuMP Program	8
2.5 SolverStudio	9
3 Methods	10
3.1 SolverStudio	10
3.2 Process Flow Diagram	10
3.3 Model Definitions	12
3.3.1 Set Definitions	12
3.3.2 Decision Variable Definitions	13
3.3.3 Definition of Constants	14
3.4 Model Formulation	15
3.4.1 Silo Volumes and Truck Constraints	15
3.4.2 Initial Model Machine Constraints	17
3.4.3 Revised Model Machine Constraints	18
3.4.4 Cleaning Constraints	19
3.4.5 Wastewater Constraints	20
3.4.6 Objective Functions	21
3.5 Rolling Horizon Algorithm	22
4 Results	23
4.1 Model Results	23
4.2 Solver Results	25
4.3 History Results	26
4.3.1 Initial History Run	26
4.3.2 Revised History Run	27
5 Discussion	28
5.1 Model Recommendation	28
5.2 Model Limitations	29
5.3 Future Work	29

5 Conclusion	.30
7 Bibliography	.31
Appendix A	.33
Appendix B	.36
List of Figures	
Figure 1 - Rescheduling via a rolling horizon approach [2]	8
Figure 2 - Process Flow Diagram of the plant.	.11
Figure 3 - Process Flow Diagram of the cleaning at the plant	.12
Figure 4 - The Implemented Rolling Horizon Algorithm	.22
Figure 5 - Solution Spreadsheet for objective component 3	.23
Figure 6 - Solution Spreadsheet for objective component 4	.24
List of Tables	
Γable 1 - Condition table for machines 4a and 4b	.16
Γable 2 - Example spreadsheet for runtime constraints	.18
Γable 3 - Condition table for dirty machines	.20
Γable 4 - Results for single objective components 3 and 4	.24
Γable 5 - Solver results using objective components 2 and 3 together	.25
Γable 6 - Solver results using objective components 2, 3, and 4 together	.25
Гable 7 - History results for initial objective	.26
Γable 8 - History results for revised objective	.27

#### 1 Introduction

Production scheduling in the dairy industry has been a prominent theme over the past few decades [1]. Scheduling is the process of allocating resources, such as equipment, utilities, and labour to complete processing tasks to ensure production demands are met. Generally, scheduling decisions at the plants are often made by managers and operators in an ad hoc manner, meaning performance and efficiency heavily rely on their experience and how well they understand the processes and equipment [2].

In this project, an optimization model has been developed for a dairy plant to schedule its production processes efficiently. A spreadsheet of the current model was provided, which was used to evaluate the possible schedules at the plant manually. The implementation of an automated model would help the plant produce optimal, high-quality plant schedules. This will help reduce the involvement of a worker in decision-making. Consequently, this could help a company save on resources such as power and cost.

The optimization model will be developed using the Julia/JuMP optimization framework that interacts with the spreadsheet model to generate and display the solution schedules. The model will be solved mainly using the Gurobi solver. This is because Gurobi is a proven solver, having the fastest solve times for all model types, including Linear Programs and Mixed Integer Programs [3].

Furthermore, the practicalities of real-world plant operations were considered to ensure the solution produced can be applied to the dairy plant immediately if required for a various scenarios. This problem considered complexities such as minimizing the use of purchased water for cleaning purposes and the retention of target volume levels in the silos daily.

#### 2 Literature Review

#### 2.1 Machine Scheduling Models for Dairy Plants

Many different approaches have been taken to model machine schedules for dairy plants. Most models follow a Mixed-Integer-Linear-Programming (MILP) formulation, as they are flexible, capable, and effective [2]. These MILP formulations are often tailored to the details of the dairy plant's requirements to meet their resource, equipment, and demand requirements. This section gives insight into machine scheduling problems in the dairy industry that have been applied to a plant.

One of the main features of a standard dairy plant is that many products are similar, meaning the processing characteristics may be alike. Georgiadis et al. [2] provide good insight into grouping these products with similar properties into product families to simplify the problem by decreasing the number of active variables. The model addresses constraints such as storage limitations, machine capacity, and material balances, which are typically met in production scheduling formulations. The model's objective was to minimize the operating cost, including production, inventory, and changeover costs. The MILP model was solved using the CPLEX solver on the General Algebraic Modelling System (GAMS) interface and applied to a large dairy facility in Greece.

General scheduling methods have been deemed impractical for optimal production scheduling for dairy plants due to the nature and complexity of the systems [1]. To counteract this, Doganis and Sarimveis [1] have implemented a customized MILP model that considers various due dates, sequence-dependent set-up times, and storage management. The key characteristics of this model include that time is partitioned into periods representative of days, meaning that production tasks must start and end on the same day. Within these time periods, time is treated as a continuous variable, meaning it was not required to determine a suitable number of timeslots sequentially as other formulations may implement. This MILP model was applied to a large dairy plant in Greece and solved using the CPLEX solver on the GAMS interface.

Touil et al. [4] implemented a MILP model to address problems in production scheduling to minimize the completion time (makespan) of a multistage, multiproduct milk process. The significant constraints are sequence-dependent changeover times, machine speed, and capacity constraints. The model was applied to a leading dairy company in Morocco and solved using the IBM ILOG CPLEX solver.

#### 2.2 Rolling Horizon Algorithm

A rolling horizon approach is often implemented in production scheduling as a flexibility tool for a time-dependent model where the planning interval is moved forward in time at every iteration of the algorithm [5]. This method, when applied, would allow a plant to plan for the future sequentially.

To account for changes in the production system, such as order modifications, new orders, and cancellations to orders, Georgiadis et al. [2] implemented a rolling horizon algorithm. This reactive scheduling method provides a solution to the volatility of product orders which is quite apparent in the dairy industry, ensuring optimal schedules after new information arrives. In figure 1, it can be seen how a full scheduling horizon was completed through each

iteration of the rolling horizon algorithm. A portion of the prediction horizon would be taken as the control horizon, which was then fixed for the next iteration of the algorithm. This process would continue until the full scheduling horizon was completed.

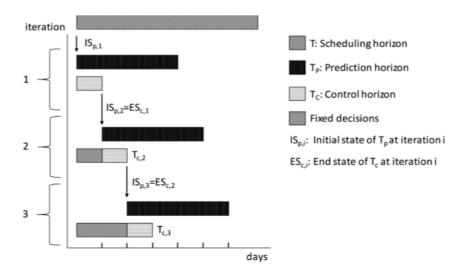


Figure 1 - Rescheduling via a rolling horizon approach [2]

#### 2.3 Constraint Programming in Scheduling Models

As stated, most models for machine scheduling tend to follow a MILP model format to ensure optimal solutions and schedules. However, these methods prioritize optimality and do not always ensure feasible solutions; hence different methods were investigated for this case. A proposed method known to prioritize feasibility over optimality is constraint programming (CP).

Meng et al. [6] implemented a constraint programming model (among many other MILP models) for the distributed flexible job shop scheduling problem (DFJSP) due to its NP-hard classification. The performance and computation times of these models were compared to justify which models are most suitable to be used to solve the DFJSP problem. The CP model considers interval decision variables and domain filtering algorithms. The domain filtering algorithm ensures constraints are feasible by removing the values of the variables for the current domain that aren't in a solution for that constraint. The proposed CP model was concluded to outperform the other algorithms in respect of efficiency and the quality of the solution. The CP model of the DFJSP was coded using IBM CPLEX Studio.

Constraint programming models can be solved using MiniZinc, a high-level program that allows the user to model constraint satisfaction [7]. MiniZinc allows users to model CP and optimization problems in a solver-independent manner while taking advantage of the large library base with pre-defined constraints. The compiled model can then be fed into several solvers, including Gurobi, which it is compatible with.

#### 2.4 JuMP Program

JuMP is a modelling language used for mathematical optimization in Julia [8]. This framework allows users to formulate and solve many problems, such as linear and integer programs, making it relevant and useful for this project. JuMP achieves performance comparable to many commercial modelling tools. This is done using Julia's advanced features

like syntactic macros and code generation [8]. The syntactic macros feature enables JuMP to provide a simple syntax for algebraic modelling without the implications of operator overloading. The code generation feature allows JuMP to generate low-level, yet efficient code as required using the LLVM compiler. Hence, the JuMP framework, in conjunction with Julia, will be used to formulate the scheduling for this project.

Another advantage of using JuMP is the many sources online with relevant column generation code online, which in turn takes on column generation theory. Applying this algorithm to a model could have a significant upside, in terms of saving on computational costs and being more efficient.

Column generation is essentially an algorithm that selects appropriate new columns to add to the base MILP, starting from an initial set of columns (decision variables) that make the problem feasible [9]. Appropriate new columns are chosen through a pricing problem that selects columns based on the criterion of having a significant impact on the objective function value. This method is particularly useful when solving large models, as optimality could be achieved without requiring enumeration of every column possible. Instead, using only the columns added to the MILP through column generation [9].

#### 2.5 SolverStudio

SolverStudio [10] is an Excel add-in that allows users to write and solve optimization models within an Excel workbook using a range of different programming languages including Julia. A unique feature of the program is that it allows data on the spreadsheet to be used in the model as constants, parameters, and variables. The add-in also prints the solution of the model back into the spreadsheet after it has been solved within Excel. SolverStudio [10] calls Julia/JuMP [8] which is compatible with a number of different solvers, including the Gurobi [3], HiGHS [11] and CBC [12] solvers allowing solutions to be found in a timely manner.

#### 3 Methods

#### 3.1 SolverStudio

Before beginning to model the problem, the SolverStudio files were updated to match the current version of Julia/JuMP, which was outdated. This included tasks such as adding catch to the end of try statements in the SSJuliaTools.jl file and fixing encoding errors by adding commands to the SolverStudio.jl file to ensure the results stored in the JuliaResults.py Python file had this appropriate encoding. Another issue that was fixed was applied directly to the model, where printing the Julia results from the model back onto the spreadsheet used outdated code. This was updated to use the correct JuMP commands for accessing the decision variable values.

#### 3.2 Process Flow Diagram

To help visualize the processing plant, the process flow diagram in figure 2 was adopted to show how flow would move through the model and to state the assumptions made. By sequentially looking at the flow diagram, it can be seen where the inputs and outputs of every silo come from. Specifically, the process flow diagram helped to develop and guide the silo volume constraints, which kept silo volume levels within bounds.

Assumptions were vital to keeping the model manageable in complexity. These were decided during consultations with the company to ensure the agreement of the details in the formulation. Meetings were arranged to do this, and regular communication was maintained to ensure they were aware and approved of any assumptions made.

An assumption made was that there was no direct flow between machines 4a and 4b into machine 5, instead, the flow from machines 4a and 4b went through the buffer tank (silo 5) to then get to machine 5. This is because the procedure at the plant was to prioritize the flow from the buffer into machine 5 when machines 4a and 4b were turned on. Hence by ensuring the flow goes into the buffer first from machines 4a and 4b, the flow from the buffer always gets depleted and is the priority. Another assumption of the model was that flow could not go directly from silo 6 into silo 7, instead, the flow from silo 6 would be directed to machine 6 before it could reach silo 7. The current arrangement at the plant allows for this silo-to-silo flow to occur. However, on speaking to members of the team familiar with the plant this direct silo-to-silo flow was looking to get removed from the plant and was hence not included in the model developed.

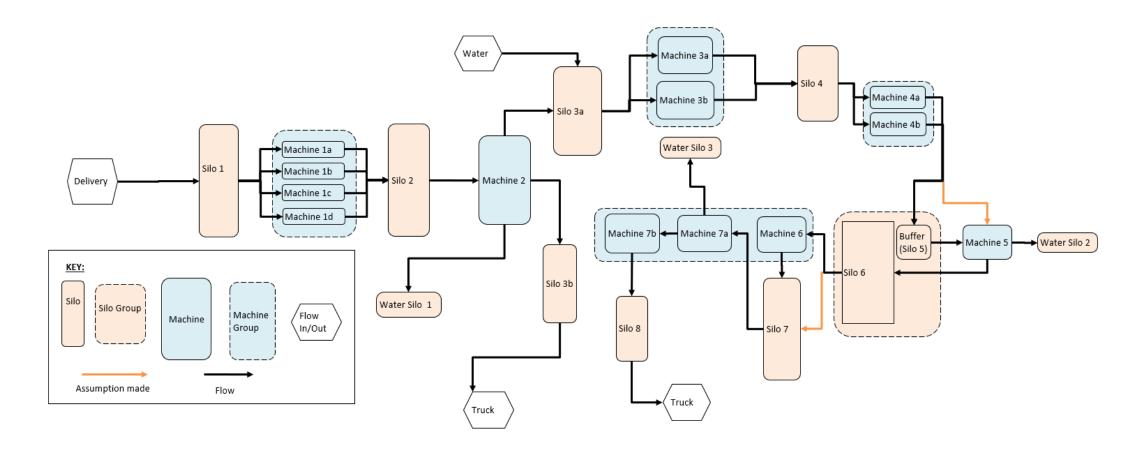


Figure 2 - Process Flow Diagram of the plant.

Note: the orange flows represent the assumptions that have not been included in the model.

Figure 3 shows a process flow diagram for the wastewater constraints. The purple arrows indicate water that flows out of machines 2, 5, and 7a into their respective water silos after being purified. The maroon arrows show the water from the respective water silos used to clean specific machines. The information given at the plant was that the water silos could only be used to clean machines close to this silo. Hence it was assumed that the following machines in the diagram were the only ones that could be cleaned from the recycled water in the water silos.

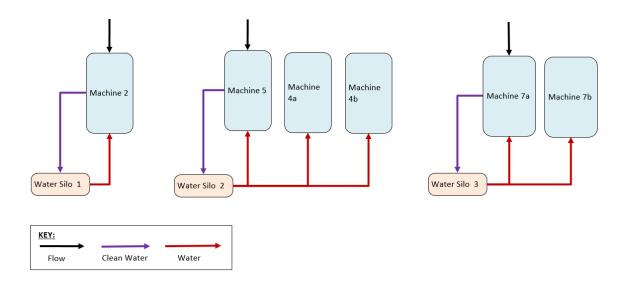


Figure 3 - Process Flow Diagram of the cleaning at the plant.

#### 3.3 Model Definitions

#### 3.3.1 Set Definitions

Time Periods : 
$$T = \{1,2,...,36\}$$
  
Stages/Silos :  $S = \{\{1,2,3a,3b,4,5,6,7,8\}\}$ 

 $P_s$  is the group of predecessor machines at stage s that feed into the silo s,

$$\begin{array}{l} P_2 = \{1a,1b,1c,1d\}, \ P_{3a} = \{2\}, \ P_{3b} = \{2\}, \ P_4 = \{3a,3b\}, \\ P_5 = \{4a,4b\}, \ P_6 = \{5\}, \ P_7 = \{6\}, P_8 = \{7b\} \end{array}$$

 $S_s$  is the group of successor machines at stage s that the silo feeds into the silo s,

$$S_1 = \{1a, 1b, 1c, 1d\}, S_2 = \{2\}, S_{3a} = \{3a, 3b\},$$
  
 $S_4 = \{4a, 4b\}, S_5 = \{5\}, S_6 = \{6\}, S_7 = \{7a\}$   
Truck Type:  $J = \{\{\text{Truck Type 1, Truck Type 2}\}\}$   
 $W = \{\{\text{Water Silo 1, Water Silo 2, Water Silo 3}\}\}$ 

#### 3.3.2 Decision Variable Definitions

 $m_{i,t} = 1$  if machine i is on for the duration of time period t, and 0 otherwise  $m_t^5$  = Continuous flow rate for machine 5 if it is on for the duration of time period t, and 0 otherwise  $m_{i,t}^{\text{start}} = 1$  if machine i starts operating at the start of period t after being off during period t-1, and 0 otherwise  $m_{i,t}^{\text{finish}} = 1$  if machine i finishes operating at the start of period t after being on during period t-1, and 0 otherwise  $p_t = 1$  if there is exactly one machine running between machines 4a and 4b for the duration of time period t, and 0 otherwise  $q_t = 1$  if there are exactly two machines running between machines 4a and 4b for the duration of time period t, and 0 otherwise  $c_{it} = 1$  if machine i is being cleaned for the duration of time period t, and 0 otherwise  $c_{i,t}^{\text{start}} = 1$  if machine i starts cleaning at the start of period t after being cleaned during period t-1, and 0 otherwise  $c_{i,t}^{\mathrm{finish}} = 1$  if machine i finishes cleaning at the start of period t after being cleaned during period t-1, and 0 otherwise  $d_{i,t} = 1$  if machine i dirty at the start of time period t, and 0 otherwise  $n_t^j$  = Number of trucks of type j arriving at the start of time period t  $v_t^s$  = Volume for silo s at the start of time period t  $u_t^w$  = Volume for Waste Water Silo w at the start of time period t  $w_{i,t}^{\rm r}$  = Recycled water used at the start of time period t for machine i  $w_{i,t}^{p}$  = Purchased water used at the start of time period t for machine i

 $m_{i,t}^{\mathrm{switch}} = 1$  if machine i starts operating at the start of period t after being off during period t-1, -1 if machine i ... finishes operating at the start of period t after being on during period t-1 and 0 otherwise

#### 3.3.3 Definition of Constants

 $F_i$  = Flow rate into the silo from the predecessor machine i

 $O_i$  = Flow rate out of the silo into the successor machine i

 $f_1$  = constant flow rate for exactly one machine running between machines 4a and 4b

 $f_2$  = constant flow rate for exactly two machines running between machines 4a and 4b

 $D_t$  = Milk delivery to the plant at the start of time period t

 $V^s = \text{Max silo volume for silo s at the start of time period t}$ 

 $V^{j}$  = The truck volume for truck j

 $r_i^{\text{max}} = \text{Maximum runtime for machine i}$ 

 $r_i^{\min}$  = Minimum runtime for machine i

 $c_i^{\text{length}} = \text{Cleaning runtime length for machine i}$ 

 $M_i$  = Water used in the cleaning for machine i

 $g^s$  = Target volume goal for silo s

#### 3.4 Model Formulation

#### 3.4.1 Silo Volumes and Truck Constraints

The first set of constraints that were developed were the silo volume constraints. These constraints linked the set of machines according to the full process flow diagram to ensure a mass balance throughout. Generally, the silo volume constraints followed the generic mass balance equation below that was adapted for specific silos.

Inflow - Outflow + Carryover Volume = Volume

$$D_{t} - \sum_{i \in S_{1}} (O_{i} \cdot m_{i,t}) + v_{t-1}^{1} = v_{t}^{1} \quad \forall \ t \in T$$

(1)

$$\sum_{i \in S_1} m_{i,t}^{\text{start}} \le 2 \quad \forall t \in T$$
(2)

$$0 \le v_t^s \le V^s \quad \forall \ t \in T, s \in S$$

(3)

The first step of the production plan was to model the delivery of flow into the first silo, also known as stage 1 of the plant. Constraint 1 has an inflow volume  $D_t$  delivered at the start of time period t and subtracts the sum of the outflow flow rate  $O_i$  for every machine  $i \in S_1$  that follows the silo if machine i is turned on for the duration of time period t. The volume at the start of the previous time period  $v_{t-1}$  was then added to get the volume of the silo at the start of the current time period  $v_t$ .

Constraint 2 was created to ensure that the sum of machines  $i \in S_I$  that follows silo 1 if machine i starts operating at the start of time period t after being off for time period t-1 is less than equal to two. Therefore, only two of the machines in this group can start at any one point in time. Constraint 3 ensures that the volume of each silo  $v_t$  for the duration of time period  $t \in T$  for all silos  $s \in S$  stays between 0 and the max silo volume of silo s, V. This constraint was applied for all the stages of the production constraints.

$$\sum_{i \in P_S} \left( F_i \cdot m_{i,t} \right) - \sum_{i \in S_S} \left( O_i \cdot m_{i,t} \right) + v_{t-1}^S = v_t^S \quad \forall t \in T, s \in S$$

Constraint 4 was the general silo volume constraint that was applied to silos 2, 3a, and 7 and adapted appropriately for the rest of the silos. The only difference to the above delivery formulation is that instead of a delivery inflow volume into the silo at the start of time period t, the inflow volume was the sum of the inflow flow rate  $F_i$  for every machine  $i \in P_s$  that precedes the silo if machine i is turned on for the duration of time period t. Note: slack and surplus variables were added to each of the general silo volume constraints for feasibility purposes if a solution for the model couldn't be found initially.

$$\sum_{i \in P_4} (F_i \cdot m_{i,t}) - (f_1 \cdot p_t + (f_2 - f_1) \cdot q_t) + v_{t-1}^4 = v_t^4 \quad \forall t \in T$$

$$(f_1 \cdot p_t + (f_2 - f_1) \cdot q_t) - m_t^5 + v_{t-1}^5 = v_t^5 \qquad \forall t \in T$$

$$\sum_{i \in S_4} m_{i,t} = p_t + q_t \qquad \forall t \in T$$
(7)

$$p_t \ge q_t \qquad \forall t \in T$$

	Ex. 1	Ex. 2	Ex. 3
$p_t$	0	1	1
$q_t$	0	0	1
Machine(s) Running	0	1	2
Flow	0	$f_1$	$f_2$

Table 1 - Condition table for machines 4a and 4b

Constraints 5-8 were created for a portion of the plant where the machines carrying flow between silos 4 and 5, machines 4a and 4b, had different conditions to the other machines in the plant. The conditions in table 1 show the effect that the number of machines running had on the outflow of silo 4 and inflow of silo 5.

The outflow for constraint 6 works slightly different to the general silo constraint. Instead of summing over constant flow rates multiplied by the machine's running variable  $m_{it}$ , the outflow of constraint 6 is the continuous flow rate  $m_t^5$  where 0 is the lower bound and also an indicator that the machine is off. This variable is slightly different to the other machine running variables  $m_{it}$ .

$$m_t^5 - \sum_{i \in S_6} (O_i \cdot m_{i,t}) + v_{t-1}^6 = v_t^6 \quad \forall t \in T$$

Constraint 9 was another implementation of the general silo volume constraint, where the outflow follows the same formulation only for the specific set of machines following silo 6. Much like the outflow of constraint 6, the inflow of constraint 9 is the continuous flow rate of machine 5  $m_t^5$ .

$$\sum_{i \in P_S} (F_i \cdot m_{i,t}) - (V^j \cdot n_t^j) + v_{t-1}^s = v_t^s \quad \forall t \in T, s \in S, j \in J$$

$$\sum_{j \in J} n_t^j \le 2 \quad \forall t \in T$$
(11)

Constraint 10 was the final implementation of the general silo volume constraint, which was applied to silos 3b and 8. This case was slightly different as instead of an outflow going into another silo it would be filling up a truck. The outflow was specified as truck j's volume  $V^j$  multiplied by the number of trucks  $n_i^j$  of type j that arrived at the start of time period t.

This was done to ensure that trucks were filled up at a flow rate equivalent to its volume so that the truck, at the end of time period t, be ready to leave the plant thereafter. Constraint 11 ensured that the number of trucks  $n_t^j$  of type  $j \in J$  that had arrived at the plant at the start of time period t was no larger than 2 meaning that no more than two trucks could arrive to the plant at any one point in time.

#### 3.4.2 Initial Model Machine Constraints

Initially a set of machine constraints were developed to ensure correct runtimes were set in the model for machines which were later revised with simpler formulations for the model. The following were the initial runtime constraints used in the model to generate solutions.

$$m_{i,t}^{\mathrm{switch}} = m_{i,t} - m_{i,t-1} \quad \forall t \in T, i \in S_s$$

Constraint 12 was created as a switch constraint, which was calculated by finding the difference between the machine running variable  $m_{i,t}$  for machine  $i \in S_s$  at the start of time period t and t-1 for all  $t \in T$ . The machine switch variable  $m_{i,t}$  switch would have a 1 where the machine started operating at the start of time period t and a -1 where the machine stops operating at the start of period t after being on during time period t-1 for  $i \in S_s$ .

$$\sum_{j=t}^{j=t+r_{i}^{\min}-1} m_{i,j} \geq r_{i}^{\min} \cdot m_{i,t}^{\text{switch}} \quad \forall \ t \in T, \ i \in S_{s}$$

$$\sum_{j=t+r_{i}^{\max}} \sum_{j=t}^{j=t+r_{i}^{\max}} m_{i,j} \leq r_{i}^{\max} \quad \forall \ t \in T, \ i \in S_{s}$$

$$(13)$$

Constraints 13 and 14 were the model's minimum and maximum runtime constraints. These constraints ensured that the runtime for machine  $i \in S_s$  was greater than the minimum machine runtime  $r_i^{\min}$ , and no greater than the maximum runtime  $r_i^{\max}$ . Setting minimum and maximum runtime constraints were important to the model to ensure realistic solutions were generated. By having no minimum or maximum runtime constraints, the machines wouldn't be required to run nor stop running which is highly unrealistic at a production plant.

$$\sum_{j=t}^{j=t+r_i^{\min}-1} m_{i,j} \le c_i^{\text{length}} \cdot \left(1 + m_{i,t}^{\text{switch}}\right) \quad \forall \ t \in T, \ i \in S_s$$
(15)

Constraint 15 ensured that that there was the suitable amount of time between a machine  $i \in S_s$  being turned off and on again for cleaning to occur.

#### 3.4.3 Revised Model Machine Constraints

After reviewing the initial machine constraints, it was decided that they could be made tighter and improved using a different method with separate start and stop decision variables. The revised model used the following revised machine runtime constraints.

$$m_{i,t} \leq m_t^5 \quad \forall t \in T, i \in S_5$$
 (16) 
$$f_2 \cdot m_{i,t} \geq m_t^5 \quad \forall t \in T, i \in S_5$$
 (17)

Constraints 16 and 17 were conversion constraints of the continuous flow rate runtime variable  $m_t^5$  into the binary machine runtime variable  $m_{i,t}$  for machine 5. This was done to ensure that the machine runtime constraints could formulate the runtime constraints.

$$m_{i,t-1} + m_{i,t}^{\text{start}} - m_{i,t}^{\text{finish}} = m_{i,t} \quad \forall \ t \in T, \ i \in S_s$$

$$m_{i,t}^{\mathrm{start}} \leq 1 - m_{i,t-1} \quad \forall \ t \in T, \ i \in S_s$$

$$m_{i,t}^{\mathrm{finish}} \leq m_{i,t-1} \quad \forall \ t \in T, \ i \in S_s$$

(20)

T	1	2	3	4	5	6
$m_{i,t}$	0	1	1	1	1	0
$m_{i,t}^{ m start}$	0	1	0	0	0	0
$m_{i,t}^{ m finish}$	0	0	0	0	0	1

Table 2 - Example spreadsheet for runtime constraints

The combination of constraints 18 to 20 ensure that the machine start time  $m_{i,t}$  start for  $i \in S_s$ can only be 1 if machine i starts operating at the start of period t after being off during time period t-1. They also ensure that the machine finish time  $m_{i,t}$  finish for  $i \in S_s$  can only be 1 if machine i finishes operating at the start of period t after being on during time period t-1. The following can be seen in table 2.

$$\sum_{j=t+r_{i}^{\min}+1}^{j=t} m_{i,j}^{start} \leq m_{i,t} \quad \forall t \in T, i \in S_{s}$$

$$j=t+r_{i}^{\max}$$

$$\sum_{j=t}^{j=t+r_{i}^{\max}} m_{i,j} \leq r_{i}^{\max} \quad \forall t \in T, i \in S_{s}$$

$$(21)$$

Constraints 21 and 22 were the model's minimum and maximum runtime constraints. These constraints ensured that the runtime for machine  $i \in S_s$  was greater than the minimum machine runtime  $r_i^{\min}$ , and no greater than the maximum runtime  $r_i^{\max}$ .

$$m_{i,t} = m_{j,t-k} \quad \forall \ t \in T, \ i,j \in S_s, \ i \neq j, \ k \in (1,2)$$

Constraint 23 was created for the specific case where machines had to be turned on 1-2 hours after each other. This constraint was applied to machines 6, 7a and 7b, to ensure that machine 7a started one hour after machine 6, and machine 7b started one hour after machine 7a.

$$m_{i,t} + c_{i,t} \le 1 \quad \forall t \in T, i \in S_s$$

Constraint 24 ensures that cleaning could not be scheduled when the machine was running by having the sum of the machine running variable and cleaning running variable for machine  $i \in S_s$  less than equal to 1.

#### 3.4.4 Cleaning Constraints

Cleaning was introduced in the model and incorporated similar constraints to the revised machine start and stop variables to ensure consistency.

$$c_{i,t-1} + c_{i,t}^{\text{start}} - c_{i,t}^{\text{finish}} = c_{i,t} \quad \forall \ t \in T, \ i \in S_s$$
(25)

$$c_{i,t}^{\mathrm{start}} \leq 1 - c_{i,t-1} \quad \forall \ t \in T, \ i \in S_s$$

$$c_{i,t}^{\mathrm{finish}} \leq c_{i,t-1} \quad \forall \, t \in T, \, i \in S_s$$

(29)

$$\sum_{j=t+c_i^{length}+1}^t c_{i,j}^{start} \le c_{i,t} \quad \forall \ t \in T, \ i \in S_s$$

$$\sum_{j=t}^{j=t+c_i^{length}} c_{i,j} \leq c_i^{length} \quad \forall \ t \in T, \ i \in S_s$$

The cleaning runtime constraints, constraints 25-29, are identical to the revised machine runtime constraints, constraints 18-22. This was done to ensure that the cleaning start time  $c_{i,t}^{\text{start}}$  and cleaning finish time  $c_{i,t}^{\text{finish}}$  work like the machine parallels respectively. Constraints 28 and 29 were also added to ensure cleaning runtime does not exceed the required cleaning time  $c_i^{\text{length}}$ .

$$d_{i,t-1} - c_{i,t}^{\rm start} + m_{i,t}^{\rm finish} \le d_{i,t} \quad \forall \ t \in T, \ i \in S_s$$

$$d_{i,t} + m_{i,t}^{\mathrm{start}} \leq 1 \quad \forall \ t \in T, \ i \in S_s$$

T	1	2	3	4	5	6
$d_{i,t}$	0	1	1	1	0	0
$c_{i,t}^{\mathrm{start}}$	0	0	0	0	1	0
$m_{i,t}^{ m finish}$	0	1	0	0	0	0

Table 3 - Condition table for dirty machines

Constraint 30 ensures that when machine  $i \in S_s$  finishes running at the start of time period t, it is considered dirty at the start of time period t and only becomes clean (not dirty) when the cleaning procedure begins at the start of a different time period t. This can be seen in table 3.

#### 3.4.5 Wastewater Constraints

To quantify the amount of water being used in the cleaning, and where it was coming from, wastewater constraints were defined.

$$w_{i,t}^{\mathrm{r}} + w_{i,t}^{\mathrm{p}} = M_i \cdot c_{i,t} \quad \forall \ t \in T, \ i \in S_s$$
(32)

Constraint 33 ensures the amount of recycled water  $w_{i,t}^r$  and purchased water  $w_{i,t}^p$  used at the start of time period t for machine i was equivalent to the water used in the cleaning process  $M_i$  for for  $i \in S_s$  when the cleaning process was underway.

$$\sum_{i \in 2} (F_i - O_i) \cdot m_{i,t} - \sum_{i \in 2} w_{i,t}^{r} + u_{t-1}^{w} \ge u_t^{w} \quad \forall t \in T, w \in W$$
(33)

$$\sum_{i \in 5} (1 - 1/0_i) \cdot m_{i,t}^5 - \sum_{i \in (4a,4b,5)} w_{i,t}^r + u_{t-1}^w \ge u_t^w \quad \forall t \in T, w \in W$$

$$\sum_{i \in 7a} (F_i - O_i) \cdot m_{i,t} - \sum_{i \in (7a,7b)} w_{i,t}^r + u_{t-1}^w \ge u_t^w \quad \forall t \in T, w \in W$$
(35)

The water used for cleaning in the plant came from 3 different silos, silos 2, 5 and 7a, all in separate water storage tanks. This water was then used to clean silos close to this storage. In the formulation, the water that came from silo 2, was used to clean silo 2 (constraint 34), the water that came from silo 5, was used to clean silos 4a, 4b and 5 (constraint 35), and the water that came from silo 7a, was used to clean silos 7a and 7b (constraint 36). Cleaning was scheduled for the other machines, but the recycled water usage was not considered.

Constraints 34 to 36 followed the generic mass balance equation for the 3 different water silos. The inflow was formulated as the remainder of flow from the silo that didn't proceed to be processed as the key milk flow. This was written slightly differently in constraint 35 due to the machine decision variable for machine 5  $m_t^5$  already being a flow. The outflow for all 3 constraints was the recycled water  $w_{i,t}^T$  returning back into the machine for cleaning for machine i stated in the respective constraints. The volume of the water silo at the start of the previous time period  $u_{t-1}$  was then added to get the volume of the water silo at the start of the current time period  $u_t$ .

#### 3.4.6 Objective Functions

A wide variety of different objective function components were considered in the problem, all with different goals of the plant taken into consideration.

$$\operatorname{Min} \sum_{t \in T} \sum_{s \in S} (\operatorname{suplus}_t^s - \operatorname{slack}_t^s)$$

(36)

Objective Component 1 was written to minimize the difference of surplus and slack variables when added to the silo volume constraints. This component only came into action when a solution couldn't be found initially.

$$\min \sum_{t \in T} v_t^5$$

(37)

Objective component 2 minimized the flow in silo 5, as it was found to provide a faster turnover in the silo, which had a slightly different formulation to the rest of the silos.

$$\min \sum_{s \in S} |g^s - v_t^s|$$

(38)

Objective component 3 minimized the difference between a target volume goal for silo s  $g^s$  and the volume of silo s at the start of time period t. This objective aimed to retain a set amount of volume in each silo s at the start of time period t. This objective component helped regulate flow in the model and maintain a steady working operation at the plant.

$$\min \sum_{t \in T} \sum_{i \in S_S} w_{i,t}^{p}$$

(39)

Objective component 4 minimized the amount of purchased water  $w_{i,t}^p$  that was used to clean the silos. Ensuring that the priority of water used to clean the silo was the recycled water from the water silos when available and purchased water only after this if required.

#### 3.5 Rolling Horizon Algorithm

An adaptation of the rolling horizon algorithm was created to ensure a plan could be made over a couple day period instead of only 36 hours. The algorithm worked by locking the machine decision variables in the last 12-hour time period into the first 12 hours of the next model, then re-solving for the rest of the 24-hour period. This was adapted to give more insight to the site on how a week's schedule may look for the plant.

The VBA Macro tool was used to record two procedures to ensure the correct data was used going forward. The first being that the carryover volumes for the silos were adjusted to be the volume of the silo in the 24th hour. The second macro recorded the decision variables being moved from the last 12-hour time period into the first 12-hour time period before the rest of the decision cells were cleared. Also recorded in the second macros were the milk delivery volumes from the first 12 hours being swapped with the last 12-hour time period to ensure the following day had the correct milk delivery set. These two macros were added to the spreadsheet as buttons for ease of access where the user would first click the reset carryover volumes button and then click the reset decision variables button. Julia code was written to lock these decision cells in the model by checking whether they weren't empty and locking them by setting a hard constraint to the value in the spreadsheet if so.

Figure 4 shows how the adapted rolling horizon algorithm worked to solve a set of solutions for a scheduling horizon. The first iteration solves for a model, which is the first prediction horizon. Every iteration from here, until the scheduling horizon has been completed, takes the solution in the control horizon (the last 12-hour period of the prediction horizon) and locks it in the next solutions prediction horizon as the first 12-hour period. In figure 4 it can be seen how the rolling horizon algorithm solves a set of solutions for a scheduling horizon of 7 days.

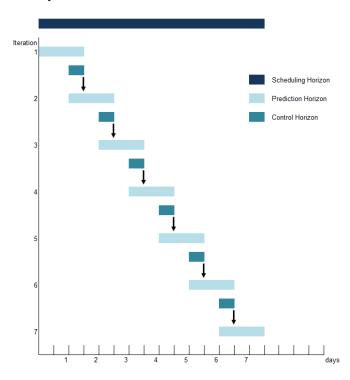


Figure 4 - The Implemented Rolling Horizon Algorithm

#### 4 Results

The model was solved using different objectives, for the same starting silo volumes, and milk delivery timeframe. A template results spreadsheet was created with all the relevant features of the model to show the recommended machine decision variables over the 36-hour time period and the varying silo volume over this time. A variety of different solvers, including Gurobi, HiGHs, and CDC, were also compared to verify which solver was the best for this problem. History was also considered to generate a week's worth of scheduling for the model so that workers may see how schedules may look for the plant and the fluctuating silo levels over time.

#### 4.1 Model Results

When finding solutions for the model, a consistent set of termination rules were set to ensure decent solutions, or if this was not possible, then a solution within a set timeframe was found, if any. Runs were terminated when a MIP gap of less than 5% was achieved, between the optimal LP and incumbent solutions, as this is indicative of a decent solution and would allow solve times to decrease. If this was not possible in a timely manner, a run was also terminated after 100 seconds. The first objective tested used the 3<sup>rd</sup> objective component.

The initial silo volume for silo s was set as the target volume goal  $g^s$ , and the start of the  $24^{th}$  time period was the volume optimized for silo s.

Min 
$$\sum_{s \in S} |v_0^s - v_{24}^s|$$

	-	_	-	-	_	_	-	_	_			40	40		-45	-40	477	40	- 40		- 0.4			0.4							- 0.4					
Time	_1	2	3	- 4	5	- 6		8	9	10	11	12	13	14	15	16	1/	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Milk delivery		153	153	153	153	153	153	153	153		153	0	0	0			214	178	156	99	232	40	0	0	155	155	155	155	155	155	155	155	155	0	0	. 0
	216	214	153	91	65	98	191	285	378	472 !	530 :	375	220	65	38	85	84	107	203	242	414	334	274	214	309	368	368	367			546	605	700			640
Machine 1a	_ 1	- 1	- 1	- 1	- 1	C	C	C	- 1	- 1	- 1	- 1	- 1	_1	- 1	- 1	c	c .	С	U	- 1	1	1	_ ]	_ 1	- 1	11	B (	C (	3	U	U	U	0	0	U
Machine 1b	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	С	0 0		0	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	0	0	미	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1 c	c		2
Machine 1c	- 1	- 1	- 1	С	C	C	0	0	0	- 1	- 1	1	- 1	- 1	- 1	- 1	1	С	С	С	0	0	0	0	1	- 1	- 1	- 1	- 1	- 1	- 1	0 0	0		0	0
Machine 1d	- 1	- 1	- 1	- 1	C	c	С	0	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	С	С	С	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Silo 2 Volume	347	394	442	458	421	474	528	582	636	576	570 5	563	557	551	599	646	640	549	458	511	619	528	436	345	285	279	273	213	153	93	178	232	286	286	286	286
Machine 2	- 1	- 1	- 1	- 1	- 1	c	c	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	0 (	8	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	0 (	c	0	0	0	0
Water	- 1	- 1	- 1	- 1	- 1	0	0	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	0	0	0	0	0	0
Water Silo 1	9	19	28	37	47	47	47	47	47	56	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	168	183	244	304	365	332	254	176	98	113	129	144	160	175	191	207	222	238	253	175	97	113	128	144	159	208	257	305	354	402	357	312	267	222	177	177
Machine 3a	- 1	- 1	c	С	c	c	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1.0	5
Machine 3b	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1	С	C I	D (	С	0	0	0	0	0	0	0
Silo 3b Volume	45	58	70	30	42	16	16	16	16	29	15	27	13	26	38	50	10	23	35	35	35	21	33	46	58	70	82	95	107	119	119	119	119	119	119	119
Truck 1	0	0	0	2	0	- 1	0	0	0	0	- 1	0	- 1	0	0	0	2	0	0	0	0	- 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	429	452	409	366	323	280	222	164	105	47	71	94	118	142	83	107	131	154	178	202	225	249	273	394	418	387	355	324	293	261	230	198	167	234	300	300
Machine 4a	0	0	0	0	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	С	С	С	С	0	0	0	0	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1 c		0 (	5
Machine 4b	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1 0	0 0		0 (	С	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1	С	С	С	c	0	0	0	0	0	0	0	0	0
Silo 5 volume	0	0	0	0	0	98	278	278	278	278	278	196	114	32	0	0	0	98	196	294	294	212	130	48	0	0	0	0	0	0	98	196	294	294	294	294
Machine 5	- 1	- 1	- 1	- 1	- 1	c	С	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	С	c	0	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	- 1	1	0 (	С	0	0	0	0
Water Silo 2	9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	166	210	205	178	152	81	10	21	31	113	195 2	227	238	226	237	210	184	113	113	113	144	155	166	117	90	64	38	82	127	171	171	171	171	171	171	171
Machine 6		С	f	- 1	- 1	- 1	- 1	- 1	- 1	0 0	: f		- 1	- 1	- 1	- 1	- 1	- 1	С	e I		- 1	- 1	- 1	- 1	- 1	1 1	в (	С	0	0	0	0	0	0	0
	-65	0	0	61	57	53	49	45	41	-24	0	0	61	57	53	49	45	41	-24	0	0	61	57	53	49	45	41	-24	0	0	0	0	0	0	0	0
Machine 7a	- 1	C	С	0	- 1	- 1	- 1	- 1	- 1	1 0	0 0		0	- 1	- 1	- 1	- 1	- 1	- 1	0 (	В	0	- 1	- 1	- 1	- 1	- 1	1	c (	3	0	0	0	0	0	0
machine 7b	- 1	- 1	С	С	0	- 1	- 1	- 1	- 1	- 1	1 c		0	0	- 1	- 1	1	- 1	- 1	- 1	В	С	0	- 1	- 1	- 1	- 1	- 1	1 (	0 0	С	0	0	0	0	0
Water Silo 3	41	41	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Silo 8 Volume	25	12	12	12	12	23	10	21	32	19	30	30	5	5	16	28	39	50	61	72	23	23	23	10	21	32	43	6	17	17	17	17	17	17	17	17
Truck 2	0	- 1	0	0	0	0	- 1	0	0	- 1	0	0	- 1	0	0	0	0	0	0	0	2	0	0	- 1	0	0	0	2	0	0	0	0	0	0	0	0

Figure 5 - Solution Spreadsheet for objective component 3

$$\min \sum_{t \in T} \sum_{i \in S_S} w_{i,t}^{p}$$

Time	1	2	3	4	5	- 6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Milk delivery	0	153	153	153	153	153	153	153	153	153	153	0	0	0	128	262	214	178	156	99	232	40	0	0	155	155	155	155	155	155	155	155	155	0	0	0
Silo 1 Volume	216	369	463	556	590	623	596	570	543	577	610	490	370	250	318	520	614	672	708	687	859	839	779	779	934	968	943	917	892	866	841	815	850	850	850	850
Machine 1a	0	0	0	0	0	1	1	1	1	- 1	1	1	1	C	C	C	0	0	0	0	0	0	0	0	1	1	1	- 1	1	1	1	c	C	С	0	0
Machine 1b	0	1	1	1	1	1	- 1	1	C	C	C	1	1	1	1	1	- 1	1	1	C	C	С	0	0	1	1	1	1	1	- 1	1	1	c	0	С	1
Machine 1c	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Machine 1d	0	0	0	1	1	1	- 1	1	1	1	1	C	c	C	0	1	1	1	1	1	1	1	c	С	C	1	1	- 1	1	- 1	1	1	c .	c i	С	1
Silo 2 Volume	353	407	460	423	385	402	418	434	397	359	322	284	247	155	209	317	424	531	639	548	456	365	220	75	38	54	70	86	103	119	280	388	388	243	98	60
Machine 2	0	0	0	1	1	1	1	1	1	1	1	1	1	1	C	C	0	0	0	1	1	1	1	1	1	1	1	1	1	1	C	C	0	- 1	1	- 1
Water	0	0	0	- 1	1	1	1	1	1	- 1	1	1	1	1	0	0	0	0	0	1	- 1	1	- 1	1	1	1	1	- 1	1	1	0	0	0	1	- 1	- 1
Water Silo 1	0	0	0	9	19	28	37	47	56	65	75	84	93	100	50	0	0	0	0	9	16	25	35	44	53	63	72	81	91	100	50	0	0	0	0	0
Silo 3a Volume	107	62	17	66	114	163	211	260	275	291	306	322	337	353	275	230	185	140	95	111	126	142	157	173	188	204	252	301	350	398	353	308	263	312	405	466
Machine 3a	1	- 1	1	1	1	1	- 1	1	1	1	1	1	1	1	1	1	- 1	- 1	- 1	1	1	1	1	- 1	1	1	- 1	1	_ 1	1	- 1	1	1	11	С	c
Machine 3b	С	С	C	C	0	0	0	0	1	1	- 1	- 1	- 1	1	1	C	C	C	C	1	1	1	1	1	1	1	С	С	С	С	0	0	0	0	0	1
Silo 3b Volume	33	33	33	45	6	18	30	42	55	67	79	65	26	38	38	38	38	12	12	24	36	49	61	73	85	98	110	70	56	69	17	17	17	3	15	27
Truck 1	0	0	0	0	2	0	0	0	0	0	0	- 1	2	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2	1	0	2	0	0	1	0	0
Silo 4 Volume	374	342	311	280	346	413	480	448	472	496	519	543	567	590	614	582	649	716	782	904	1026	1147	1269	1293	1316	1340	1309	1277	1246	1215	1183	1152	1219	1285	1285	1340
Machine 4a	1	1	1	1	C	C	C	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Machine 4b	0	0	0	0	0	0	0	1	1	1	- 1	1	1	1	1	1	C	C	C	C	0	0	0	- 1	1	1	1	1	1	1	1	1	c	C	С	c
Silo 5 volume	98	196	196	114	32	0	0	0	0	0	0	0	98	196	196	114	32	0	0	0	0	0	0	0	98	196	196	114	32	0	0	0	0	0	0	0
Machine 5	0	0	- 1	1	- 1	1	- 1	- 1	1	1	1	1	C	С	1	1	- 1	1	- 1	1	- 1	- 1	- 1	1	c	С	- 1	1	1	1	- 1	1	1	1	- 1	- 1
Water Silo 2	0	0	1	99	100	100	51	0	53	99	99	100	50	0	49	49	0	26	0	0	1	1	47	100	50	0	30	42	98	99	99	100	80	30	0	0
Silo 6 Volume	121	121	203	285	299	299	299	294	267	241	214	188	117	46	128	210	224	224	224	224	224	224	224	269	269	269	350	382	370	344	317	291	220	149	149	149
Machine 6	0	0	0	0	0	0	0	f	1	1	1	1	- 1	1	C	C	0	0	0	0	0	0	0	0	0	0	0	f	1	1	1	1	1	1	С	С
Silo 7 Volume	0	0	0	0	0	0	0	0	61	57	53	49	45	41	-24	0	0	0	0	0	0	0	0	0	0	0	0	0	61	57	53	49	45	41	-24	0
Machine 7a	0	0	0	0	0	0	0	0	0	- 1	- 1	1	1	1	1	C	C	0	0	0	0	0	0	0	0	0	0	0	0	- 1	1	1	1	1	- 1	C
machine 7b	_1	0	0	0	0	0	0	0	0	0	1	_ 1	. 1	1	1	1	C	C	0	0	0	0	0	0	0	0	0	0	0	0	- 1	- 1	1	- 1	1	1
Water Silo 3	0	0	0	0	0	0	0	0	0	41	82	100	100	100	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	41	59	59	59	9	50	0
Silo 8 Volume	25	- 1	1	1	1	1	1	1	1	1	12	23	34	45	56	68	68	68	68	68	68	68	68	68	68	19	19	19	19	19	30	16	28	39	50	61
Truck 2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	1	0	0	0	0

Note: The purchased water in results table 4, is only the amount of purchased water that was used for machines that could've used recycled water — which consists of machines 2, 4a, 4b, 7a and 7b.

Figure 6 - Solution Spreadsheet for objective component 4

Objective					Mo	odel Output				
Used	Solver	Incumbent Solution Objective	Explored Nodes	Objective Function Value	Optimality Gap (%)	Solve Time (seconds)	Feasible without slack & surplus variables	Recycled water used (m <sup>3</sup> )	Purchased water used (m <sup>3</sup> )	$\frac{\sum_{s \in S}  v_0^s - v_{24}^s }{(m^3)}$
Objective component 3 – Target Volume	Gurobi	0	91,357	61.54	50.75	100	Yes	0	2,150	61.54
Objective component 4 - Wastewater	Gurobi	0	1,396	100	0	6.93	Yes	1,150	100	1,993

Table 4 - Results for single objective components 3 and 4

Table 4 shows that the model prioritizing objective component 4 had a better solve time at 6.93 seconds whereas the model prioritizing objective component 3 solved in 100 seconds. This correlates to the number of explored nodes, as the objective component 4 model only explored 1,396 nodes to the 91,357 of the objective component 3 model. The optimality gap of the model prioritizing objective component 3 was relatively high, with an optimality gap of 50.75%, whereas it was 0% for the model prioritizing objective component 4 – meaning no difference between the incumbent and optimal LP solution, and that an optimal solution was found. The incumbent solution objective was 0 for both models signalling that the model formulations were somewhat poor.

There was a noticeable difference in the amount of recycled water used for each model, with the objective component 4 model recycling 1,150 m³ of water for cleaning. In contrast, the objective 3 component model didn't consider using recycled water. The objective component 4 model only had to purchase 100 m³ of water to complete the cleaning, in comparison to the objective component 3 model that purchased 2,150 m³ of water to do this. It seems that the model prioritizing objective component 4 wasn't regulating flow as much throughout the model (seen in the rising silo volumes) to ensure less cleaning with purchased water was done. The model prioritizing objective component 3 was better at regulating flow for the first 24 hours but seemed to also have some troubles after this.

#### 4.2 Solver Results

Gurobi, HiGHs, and Cbc, were tested on the model to see what solver would give the best results in a 100 second timeframe. These solvers were experimented on two objective functions to give a wide perspective of solutions for different cases.

Min 
$$\sum_{t \in T} v_t^5 + \sum_{s \in S} |v_0^s - v_{24}^s|$$

			M	Iodel Output		
Solver	Incumbent	Explored	Objective	Optimality	Solve	Feasible without
	Solution	Nodes	Function	Gap (%)	Time	slack & surplus
	Objective		Value		(seconds)	variables
Gurobi	0	86,215	60.38	71.49	100.04	Yes
HiGHs	0	14,021	316.48	99.37	99.9	Yes
Cbc	0	7,305	263.83	100	99.02	Yes

Table 5 - Solver results using objective components 2 and 3 together

$$\operatorname{Min} \ \sum_{t \in T} v_t^5 \ + \ \sum_{s \in S} |v_0^s - v_{24}^s| \ + \sum_{t \in T} \sum_{i \in S_s} w_{i,t}^p$$

			M	lodel Output		
Solver	Incumbent	Explored	Objective	Optimality	Solve	Feasible without
	Solution	Nodes	Function	Gap (%)	Time	slack & surplus
	Objective		Value		(seconds)	variables
Gurobi	0	5,299	273.17	34.89	100.04	Yes
HiGHs	0	995	592.68	99.66	100	Yes
Cbc	0	4,128	/	/	98.87	No

Table 6 - Solver results using objective components 2, 3, and 4 together

In the first experiment in table 5, all the solvers found a solution using the model. Gurobi found the best objective function value in comparison to HiGHs and Cbc, which was helped by exploring 6 and 12 times more nodes than HiGHS and Cbc, respectively, in the 100 second timeframe. The optimality gap was also lower and therefore better for Gurobi compared to the other solvers. The second experiment in table 6 also paints a similar picture with Gurobi exploring the most nodes and finding the best objective function value. Although it was able to explore more nodes than the HiGHS solver, the Cbc solver could not find a solution in the second experiment. In neither experiment did any solver find an incumbent solution objective, meaning the model formulations were somewhat poor for the models.

#### 4.3 History Results

#### 4.3.1 Initial History Run

The addition of history into the model allows the worker the ability to generate a weeks' worth of scheduling prior using the adapted rolling horizon algorithm. In the following example, objective components 1, 2 and 3 were minimized and 7 solutions generated over a 7-day period. The fourth objective component was not included in this objective for simplicity purposes, as well as it being seen in figure 6 that it wasn't regulating flow well throughout the plant which would cause problems through a 7 day history run where flow regulation was vital to good solutions being created.

Note: Objective component 3's target volume goal  $g^s$  for silo s was set as the initial silo volume  $v_0^s$  for silo s, and this did not change with every iteration of the model as it was set as a constant before the first run.

$$\operatorname{Min} \ \sum_{t \in T} \sum_{s \in S} \left( \operatorname{suplus}_t^s - \operatorname{slack}_t^s \right) \ + \operatorname{Min} \ \sum_{t \in T} v_t^5 \ + \operatorname{Min} \ \sum_{s \in S} |g^s - v_{24}^s|$$

				Mo	odel Output			
Day	Solver	Incumbent	Explored	Objective	Optimality	Solve	Feasible	$\sum_{s \in S}  v_0^s - v_{24}^s $
		Solution	Nodes	Function	Gap (%)	Time	without	$(m^3)$
		Objective		Value		(seconds)	slack &	
							surplus	
							variables	
1	Gurobi	20.785	61,169	155.34	50.35	100.02	Yes	155.34
2	Gurobi	9.855	659,305	87.64	0.3172	100.02	Yes	87.64
3	Gurobi	718.82	2,637	881.44	0	2.45	Yes	685.44
4	Gurobi	710.71	1	774.41	0	2.64	Yes	774.41
5	Gurobi	1,309.77	3,567	1,331.24	0	1.25	Yes	1331.24
6	Gurobi	1,760.34	46,795	2,040.37	0	11.20	Yes	1813.37
7	Gurobi	1,687.61	9,339	1,743.27	0	1.82	No	449.48

Table 7 - History results for initial objective

It is evident via table 7 that over the course of the 7-day period that the rise of the objective is quite large in the model. This drastic rise can also be seen in the absolute difference between target volume and actual volume, where the absolute differences between the target volume and actual volume in 24<sup>th</sup> hour for all silos *s* rose gradually over time with each iteration of the model. For the first 6 days it is clear that the objective is dominated by the 3<sup>rd</sup> objective component, and on the last it is dominated by the 2<sup>nd</sup> objective component. This is clear in Appendix A.7 where the cumulative volume in silo 5 throughout the schedule

is quite large. The 1<sup>st</sup> objective component only comes into play on the 7<sup>th</sup> day when a feasible solution could not be found, and hence slack and surplus variables were added. The effect of this component is similar to that of the 3<sup>rd</sup> objective component on the 7<sup>th</sup> day. The optimality gap tends to decrease with each iteration of the rolling horizon algorithm, with days 3-7 having an optimality gap of 0%. This means that the model found an optimal solution for these days, which isn't quite what would be expected with the output high objective function values. Given that these values were considered "optimal" the model was revisited to find and inspect this issue.

#### 4.3.2 Revised History Run

The 3<sup>rd</sup> objective was adapted to ensure the 36<sup>th</sup> hour was now being optimized instead of the 24<sup>th</sup> hour which may have been causing some error in the initial history run. This is because the following 12 hours (after the 24<sup>th</sup> hour) were not being optimized by this objective and were the cells being locked into the first twelve hours of the next model. Therefore, a second batch of history results were solved for this change. The new objective was:

$$\operatorname{Min} \sum_{t \in T} \sum_{s \in S} (\operatorname{suplus}_t^s - \operatorname{slack}_t^s) + \operatorname{Min} \sum_{t \in T} v_t^5 + \operatorname{Min} \sum_{s \in S} |g^s - v_{36}^s|$$

				Mo	odel Output			
Day	Solver	Incumbent	Explored	Objective	Optimality	Solve	Feasible	$\sum_{\substack{s \in S \\ 3}}  v_0^s - v_{36}^s $
		Solution	Nodes	Function	Gap (%)	Time	without	$(m^3)$
		Objective		Value		(seconds)	slack &	
							surplus	
							variables	
1	Gurobi	0	42,376	97.95	38.76	100.02	Yes	96.95
2	Gurobi	13.58	15,639	59.41	0	17.55	Yes	58.41
3	Gurobi	488.40	64,249	681.99	0	31.63	Yes	220.99
4	Gurobi	160.43	489,927	201.35	4.71	100.00	Yes	201.35
5	Gurobi	98.00	300,024	136.89	1.0661	100.02	No	5.77
6	Gurobi	51.64	298,397	141.76	17.37	100.03	No	38.67
7	Gurobi	0	167,067	31.61	3.21	100.02	No	0

Table 8 - History results for revised objective

The results in table 8 show that the objective values stay fairly consistent with a slight rise on the third day, which is due to the effect of the second objective component. This objective then falls back to what would be expected and desired following this peak. The third objective component stays fairly consistent and doesn't rise over a 221 m³ cumulative difference between the silo volumes and the target volumes over 7 days. This model does a better job of regulating flow throughout the model that we can generally see in Appendix B.1-7. The model requires the full 100 seconds as solve time for 5 out of the 7 schedules generated, which is expected with the explored nodes going beyond 10⁵. The first four solution schedules were feasible without slack & surplus variables, whereas the last three were not and required surplus and slack variables in the silo constraints to complete the run. This indicates minor breaches to the silo volume constraints towards the back end of the rolling horizon algorithm. With generally good solutions being found (as shown by the respective objective functions aside from day 3), the optimality gaps for the runs suggest that the solutions are nearly optimal for most of the schedules created.

#### 5 Discussion

#### 5.1 Model Recommendation

Gurobi has proven to be the best solver to be used on the model. This was because in both the tests conducted, it found the best objective values and had the lowest optimality gaps from the list of solvers. This was to be expected by the research done and was backed up by the Gurobi Performance Benchmarks in [3].

From the model results, it can be seen that the model is working and generating suitable, feasible solutions in a timely manner for a range of different objectives and combinations of these objectives. However, it was noticed that the models with one objective gave results that skewed heavily in terms of prioritizing this and lacked in other considerations, such as regulating flow. The purpose of the different objectives was to provide the plant flexibility with what they wanted to prioritize, whether that be aiming for a target volume in the silos or minimizing the use of purchased water for cleaning. It was also found that combining these objectives could provide reasonable, timely, and more realistic solutions that considered the different aspirations of the plant. As with many machine scheduling problems, there is never a one-size fit all solution, instead, it is circumstantial to what the plant would like to achieve.

The recommended use of the model would mainly be in conjunction with the worker's pre-knowledge of the plant. For example, the model generates provisional schedules that the workers can review and adjust (if required) to suit their intentions for the day. Blindly following the model created could have drastic implications (as with any AI related problems) and should not be the intended use for the project. This is due to some of the assumptions that have been made that may cause the model to misrepresent some parts of the plant, which were less so evident in one-off uses but more apparent with history.

The rolling horizon algorithm was an added feature to show how the model developed the solution spreadsheet over a week. From the results, it can be seen that the revised objective provided a much better set of solutions for a week at the plant than the initial run. This can be seen where in comparison to the initial run, generally, the objectives are much smaller and more consistently values that would be expected. This highlights the importance of pre-knowledge at the plant being used to verify the model, as if the initial run was used at the plant, it would've caused problems.

It would also not be recommended to use the history feature for an extended period of time due to the error that is added with each solution spreadsheet created. This is seen vividly for the revised history run, where in order to get feasible solutions for the 5<sup>th</sup>, 6<sup>th</sup>, and 7<sup>th</sup> days, slacks and surpluses must be added to the silo volume formulation to ensure feasibility. This indicates that a small breach of these constraints was allowed to ensure feasibility. This is because with every solution spreadsheet generated, the difference between the model and spreadsheet silo volumes rises, and the model no longer represents the situation at the plant as accurately by the end of the week. The recommended use of the rolling horizon algorithm would be to use it for a 5 days or as long as it produces feasible solutions up to the 5<sup>th</sup> day (using the recommended model use instructions above) before re-solving for an initial spreadsheet without using history. This would ensure that the effects of carryover error be mitigated over a longer period of time, and more consistent schedules are created with the freshly generated schedules made every 5 days.

#### 5.2 Model Limitations

As mentioned, it would be unwise to blindly follow the solution spreadsheets generated by the model and instead would require a supervisor to verify with their pre-existing knowledge. This is because some assumptions have been made in the model, which may misestimate volume levels in some circumstances, especially over time using history. One assumption that may have such effect was assuming no flow can go from machines 4a and 4b into machine 5, and instead being directed through silo 5 before being able to go into machine 5. This in practice, is slightly different to how the model was formulated. The effect of this assumption was minimized by creating an objective function that minimized the flow in silo 5. This ensured that the turnover rate in the silo was relatively quick.

Another assumption made was that milk deliveries were reused at the plant. On every iteration of the rolling horizon algorithm, the milk delivery volumes from the first twelve hours were swapped with the back twelve hours (between 25-36) to ensure the following day had the correct milk delivery set. This limited the model and the history results as tests weren't conducted on any differing milk delivery volumes, which in reality could vary on a day-to-day basis depending on the plant's needs.

Similarly, another limitation to the model was that it was only tested on one set of initial carryover volumes. This limited the model as there weren't any tests done on input data sets to ensure feasibility when there were different inputs into the model. The history component does give a glimpse into using different inputs in the model, as every day there was a new set of inputs to the model. However, in saying this disaster scenarios weren't considered meaning that the model wasn't pushed to its limits. Hence, it is uncertain to what extent the model can be applied to, to still be able to give feasible solutions consistently.

A limitation of the target objective is that the objective is only setting a target volume for the 36<sup>th</sup> hour and has not considered solutions in between. This leads to the solutions providing optimality for the 36<sup>th</sup> hour and not throughout the whole model. The potential impact this could have had wasn't apparent in the revised history run due to the low target volumes which were set as constants using the initial volume in the silos. This regulated the flow in the model quite well and didn't show the potential effect of this objective. However, the impact of this may be more apparent if the model was used with different inputs which were larger. Hence, it would also be recommended to use smaller and more appropriate target volumes for all silos to regulate flow better.

#### 5.3 Future Work

Future work to be done includes:

- Testing the model on disaster scenarios to see the extent that the model will work for.
- Using bi-objective functions on the model to optimize multiple aspects of the model.
- Adding cleaning constraints on all the machines that reflect the plant.
- Trialling a new horizon for the rolling horizon algorithm with the 13-24<sup>th</sup> hours being used as the following days first twelve hours.

#### 6 Conclusion

A mixed integer linear program has been used to generate high quality plant schedules for a production plant in NZ. The intended use of the models was to generate these solutions schedules to be applied to the plant directly. Upon inspection, the recommended use of the models would be to act as a provisional schedule to the plant, being verified by a supervisor, before being implemented in practice. This is due to some of the assumptions that have been made, that over the course of a week period may cause the model to begin to misrepresent the silo volumes. A range of different objective functions were tested, such as minimizing the use of purchased wastewater for cleaning purposes, and the preservation of volume levels in the silos on a daily basis. The plant is given full permission to use these objectives as they see fit to determine the best schedules for the site. They are also given full permission to develop the prototype further as they wish.

### 7 Bibliography

- [1] P. Doganis and H. Sarimveis, "Optimal production scheduling for the dairy industry," 2007. [Online]. Available: https://link.springer.com/content/pdf/10.1007/s10479-007-0285-y.pdf.
- [2] G. P. Georgiadis, G. M. Kopanos, A. Karkaris, H. Ksafopoulos and M. C. Georgiadis, "Optimal Production Scheduling in the Dairy Industries," 2019. [Online]. Available: https://pubs.acs.org/doi/abs/10.1021/acs.iecr.8b05710.
- [3] G. Optimization, "Gurobi 9.0 Performance Benchmarks," 2020. [Online]. Available: https://www.gurobi.com/wp-content/uploads/2020/02/Performance-Gurobi-9.0-1.pdf.
- [4] A. Touil, A. Echchatbi and A. Charkaoui, "An MILP Model for Scheduling Multistage, Multiproducts Milk Processing," 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896316311703.
- [5] AIMMS, "Implementing a Model with a Rolling Horizon," AIMMS, 6 September 2022. [Online]. Available: https://documentation.aimms.com/language-reference/advanced-language-components/time-based-modeling/implementing-amodel-with-a-rolling-horizon.html.
- [6] L. Meng, C. Zhang, Y. Ren, B. Zhang and C. Lv, "Mixed-integer linear programming and constraint programming formulations for solving distributed flexible job shop scheduling problem," 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360835220300814.
- [7] MiniZinc, "Welcome to MiniZinc," Monash University, Data 61 Decision Sciences, University of Melbourne, [Online]. Available: https://www.minizinc.org/.
- [8] I. D. a. J. H. a. M. Lubin, "JuMP: A Modeling Language for Mathematical Optimization," *SIAM Review*, vol. 59, no. 2, pp. 295-320, 2017.
- [9] JuMP, "Column Generation," JuMP, [Online]. Available: https://jump.dev/JuMP.jl/stable/tutorials/algorithms/cutting\_stock\_column\_generation/#:~:text=Column%20generation%20is%20an%20iterative,having%20to%20enumerate%20every%20column..
- [10] A. J. Mason, "SolverStudio: A New Tool for Better Optimisation and Simulation Modelling in Excel," INFORMS Transactions on Education 14(1):45-52., 2013. [Online]. Available: https://doi.org/10.1287/ited.2013.0112.
- [11] Q. Huangfu and J. A. J. Hall, "Parallelizing the dual revised simplex method," Mathematical Programming Computation, 10 (1), 119-142, 2018. [Online]. Available: DOI: 10.1007/s12532-017-0130-5.

[12] J. Forrest, T. Ralphs, H. G. Santos, S. Vigerske, L. Hafer, B. Kristjansson, M. Lubin,
 S. Brito and M. Saltzman, "coin-or/Cbc: Release releases/2.10.8," zenodo, 2022.
 [Online]. Available: https://doi.org/10.5281/zenodo.6522795.

## Appendix A

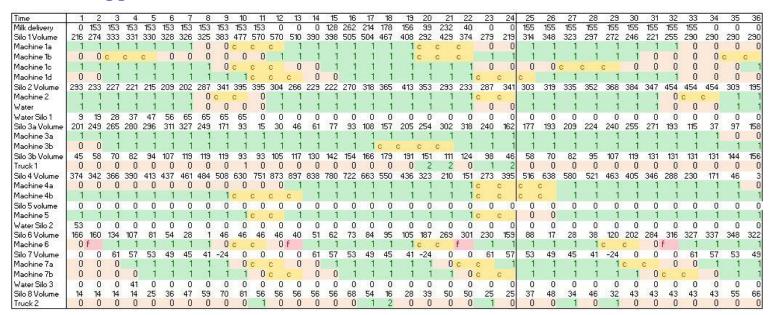


Figure A.1: Day 1 Spreadsheet Solution

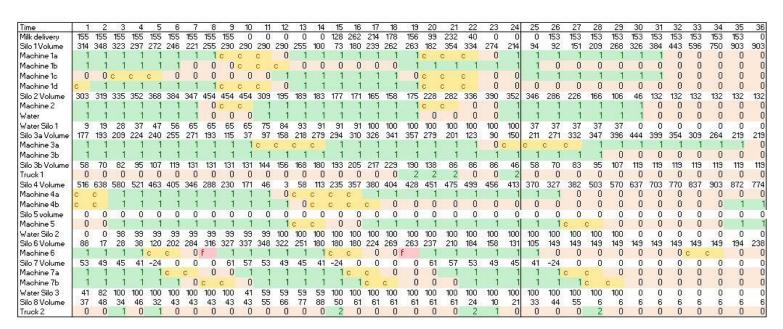


Figure A.2: Day 2 Spreadsheet Solution

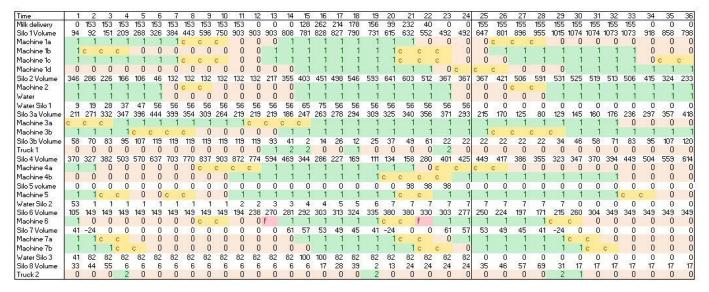


Figure A.3: Day 3 Spreadsheet Solution

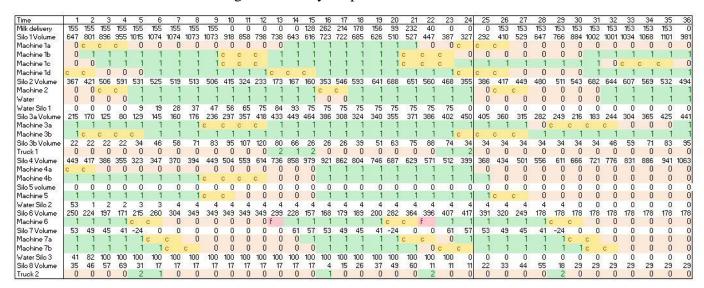


Figure A.4: Day 4 Spreadsheet Solution

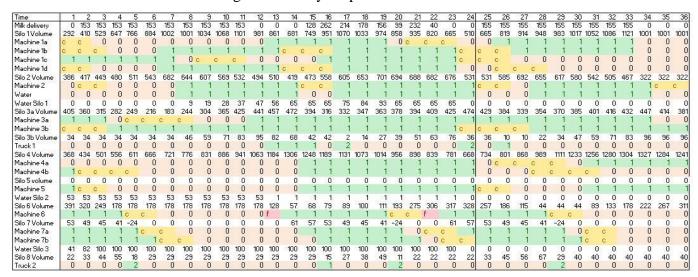


Figure A.5: Day 5 Spreadsheet Solution

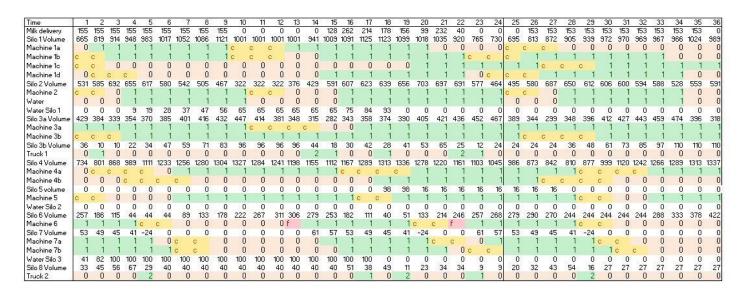


Figure A.6: Day 6 Spreadsheet Solution

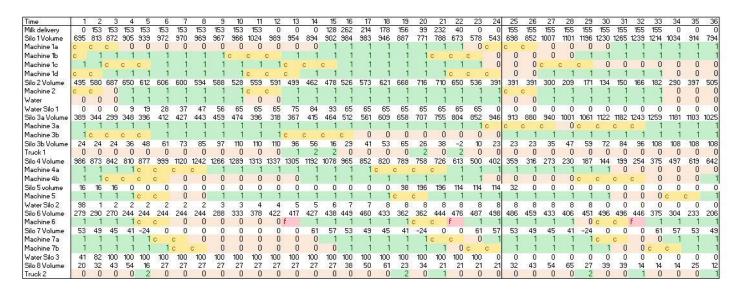


Figure A.7: Day 7 Spreadsheet Solution

## Appendix B

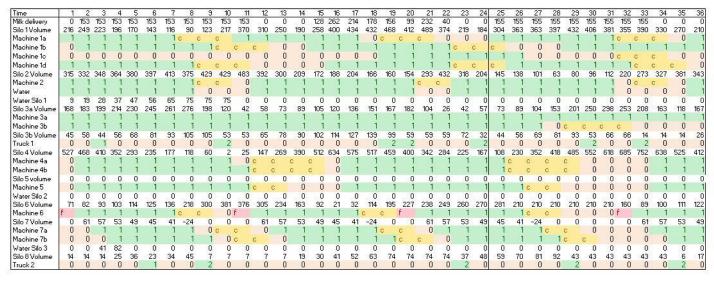


Figure B.1: Day 1 Spreadsheet Solution for updated objective

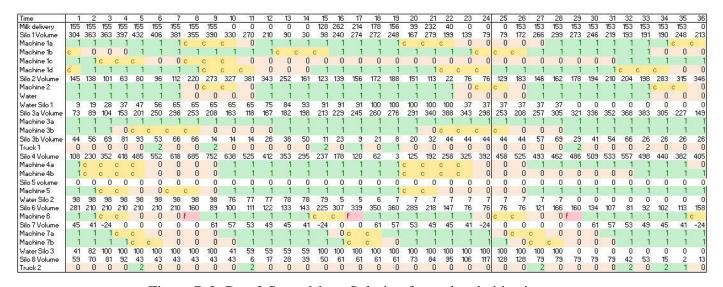


Figure B.2: Day 2 Spreadsheet Solution for updated objective

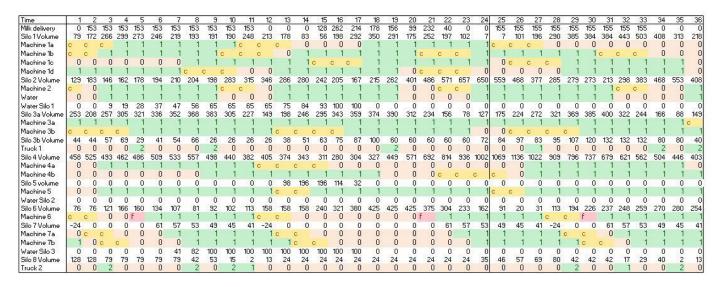


Figure B.3: Day 3 Spreadsheet Solution for updated objective

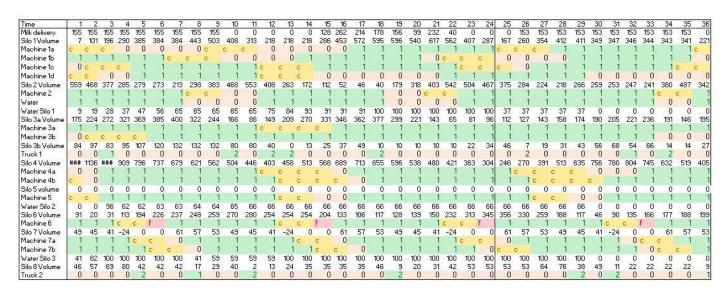


Figure B.4: Day 4 Spreadsheet Solution for updated objective

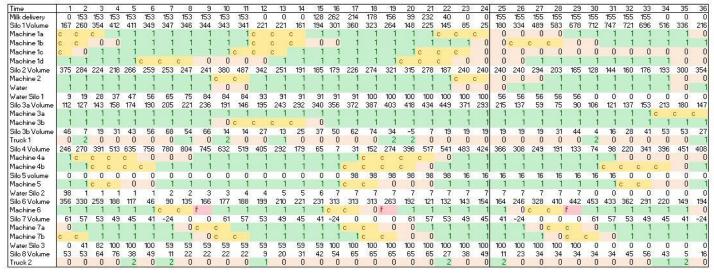


Figure B.5: Day 5 Spreadsheet Solution for updated objective

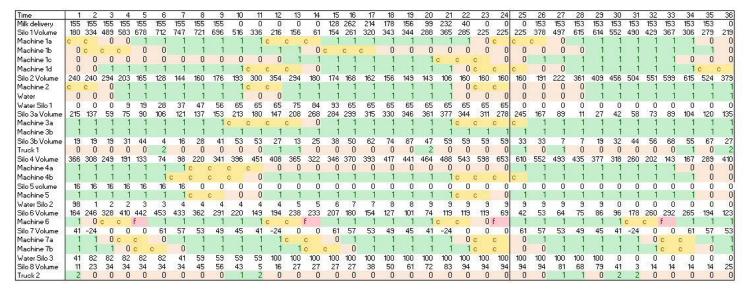


Figure B.6: Day 6 Spreadsheet Solution for updated objective

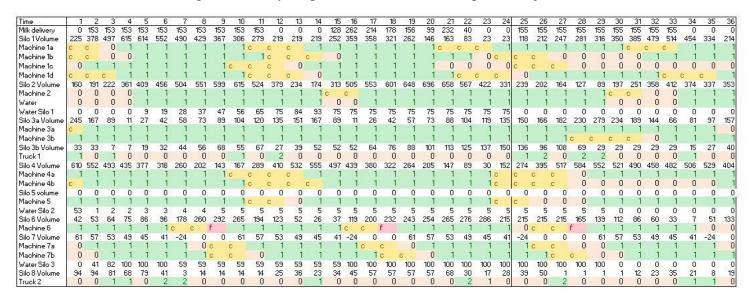


Figure B.7: Day 7 Spreadsheet Solution for updated objective