REWARD-Data Analysis and Water Flow Scheduling

Karthik Sundaram[1] Sai Jishna Pulluru[2] Renzo Akermann[3]

Abstract— **1) Analysing the data (for a time span of 7 days) of a cheese process plant which involves: a)producing Gantt charts of the schedule of each batch process of the different cheese lines b)plots of the fresh water consumption assuming a discrete water flow c)plot of whey(waste) input on a time horizon assuming a function for flow in post-processing(purifying) d) calculating the amount of water consumption during each CIP run and plotting in Tableau**

This part is important to get a sense of the raw date log files to graphically represent the batch schedule after manipulating the data and constraining the with small snippets of codes. Also it is used to find out the minimum buffer storage capacity of the tank.

2)Mathematical modelling for scheduling of a multipurpose plants that has both batch and continuous processes, based on the resource-task network (RTN) , global time point continuous representation, for different scenarios of cleaning and storage. It also involves post processing of plotting the gantt charts.

A skeleton model without a water flow and with only the various resources and tasks was created initially and a water layer of storage and cleaning has been added to this primary framework. It has incorporated both , the continuous and the batch models in a unified framework.

3)Analysing the possibility of applying pinch analysis as a heuristic for implementing in the scheduling model.

I. INTRODUCTION

The main aim of this approach is to address and redress the situation of water scarcity using water reuse . As can be seen

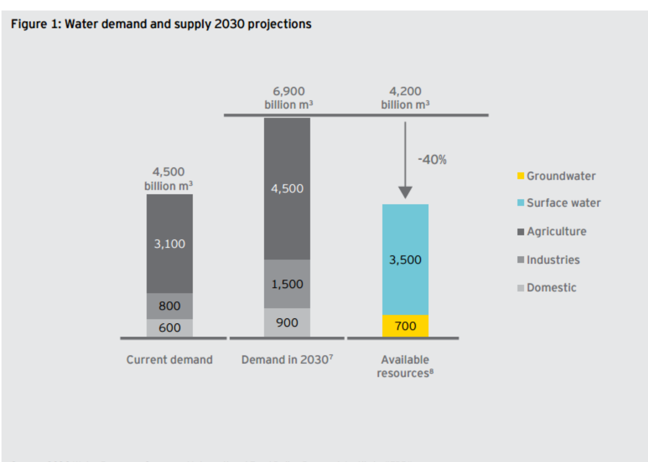

Fig. 1. Water Scarcity Diagram

in the above case water demand is seen to increase according to the forecasts whereas the supply is remaining a constant. Hence in order to sustain in the long run economically and as a part of the environmental conscious motive, all the companies should employ and concentrate more on waste water management and also improve their processes to reduce the initial water consumption. This can be done by a lot of data analysis techniques and mathematical modelling to optimise the water usage. For doing this the first steps should be to analyse the current scenario by representing the whole company in a mathematical model and collecting data including the time logs and amount of water consumed with respect to its contaminant concentration.

Waste 1) whey production 2) cheese rinse

3 types of water1) processing water 2) cleaning water 3) sanitary water

3 steps in CIP 1) pre rinse 2) caustic wash 3) final cold rinse

Waste processing1) discharge treatment 2) waste disposal 3) onsite treatment

Then as is my first part of my project , analysis should be done to get a pictorial representation of the scheduling of water flow. The earlier waste water treatment methods are used to ensure that the proper levels of allowed discharge into the water bodies .As it is directly proportional to the amount of waste water being treated, the cost goes up with more water discharge. Hence it is considered wise to produce less waste water which needs a costly treatment. This can be possibly done by reducing the waste water production by using efficient methods scientifically or rather produce lesser quantities of the same and hence reduce fresh water consumption. But ,the inherent discontinuous nature of batch processes complicates the development of a methodology to minimise wastewater, since the methodology not only has to take inlet and outlet concentrations into consideration, but also the times at which water is used during the operation unlike the continuous. Its of major importance to analyse the process and classify them as continuous or batch and adopting the correct set of time point system like RTN or STN and make use of the appropriate insight based techniques or the mathematical modelling technique or a combination of both as is attempted in this work, where pinch analysis, a insight based technique is tried for its use as a heuristic for getting a good enough initial solution to optimise further. Minimisation of water can be done by initially scheduling or using a fixed schedule and then minimising or both can be done simultaneously. Due to complexity of the model a MINLP model arises, but in this work where modelling has been done, the non-linear terms have been neglected in the objective function.

The project at hand is of a cheese plant (Arla, Denmark) where water minimisation is desired. It has 3 cheese lines and an RO plant where the input whey or the initial waste water is sent for further purification and reuse for further usage. Planning schedule: Lead time monad to moday Plan Tuesday to Monday Transportation plan Wednesday to Monday Master plan Sunday to moday (more than a week)

It can be approached primarily by organising, analysing and interpreting the data collected and drawing inferences and conclusions like the optimum tank capacity in this project. After the gantt plots of the different schedule has been plotted , a rough idea of the relevance of the different models to choose the appropriate time representation (continuous or discrete) , material balances(STN or RTN), event representation like global time points(discrete or continuous) or unit specific or time slots or precedence relations. Then mathematical modelling optionally with heuristics is developed accordingly.

This research is part of the REWARD project, a large interdisciplinary research project, combining real-time monitoring, sensor development and quality and safety guidelines in pro-active industrial water management funded by the Danish Council for Strategic Research.

II. Analysing the data and finding the efficient capacity for storage

A. THE FACTORY LAYOUT OF THE ARLA PLANT

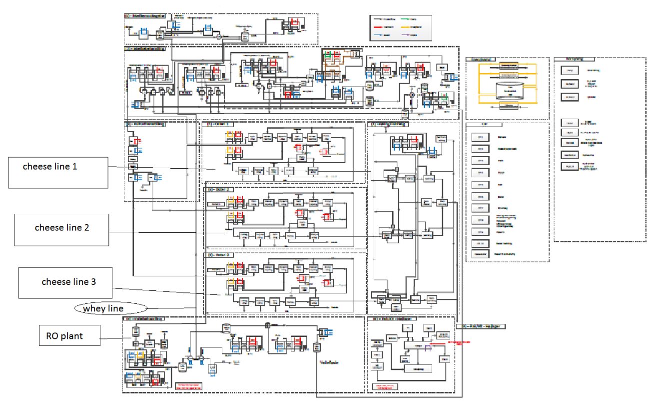

Fig. 2. Factory Layout

The layout of the project includes three cheese lines, a RO plant and 10 Cleaning in Process units which are used to clean the units.

B. DIFFERENT PROCESSES IN THE ARLA CHEESE PLANT

1) Schedule of each batch process of the different cheese lines: Producing gantt charts with the overall data available between different quantities on the y axis with respect to time gives a basic idea of how the processes are occurring on a uniform time scale and also to get a rough idea whether the solution generated in the mathematical model is in line with the actual schedule. Also ,because data generally available in the real time would not be ideal, hence plotting such charts by combining all the relevant data from different files gives an idea about such a deviation as this cannot be interpreted from just the log of the start and

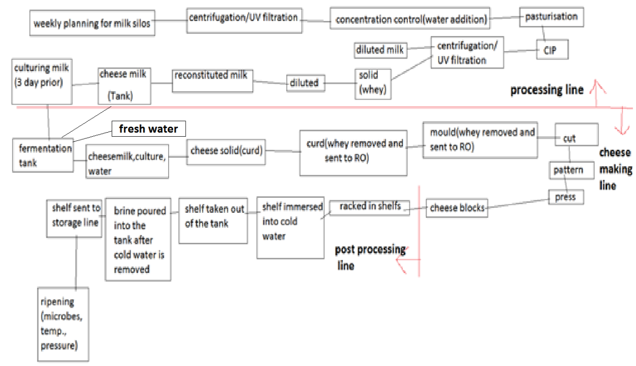

Fig. 3. Arla plant- processes

end dates. Different excel files which had all the necessary information were clubbed together by using the Vlookup function. Calculation for plotting the gantt chart on a uniform horizon: (DATEDIFF('second',[Log Open Date],[Log Close Dt]))/86,400 For each cheese line: As can be seen above,

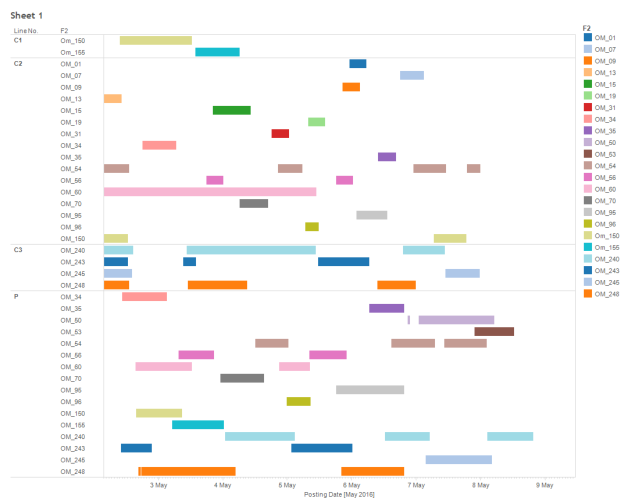

Fig. 4. Gantt chart-cheese plant

Cheeseline 2 has lot of processes scheduled throughout the week, whereas as cheeseline 1 is not .This can be rectified by redistributing the processes in cheeseline 2 into chesseline 1 which has only 2 type of raw materials input to produce only 4 campaigns running 1 in OM150 and 3 in OM155.This is possible because line 1 is free after 4th May.If the cheesline 1 has a limitation of handling only select raw materials, the process should be made dynamic where different raw materials can be handled by changing a set of parameters.

For all the cip lines:

2) Plots of the fresh water consumption: After importing the excel file containing the fresh water consumption, for all the processes, it is plotted as a bar graph: As can be seen from the fresh water and whey graph, 2nd May,7th May has the least consumption and production of water and the remaining days more or less have the same usage of fresh water. It

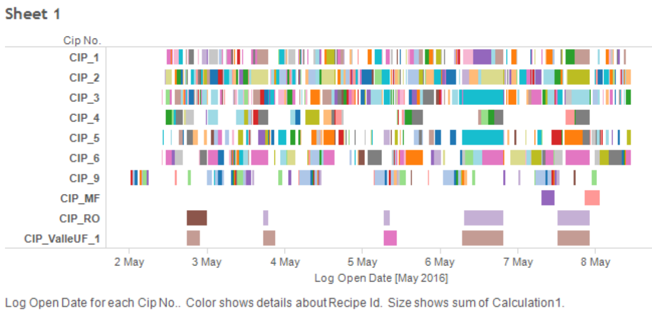

Fig. 5. CIP-Gantt chart

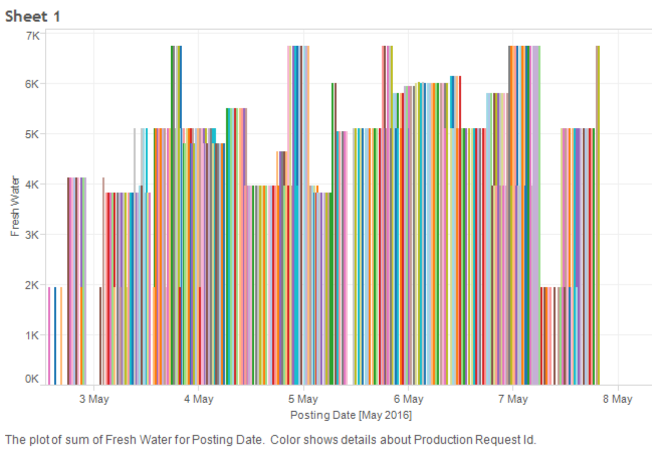

Fig. 6. Fresh Water Plot

would be better if 2nd and 7th May has more processes from the processes planned for the other day, so that the storage or handling of fresh water capacity can be reduced and hence cost savings. Load on 3rd may could be reduced by evenly distributing the processes to the other days.

day in may	fresh water used
7	251,700
6	5,368,004
5	5,192,770
4	5,836,267
3	7,459,441
2	3,598,607

Fig. 7. Fresh Water Table

3) whey(waste) input over time: Function assumed: F(t) - Quantity of whey that is given as input for the unit that is purified as RO water(whey is generated and given as input to RO plant at the rate of T=2hours per process) Q- total quantity of whey produced in the whole process T is split into 1 time point/minute (discretised)

$F(t)=0; t < T/4$

$=Q/2; t=T/4$

$=Q/180; T/4 < t < T$

The reason for taking the above form of input is be-

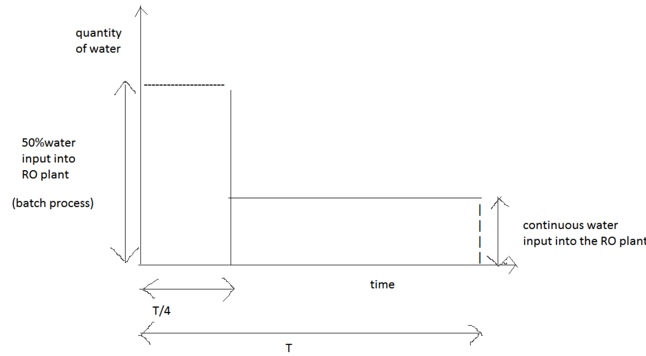

Fig. 8. Whey output graph

cause initially fermentation takes place in any campaign of batch processes, hence till T/4,no whey is produced. At t=T/4,similar to a batch process ,whey is output which constitutes 50percent of water or whey that comes out of the processes. After that from t=T/4 till T water is given as output to the RO plant as a continuous process.

After calculations in excel using Vlookup function, the whey quantity is calculated for each of the time points defining and breaking the function into different cases.(at T/4, T/4 to T). For accommodating the different time horizons of each process into a unified scale, each log date is assigned a unique number equivalent and is imported into Matlab for further processing of the data: (Appendix 1) This csv file is then plotted on Tableau to generate the overall whey input vs time for the concerned length of 7 days for the different cheese lines. As can be seen from the above graph, whey

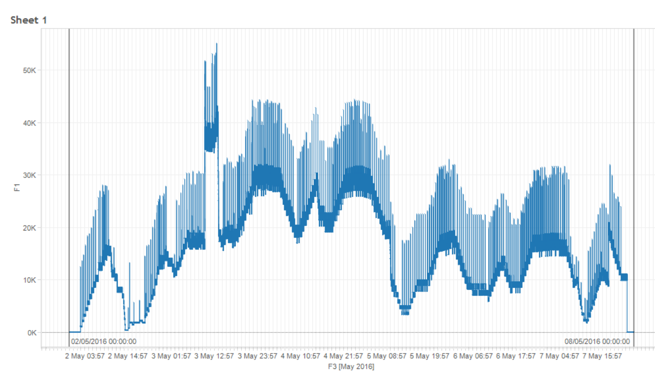

Fig. 9. Whey plot - week period

production is the highest on 3rd may corresponding to the highest usage of fresh water.

As the quantity follows the following equation:

Raw milk-cream=cheese milk Cheese milk+ fresh water +culture milk +miscellaneous-cheese=whey, A direct correlation can be seen between fresh water and whey production more or less with fresh water being input at the start of the batch under the particular campaign but whey is input following the function assumed here.

Similarly for each of the 3 cheese lines having different set of batches the following code and the likes would be useful: (Appendix 2) The buffer storage capacity of tanks in the RO plant which gets the input as whey is having a maximum capacity already considered in the plant on a 20 hr. period. According to the above results from the data being analysed and from the matlab code, it has been found that for a one hour buffer, for storing 20hrs. of water 77.5 tonne capacity tank is required. This was done by calculating the 20hr total consumption with different staring points starting from the local maximum for a considerable period with adjacent points also being more or less the same water flow rate. (Appendix 3)

	Water consumption for the week
CIP_1	658.5734154
CIP_2	956.3107768
CIP_3	1,116.508535
CIP_4	849.6891948
CIP_5	649.7761057
CIP_6	149.0969387
CIP_8	12.02448243
CIP_9	333.6409185
CIP_10	1,106.304243

Fig. 10. Water consumption for the week

4) water consumption during each CIP run : From the data it can be inferred that water consumption for CIP was highest for Cip-3 which is for valle (whey) and least water is used for Cip-9 which is for frame washing.

Mapping of the CIP logs was done. Certain calculations, using average consumption patterns were used to calculate an approximate amount of water consumption during each CIP run.

Assuming a constant function, the data is plotted in Tableau. To plot on a uniform time horizon, a scale was generated with each second having a unique general value equivalent of the date time format. The discretisation of the time horizon is done by,

F2=actual consumption*502,524/duration of each run*58,163

III. Mathematical modelling

[1]

A. The relevance of choosing this model for the project

The following model was chosen because from the tableau plotted, it can be seen that a continuous and batch processes together can be represented in a single mathematical formulation like in this paper. This paper had a similar process but for

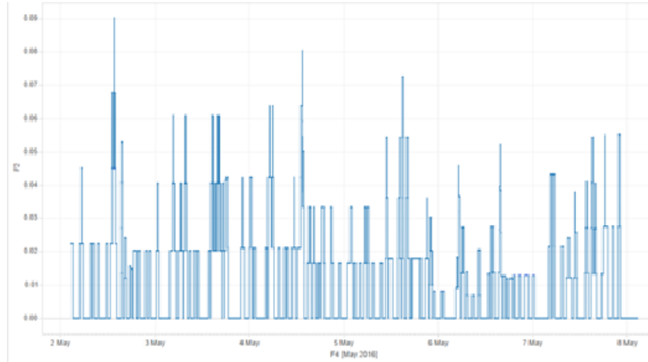

Fig. 11. Water consumption for all CIP runs

a different application whose relevance has been exploited in the cheese line scheduling model. Here different types of raw materials are continuously mixed in mixer units with certain intermediates in between them. It is then packed in different product lines to produce the final products. This constitutes the continuous process which is same as the campaigns run on the 3 cheese lines. The whey that is output for converting it into useful RO water in whey handling has both the batch process of a sudden water in flow at the 50 percent mark of the process and a constant continuous water in flow in the RO plant for the rest of the time in a continuous manner. Also both the scenarios involve cleaning-CIP (Cleaning In Process run)in the cheese line model and cleaning of the packing line in the model considered in this research paper. Also the global time point being used in the paper considered is useful in the cheese plant also because by using global time intervals, calculation for each point on the time horizon line need not be done which though might give a better solution would become tedious in a large scale real world problem like the cheese, line that is being considered here. Also as both the processes involve a lot of resources and intermediates and raw materials and also keeping a track of these resources is of significance RTN implementation to this project is of great relevance. Also the characteristic curve that was used for the whey production has a curve that has both the batch process at t=T/4 and continuous curve from T/4 to T.

Hence the formulation uses a RTN,uniform time grid representation and also a mixed integer linear programming model is used to solve for a optimal number of event points.

B. The basis and relevance of selecting the framework that has been chosen for this model from the literature

1) Castro et al.,-An Improved RTN Continuous-Time Formulation for the Short-term Scheduling of Multipurpose Batch Plants.It has been shown that the STN-based continuous-time scheduling formulation of Ierapetritou and Floudas ,1998 is less accurate, as it violates time horizon constraints in certain cases.[2]

2) Maravelias and Grossmann 2003 In this work, the approach used in practically all continuous STN models, started with a small number of intervals and then the number

of intervals was iteratively increased by one until there was no improvement in the objective function for a fixed number of iterations. Global time point approach for scheduling batch process plants disucssed in this paper was also utilised in this project. [3]

3) Iiro Harjunkoski et al, -It employed the RTN process representation since the unified treatment of resources allows for a focus on fewer sets of constraints. In a case study of the scheduling model of the Dow Chemicals, usage of RTN model led to low cost solutions ,cost reduction by waste reduction and the all the treatment units were used to their fullest efficiency. [4]

4) Munawar A. Shaik et al., 2007, -In contrast to the STN representation, all the resources such as materials, processing and storage equipment, and utilities are uniformly treated in the RTN representation. STN model often leads to storage violations in the Gantt chart in real time, unless one considers storage as a separate task. The proposed RTN formulation overcomes this intricacy with the additional sequencing constraints. Compared to the STN-based model, the proposed RTN-based model gives better LP relaxation (due to the tightening constraint used), and even though the RTN version has more continuous variables and constraints, it solves relatively faster in some instances and it performs equally well in other instances of the examples considered for the case of unlimited storage capacity. [5] 5) Mendez et al. 2006-the RTN-based formulations employ a uniform treatment and representation framework for all available resources through the idea that processing and storage tasks consume and release resources at their beginning and ending times. The major advantage of the RTN formulation over the STN counterpart arises in problems involving identical equipment. Here, the RTN formulation introduces a single binary variable instead of the multiple variables used by the STN model and hence complexity is reduced in the RTN model. [6]

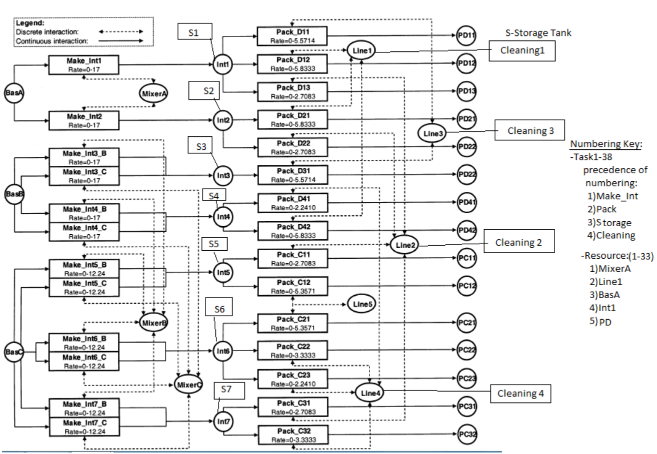

Fig. 12. Arla plant layout with cleaning and storage units included

C. Nomenclature

Sets/Indices: C, c) equipment conditions
I, i) process tasks
I b) batch tasks
I c) continuous tasks
I emr) tasks that must exceed a certain minimum rate
Ic) tasks belonging to condition c
I s) storage tasks
I ZW) tasks subject to zero-wait policies
R, r) process resources
Rcl) equipment resources requiring cleaning
Req) equipment resources (storage tanks not included)
Rint) resources corresponding to intermediate products
RPK) equipment resources corresponding to packing lines) event points

Parameters:
Rro) initial amount of resource r
H) time horizon
cif) cost of executing task i (fixed cost term)
civ) cost of handling a unit amount of material in task i (variable cost term)
cr) value of resource
cjr) cost of holding a unit amount of inventory of resource r over a unit period of time
Rr max) maximum availability of resource r
Rr min) minimum availability of resource r
Vi max) maximum amount of material processed by task i
Vi min) minimum amount of material processed by task i
deltat) maximum number of event points allowed between the beginning and end of a given task
delmin) minimum interval length
lambda) coefficient for the rate of generation of resource r by task i
mu) coefficient for the binary extent of resource r at the start of task i
mubar) coefficient for the binary extent of resource r at the end of task i
Rhomax) maximum allowable rate of task i
Rhomin) minimum allowable rate of task i Variables
Nbar) binary variable that assigns the end of task i, which began at t, to point t0
R) excess amount of resource r at event point t
T) absolute time of event point t
zetabar) total amount of material processed by task i, starting at t and finishing at t0

D. Tuple used to store 3-D data in Excel

The cplex output of Nbar is stored in the excel as a tuple. initialise a tuple for the 3D variable Nbar as sheetwrite cant be used directly. (Appendix 4)

E. MATLAB code for interfacing Cplex and Tableau

Following is the the case of No Storage: After importing the cplex decision variable Nbar into Matlab, the following code is executed to eliminate Nbar values of 0 and choose Nbar value of 1 and corresponding time point numbers are changed to the actual time.

(Appendix 5) The above file written in the csv format is used in Tableau to plot the gantt chart.

F. Cplex code and tableau plot of the output in cplex

Time horizon(H)=120 hrs.
No. of time points considered = 8
Assumption:
no cost for cleaning and storage considered in the profit. lambda,mu,mubar values should be defined as given below for any set of task, resource:

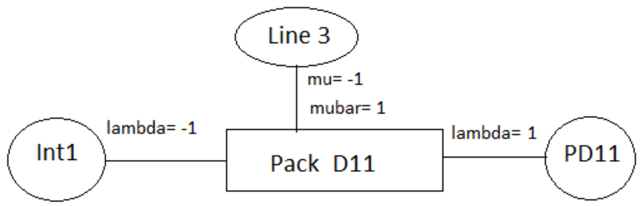

Fig. 13. RTN representation-(Appendix 6)

The four scenarios:
1.without storage and cleaning
2.with storage
3.without storage
4.with cleaning and storage

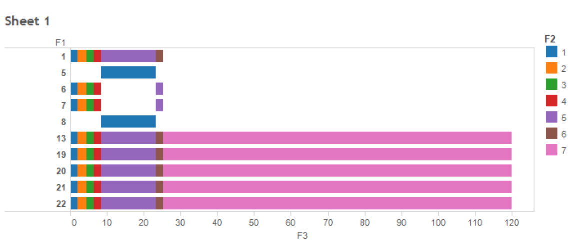

Fig. 14. Without storage and cleaning- (Appendix 7)

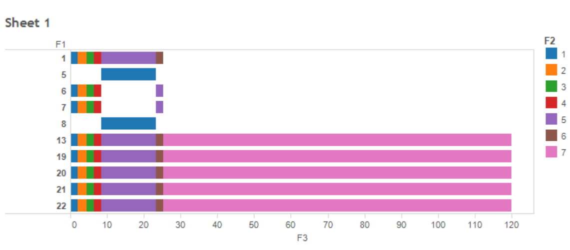

Fig. 15. With storage-(Appendix 8)

.

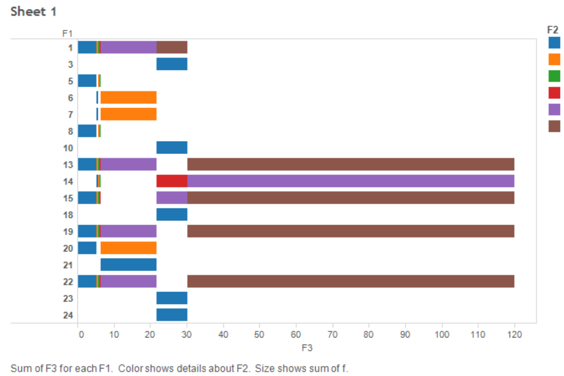

Fig. 16. Without storage-(Appendix 9)

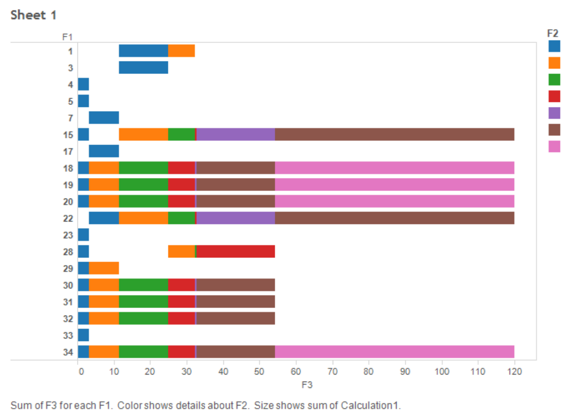

Fig. 17. With cleaning and storage-Appendix 10)

Overall comparison:

	Profit	Event points	Constraints	Continuous variables	Binary variables
without storage and with cleaning	22,381.77	8	1616	462	217
without storage and without cleaning	22,583	8	1536	462	189
with storage and without cleaning	23,274	8	1578	462	238
with storage and with cleaning	64,587.5	8	1774	490	266

Fig. 18. Comparison of all cases of the model considered

IV. Analysing the possibility of applying pinch analysis as a heuristic for implementing in the scheduling model

A. Conclusion and relevance of water surplus diagram to the cheese production model

Pinch analysis is an Insight-Based Techniques for Water Minimization in Batch Processes. Hence, it would be best to find the pinch point using the water surplus diagram as a initial solution (Heuristic) and follow the guidelines:

1)In order to achieve the flowrate targets determined above, it is necessary to observe the pinch division. This means that water sources above the pinch including fresh water may not feed demands below the pinch and may also not mix with sources that are below the pinch concentration.

2)When it is above the demand composite, this is a region with a surplus of pure water i.e. water is available with a purity greater than what is required. When it is below, this is a region with a deficit of pure water i.e. water is required with a purity higher than what is available.

both of which can be incorporated as constraints.

The pinch analysis becomes relevant and can be compared to the cheese line model.The source here being the whey and demand is the CIP line which needs water for cleaning and also the RO plant which needs water for purifying. Fresh water being available always.

In this case of the cheese line, source is formed by the whey output and for any iteration needed to alter the source curve such that there is no water deficit, fresh water is used to compensate initially such that the water surplus diagram curve always lies in the first quadrant only. Hence after the fresh water needed for the CIP line is added, the curve would have water surplus except for the pinch point. This water could then be sent to the RO plant for purification. This way, the Cip can be done with a minimum amount of fresh water. Also if RO processing is cheaper than fresh water, the initial flow rate assumed can be lesser i.e., the initially assumed quantity of fresh water can be reduced and compensated by the RO processed water that takes input from the whey line and maintains the water surplus curve in the positive first quadrant. This depends on the trade off between the fresh water cost at optimum profitable level, the amount of water that needs to be given as input to the RO plant for usage in the CIP process .This makes sure that this whole process of reusing the regenerated whey water from the RO plant is profitable. But in the case that whey is useful and valuable than the RO processing cost, line (2) will not be present and all the whey water goes to the RO plant and only then it goes to the CIP cleaning.In that case only fresh water is added to offset the negative portions of the water surplus curve.

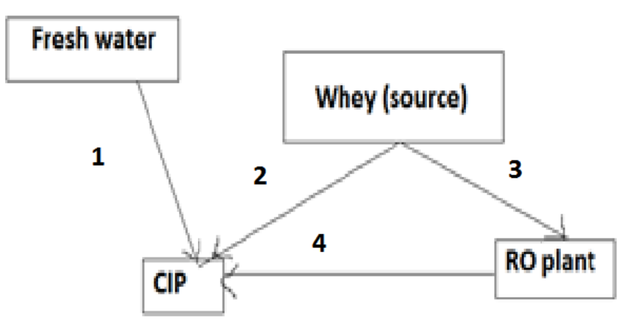

Fig. 19. Water flow path

Sorin and Bedard 1999. [7] correctly observed that the approach of Dhole et al. 1996. [8] (the conventional pinch analysis method based on heat integration method) actually defines many local pinch points, but not necessarily the true or global pinch point.To account for the different set of such processes in the actual cheese point which consists of a large number of such sets, attaining a global pinch point again becomes a difficulty.This is however overcome by the water surplus diagram. The region above pinch point has no waste water generation and the region below the pinch point does not need fresh water.

V. CONCLUSION

Fresh water consumption is unevenly spread throughout one week time horizon of concern. according to the current schedule of the Arla plant as is evident from the gantt charts plotted.

If this is evened out, the maximum capacity of the tank required would reduce which basically lies in altering the buffer time of the water tank. The current capacity as for the current plant interpreted from the data can even be reduced.

The chosen RTN, network and global time point system have proved to be effective in multiple identical equipment modelling and this model did not have any storage violation and the computation time and complexity are less. Even though the RTN version has more continuous variables and constraints, it solves relatively faster in some instances and it performs equally well in other instances of the examples considered for the case of unlimited storage capacity. Hence for all these reasons and the fact that a complex industry can be better represented with resources (even identical parallel ones) and equipments and processes separately justifies the suitability of this model for future research work on the same.

Also the whey production is maximum on one of the days disproportionately, hence scheduling the model effectively can even out the whey which can be sent to the regeneration plant and swiftly reused according to the corresponding contaminant concentration allowed.

Also there seems to be a lot of deviation from the planned timeline of execution in the Arla production schedule. This could cause problems as a plant even before the optimum possible schedule is not following the schedule and even if an optimum solution is developed, improper execution could cause cascading of water and overlapping of schedules. This could possibly be avoided by using the manufacturing system simulation which accounts for the real world uncertainties in the cheese plant.

Coming to the numerical interpretation and significance of the profits that are viable in the scenario run in this project suggests that incorporating storage containers increase the profit about three fold in comparison to the case where storage is not used. Though in both the above scenarios, cleaning is considered.

Also when storage is used the machines seem to be continuously running throughout the time horizon without having a lot of batches. This could possibly be because the number of time points considered here is 8, and more points could imply more batches instead of a predominantly continuous process throughout the timeline as it leads to the underutilisation of the parallel processes, in this case predominantly when two or three packing units are available for an intermediate (resource) preceding it only one packing unit (task) is utilised.

REFERENCES

[1] PM Castro, AP Barbosa-Pvoa, HA Matos,Simple continuous-time formulation for short-term scheduling of batch and continuous processes,2004

[2] P. Castro,A. P. F. D. Barbosa-Po voa, and H. Matos,LiAn Improved RTN Continuous-Time Formulation for the Short-term Scheduling of Multipurpose Batch Plants, 2001.

[3] Christos T. Maravelias and Ignacio E. Grossmann,New General Continuous-Time State-Task Network Formulation for Short-Term Scheduling of Multipurpose Batch Plants, 2003

[4] Iiro Harjunkoskia,Christos T. Maraveliasb, Peter Bongersc,j, Pedro M. Castrod, Sebastian Engelle, Ignacio E. Grossmannf, John Hookerg, Carlos Mndezh, Guido Sanda, John Wassicki, Scope for industrial applications of production scheduling models and solution methods,2014

[5] Munawar A. Shaik, Christodoulos A. Floudas,Unit-specific event-based continuous-time approach for short-term scheduling of batch plants using RTN framework,2008

[6] Carlos A. Mendez a, Jaime Cerda b, Ignacio E. Grossmann,State-of-the-art review of optimization methods for short-term scheduling of batch processes,2006

[7] M. Sorin, S. Bdard , The Global Pinch Point in Water Reuse Networks, 1999

[8] Y. Yorozu, M. Hirano, K. Oka, and Y. TagawaRV Dhole, N Ramchandani, RA Tainsh, M Wasilewski ,Pinch technology can be harnessed to minimize raw-water demand and wastewater generation alike, 1996