8th International IFAC Symposium on
Dynamics and Control of Process Systems
Preprints Vol.2, June 6-8, 2007, Cancún, Mexico

A HYBRID PLATFORM FOR REFINERY SIMULATION WITH CASE SWITCHES

Jianlie LI, Gang RONG[1], Yiping FENG

National Laboratory of Industrial Control Technology,
Zhejiang University, Hangzhou, 310027, China

Abstract: A new process simulation platform for refinery plant simulation and its application are introduced in this paper. Based on simulation of a crude oil distillation process, we provide a detail analysis of Flexible Multi-Case Data-drive Simulation (FMCDS), which employs the idea of discrete event simulation and would be helpful for the improvement of simulation accuracy and efficiency. The methodology used to mining the similar cases successfully reduce the number of unnecessary cases and provide a reliable case list for the platform. And a comparison between the proposed model and the static model is also presented by checking the degree of fitting between simulation data and actual plant data.

Keywords: simulation, modelling, hierarchical systems, integration, industrial control, laboratory techniques, data mining.

1. INTRODUCTION

Process simulation technology, whose main purpose is to provide an environment close to reality, has been developed for decades. As other applications of this technology, refinery models can be divided into two major categories: static model and system dynamic (SD) models(Rabelo, Helal et al., 2003). There have been many researches in both static and dynamic simulation, and accordingly some software like Aspen Plus, Aspen Dynamics and gPROMS has been developed. However, the problem of simulation efficiency cannot be ignored. Although plant-wide dynamic simulation of the refinery can provide detail information of the internal production process, mechanism and close-to-reality operation conditions, the cost and difficulty in the dynamic modelling are usually unacceptable. Using rigorous models for a long-term refinery planning might be overkill and there are still insufficiencies in the static simulation.

This paper attempts to give a brief introduction of a new simulation platform in the virtual factory laboratory system established by National Laboratory of Industrial Control Technology, Zhejiang University (China). The virtual factory laboratory system expanded the conception of simulation trainer, established industrial process model base, real-time database and management database on the basis of the integration of physical simulation and digital simulation, and built an integrated laboratory simulation platform of Process Control System (PCS) and Manufacture Execution System (MES)(Feng and Rong, 2005). The static models offered static working spots for dynamic simulation, while the dynamic simulation offered a testing platform for data processing software and advanced process control algorithm(Fang, Feng, et al. 2006). Because these is a big challenge to closely coordinate a range of individual activities to achieve overall corporate objectives(Bassett, Dave et al., 1996), and moreover, to find an acceptable tradeoff between accuracy and facility, we propose a new simulation method, i.e., Flexible Multi-Case Data-Drive simulation (FMCDS) to simulate the process, which means establishing dynamic simulation models for key units, while built the Multi-Case Static model for long- term refinery planning and other applications. The hierarchical treatment of time

[1] Corresponding author
E-mail address: grong@iipc.zju.edu.cn

and implementation of discrete event simulation enables the FMCDS to represent industrial process of refinery flexibly and generate reliable data for both planning and scheduling. Operation data from an industrial CDU is used in this paper to give a comparison between static models and the platform we proposed, and our results show that this platform enhances both efficiency and accuracy of pure static models. Simultaneously, the technology of data mining is employed to discover the familiar cases from plant data, thus providing an authentic case list for the FMCDS.

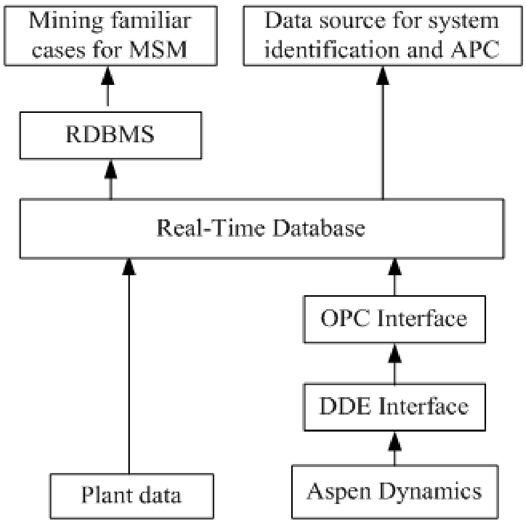

Fig.2. Data integration for FMCDS

2. DYNAMIC SIMULATION IN THE MULTI-CASE DATA-DRIVE SIMULATION

Dynamic simulation has its origin in the control engineering work of Jay Forrester(Forrester, 1958), which include variables like flow rates, temperatures, pressures, and compositions of all streams. The dynamic simulation in the FMCDS is used to provide detailed data of some key units. As illustrated in fig 1, Fang(Fang, Feng et al., 2006) has built a simplified crude oil distillation process in Aspen Dynamics. After actual unit attribution and operation parameters are carefully adjusted, these dynamic models can be used for supporting the control system design and operation optimization.

With shorter simulation intervals, i.e. 1 minute, and real-time property, the result of dynamic simulation is stored in the Real-Time database. And these data has been used to supporting the advanced process algorithm test and system identification experiment for graduate students. After extending the real-time data with spatial and temporal scales, these data are integrated into relational database management system (RDBMS) for Multi-Case Static Model (MSM). The scheme of data integration is shown in fig 2.

3. MULTI-CASE STATIC SIMULATION

Except for the dynamic simulation of some specific units and some real hardware in the platform, Multi-Case Static models of the rest units of a refinery are constructed. On the basis of static simulation platform for a refinery(Pei and Rong, 2005), these models are built to solve the problem of case switches. By calculating the yield and movement relationship of up-down stream of the operation units and tanks dynamically, the Multi-Case Static models can be aggregated to obtain the multi-period flowsheet simulation models, which is able to generate a reliable data source for research such as long-term planning and multi-period data reconciliation. The following situations are the key problems that should be considered in the simulation strategy.

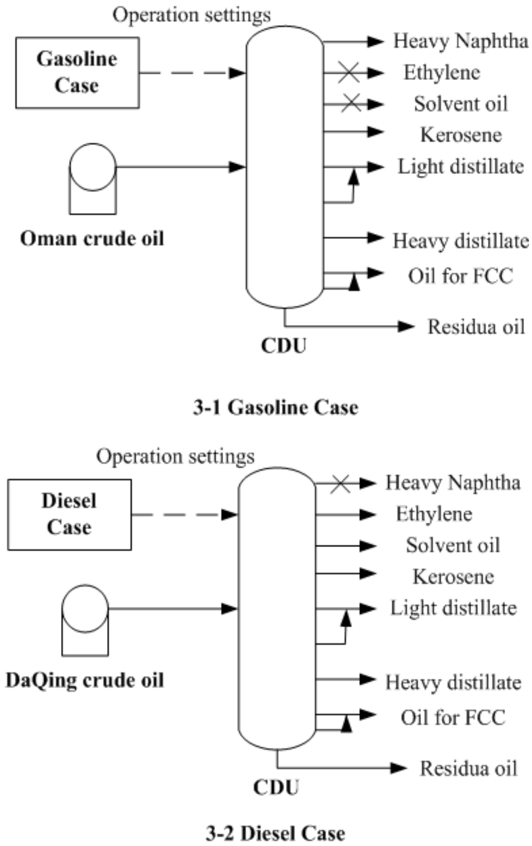

Fig. 3. Different unit configurations of cases

3.1 Case switches between shifts

When a refinery model is constructed and implemented for a long-term refinery planning, ordinary simulation strategy describes the refinery as a static model. However, there are many production cases such as diesel oil case, gasoline case for CDU and FCC according to the crude oil properties and economic parameters(Li, Hui et al., 2005). Each case has a set of predetermined operation settings, as shown in fig 3, not only the production rate but also the sidetrack configurations will change simultaneously.

Usually the operation data are calculated every 8hrs (a shift) to form statistic data in the simulation of MES, so the minimal time interval of Multi-Case Static model is also limited to 8hrs accordingly. These models are as same as static model if there are no case switches in the shift.

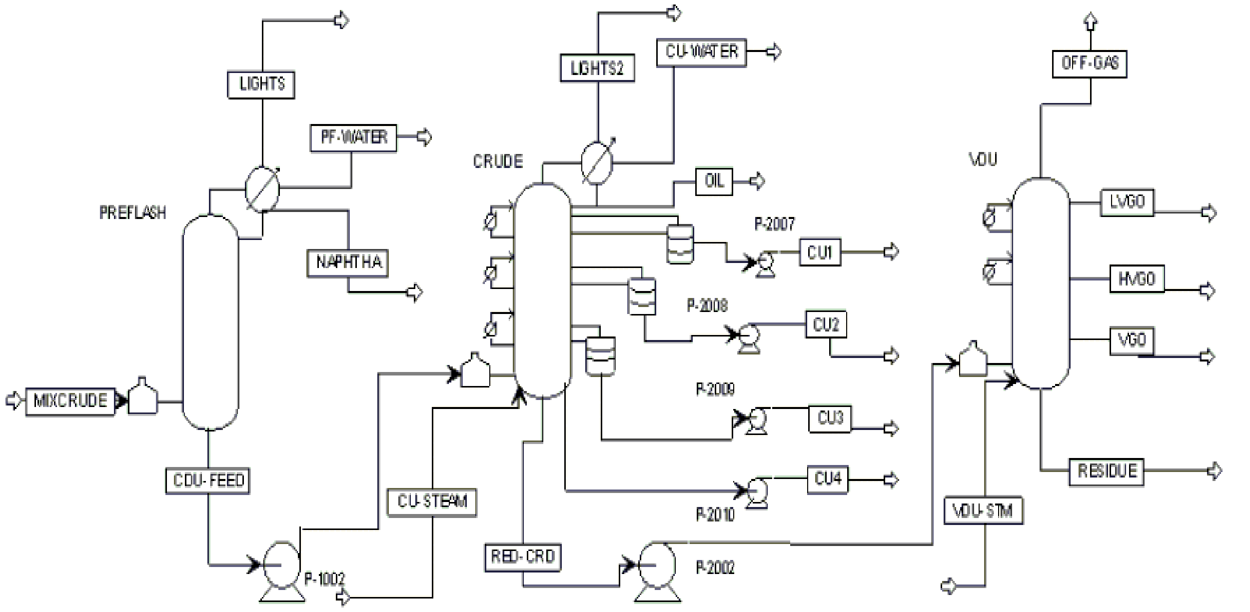

Fig.1. Flowsheet of crude oil distillation process (Redraw From [3])

Once the case switches between shifts, after the simulation of first shift is over, MSM will change to a new configuration by loading a different case from "Case list" before simulating the next shift. "Case list" is a list of familiar cases, which are identified by their unique codes. After a new case is loaded, the operation parameters of the new case are applied for the next simulation step. As illustrated in fig 3, while crude oil changes, CDU's case may switch from gasoline case to diesel oil case, which means ethylene and lubricating oil are substituted for heavy naphtha as the main products of CDU. Thus the operation settings for case in "Case list" have to include two parts of information: all productions' yield and sidetrack configuration. As unique code is used to identify the specific case, like 1 for gasoline case and 2 for diesel oil case, it can be treat as input parameter for all units. For example, input code series 111222221111 means there are 2 cases and 2 switches in 4 days. Without stopping the simulation process and changing operation settings manually, the FMCDS platform is able to simulate the process continuously.

3.2 Model aggregation and disaggregation

When the planning period expands from one shift to one day or even longer, data requirements is changing relatively. Model structure in this level, as illustrated in fig 4, is a combination of several possible shift models. Aggregate models have longer time interval and more complex structure than shift models, and the total input and output data can be generated by grouping the data after the simulation is over.

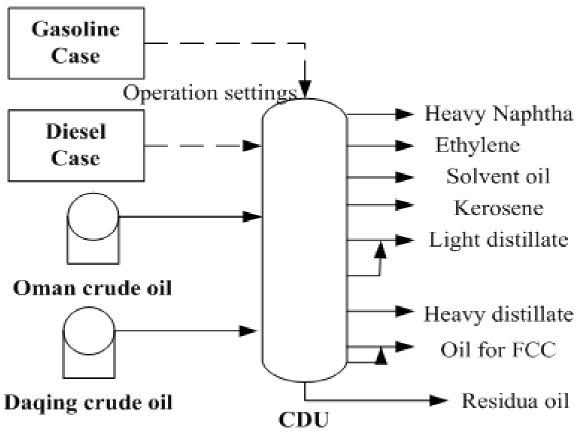

Fig .4. Aggregate model structure

Model aggregation was brought forward to support the research of data reconciliation and optimization. Data reconciliation is known as a data process technology, which improves the accuracy of process data by adjusting the measured values so that they satisfy the process constrains(Narasimhan and Jordache, 2000). Because of the economic or technical limitations, not all variables of the process are measured. And the reconciliation for the unmeasured variable is impossible if the unmeasured variable is unobservable. As there is input and output data records in the scheduling level, which describe the working conditions in an overall scale. For example, we can only find the record of daily residual oil production in the records because the demarcation of residual oil tank is a daily activity. So when KPI (Key Performance Indication) system needs to evaluate the performance of a given shift, it is better to use the daily records into reconciliation than treat residual oil production as unmeasured variables. So the problem of fulfilling the data requirement for both the scheduling and KPI system in one simulation period becomes a challenge for researchers, thus the idea of model aggre-

gation and desegregations is helpful.
By using the precise yields work out from plant data and dynamic models, shift model in the FMCDS platform provides all data that KPI system concerns about. After the shift model is built and validated, the FMCDS platform provides extensive facilities to aggregate and disaggregate these models into multi-period models. Simulation data from multi-period models are stored in the RDBMS as true values for comparison. And in order to generate emulational data close to reality, FMCDS platform add random error for all variables and gross error for a small proportion of variables, after eliminating some variables manually according to the sensor network in actual refinery to generate unmeasured variables, these data will be stored into RBDMS separately as simulation shift data. These simulation shift data can be aggregated flexibly as a data source for long-term refinery planning. By aggregating and disaggregating of these models, not only the emulational data but also the true values are provided for further researches.
As we described before, aggregated shift models with longer time interval usually have more than one case in simulation data. From the idea of model aggregation and discrete event simulation(Zeigler, 2004), the simulation strategy of case switches in shift activity is introduced hereinafter.

3.3 Case switches in shift

When the aggregated shift model is represented as a holistic model, there are usually more than one case switches in it. The input code series 111222221111 are recorded for model disaggregating and the execution time of each case can be obtained by counting the number of code appearance, 7 shifts (56 hours) for gasoline cases and 5 shifts (40 hours) for diesel oil case and the total productions can be calculated using (1).

$$P_i = \sum_{s=1}^{n} Cr_s \times Y_s^i \qquad (1)$$

Where Pi is the total output of product i, s<n is the code of case and there are n cases in the simulation data. After Cr_s (total crude oil input of case s) and Y_s^i (yield of product i of case s) together, we could get the total production of product *i* for case *s*.
There was a thought to deal with the case switches in shift problem and its main idea is to generate all possible cases by combine two basic cases into a new case in the "Case list". However, the problem of case explosion cannot be ignored. Because CDU has at least four basic cases, there are 6 new cases if the distribution of crude oil input for each case is certain. While the distribution of crude oil input is uncertain and basic cases are usually more than four, it is impossible to list all possible new cases. So a flexible simulation strategy, which employed the idea of discrete event simulation, is proposed to solve this problem by generating new cases in the simulation process.

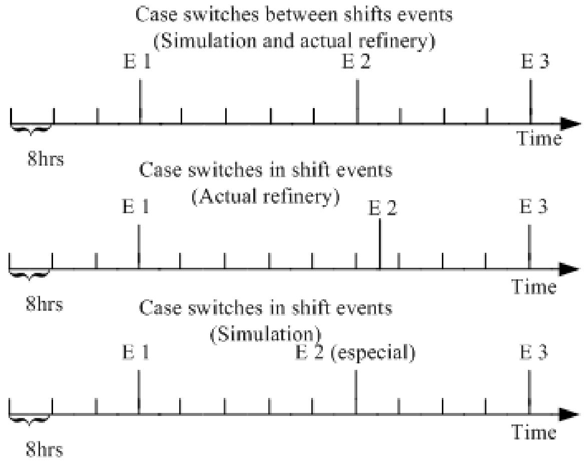

Fig.5. Representing case switches events

Discrete event processing is characterized by the ability to perceive the flow of sensory stimuli as discrete events, and to attend to both sequencing and timing of such events(Zeigler, 2004). The case switches activity in the FMCDS platform is a kind of sensory stimuli, which is crucial for the simulation results. As an example of case switches activity, the case switches between shifts activity has 8-integer multiples inter-event duration. But when the case switches in the shift, as shown in Fig 5, the inter-event duration is not that well-regulated, so we describes these events as especial events with unique codes in the simulation.

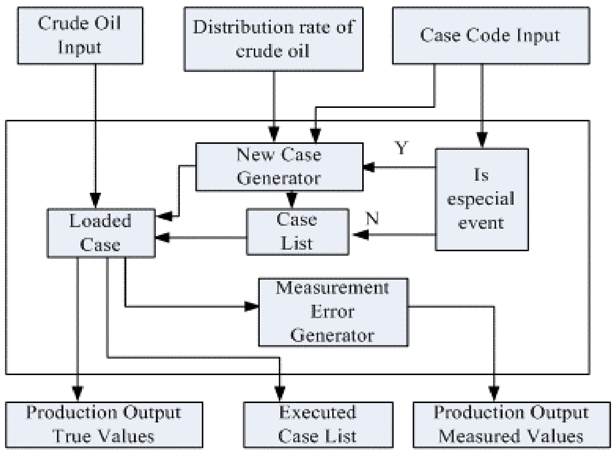

Fig.6. CDU simulation process in FMCDS platform

When the case switches event is an especial event, the FMCDS platform will ask for an additional input to continue the simulation process. This input is called the distribution rate of inputs (R), which represents the raw material cost of the case before switch. The production i's yield for the new cases is obtained by using (2).

$$Y_{es}^i = RY_{fs}^i + Y_{ls}^i - RY_{ls}^i \qquad (2)$$

Where Y^i_{es} is the yield of product i in a shift of new case *es*, Y^i_{fs} and Y^i_{ls} are the yields of production i of the latest case and switched case. After the yields are obtained, the new case is added into the "Case list" as basic cases, and the total productions still can be calculated using (1). To understand the FMCDS strategy well, Fig.6 shows the simulation process and all inputs and outputs for a CDU model.

Table 1. Comparison between the simulation data from FMCDS platform, simulation data from static models and plant data

Sidetrack	Daily recorded data of an industrial CDU			Daily statistic data of CDU from FMCDS platform[2]		Daily statistic data of CDU from static model [6][3]	
	Day 1 (No case switch)		Day 2[1] (With case switch)	Day 1	Day 2	Day 1	Day 2[4]
	Yield	Output (t/day)	Output (t/day)	Output (t/day)	Output (t/day)	Output (t/day)	Output (t/day)
Crude oil	100%	-8059	-6424	-8059	-6424	-8059	-6424
Fixed gas	0.07%	5.639	5.372	5.639	5.372	5.639	4.934
Reforming Stuff	6.1%	492	358.622	565.742	398.039	500.464	403.086
Naphtha	3.69%	297.746	156.345	155.539	78.808	0	0
Gasoline	0%	0	0	0	0	0	0
Solvent Oil	0%	0	1.531	0	1.531	0	1.531
Diesel Oil 1#	6.27%	505.333	313.774	540.759	337.009	598.784	378.451
Diesel Oil 2#	8.65%	696.726	633.096	726.922	653.147	725.31	643.77
Diesel Oil 3#	6.47%	521.614	469.433	539.953	470.978	678.568	521.218
Wax Oil	36.15%	2913.095	2334.742	2939.923	2344.131	2996.336	2396.316
Residue Oil	32.34%	2606.357	2122.574	2587.745	2106.049	2530.526	2043.161
Dirty Oil	0.07%	5.997	16.307	7.253	17.154	8.865	19.970
Loss	0.18%	14.49	11.782	14.49	11.782	14.49	11.56

[1] the distribution rate of crude oil is 4237:2187
[2] the yields is obtained from data mining
[3] the yields is obtained from standard cases
[4] this column of data is calculated from 2 batches of simulation data with different yields

Sidetrack configuration describes the movement of productions in the case. In the refinery, the physical structure of sidetracks is immovable while the production in the sidetrack is various. In order to simulate the movements in the actual refinery, the connections between units are unchangeable in the simulation, and the incidence matrix (**A**) between the productions and sidetracks for case s are defined as:

$$\mathbf{A}_{(s)} = \begin{array}{c} \begin{matrix} l_1 & l_2 & l_3 & l_4 & l_5 & l_6 & l_7 & .. & l_n \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & .. & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & .. & 0 \\ : & : & : & : & : & : & : & : & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{matrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ \vdots \\ p_m \end{matrix} \quad (3)$$

Where l is sidetrack code and p is the production code. The sidetrack outputs (**SP**) for all productions can be calculated by using (4) and (5) with constrain (6):

No case switches in the shift:

$$\mathbf{SP} = \mathbf{P} \times \mathbf{A}_{(s)} \quad (4)$$

Case switches in the shift:

$$\begin{aligned} \mathbf{SP} &= Cr \times R \times \mathbf{Y}_{(fs)} \times \mathbf{A}_{(fs)} \\ &+ Cr \times (1-R) \times \mathbf{Y}_{(ls)} \times \mathbf{A}_{(ls)} \end{aligned} \quad (5)$$

Constrain:

$$\sum_{j=1}^{n} \mathbf{A}_{ij} = \begin{cases} 1 & p_i \text{ is producted} \\ 0 & p_i \text{ isn't producted} \end{cases} \quad i = 1,2...m \quad (6)$$

Where Cr is input of crude oil in shift s, **P** is vector of production outputs, and **Y** is the vector of production yields.

In order to validate the models in the platform, plant data from a refinery in operation and simulation data from static models are obtained. Data in table 1 show that the simulate data from FMCDS platform confirm more closely to the plant data than regular static models. Moreover, the efficiency is prominently improved compare with static and dynamic models. So it can be concluded that the FMCDS platform is reliable and practical, which can be used as a useful tool for long-term refinery planning and data source for many other applications.

4. BASIC CASE MINING FOR FMCDS PLATFORM

As described before, "Case list" is a set of basic cases in the FMCDS platform and all new cases are the combination of two basic cases. So the basic cases are the foundation of FMCDS platform, thus the problem of obtain accurate basic cases is prompted as a preprocess of FMCDS. Cases are represented as vectors of production yields and sidetrack configurations in actual refineries. The standard yields for a case is determined in the laboratory and the sidetrack configuration changes because of the difference of productions. It is an easy way to get all cases in the laboratory and store them in the "Case list", however, there are lots of cases in actual process and most of them are similar. Moreover, the standard yields are not that precise because the process is highly nonlinear. To reduce the number of useful cases and improve the accuracy, technology of data mining is employed in the FMCDS platform.

Data mining is the extraction of implicit, previous unknown and potentially useful information from data(Witten and Frank, 2005). In the MES and PCS, the long-term process data of refinery is stored in the RTDB (Real-Time Database) and RDBMS for many applications. These databases, as a huge data warehouse, contain lots of useful information. Because the yields and sidetrack configurations are different according to the detail operation parameters of each shift, these values should be adjusted base on the standards. As the clustering method in the data mining technology is know an efficient way to divide the data into natural groups based on their distances, the process of case mining is described as follows:

--First, obtain a long period of process data divide by shift.

--Second, eliminate all shift data with case switches.

--Third, cluster the vector of yields by calculating yields' distances. (The method of clustering is presented in (Witten and Frank, 2005)

--Four, obtain the average yields and familiar sidetrack configurations from clustered case groups.

5. CONCLUSION AND FUTURE WORKS

Process simulation technology is more and more widely used in process industries. Various requirements like efficiency, accuracy and comprehensive are needed for a successful simulation platform. In order to fulfill more requirements from different users, lots of research fields like system identify, data integration, data mining has to be involved in the research. The FMCDS platform in virtual plant laboratory introduced in this paper is on the basis of dynamic simulation, static simulation and discrete event simulation, which take the demand of experiment teaching and comprehensive researches in to consideration. The FMCDS platform is and will play a positive role in the virtual plant laboratory and facilitate the relative research from theory to practice.

As an integrated platform of dynamic, static and discrete event simulation, the FMCDS platform has supported for many researches like scheduling, system identification and hierarchy data reconciliation, etc. However, it is just a new and immature platform, and it can be improved in many ways. First of all is the flexibility of modeling, the common units and some certain simulation strategy could be componentized by using some standards like CAPE-OPEN(Soares and Secchi, 2004) and ISA95 (ANSI/ISA, 2000) . And next is to make the simulation process intuitionistic, technology like Graphic User Interface and Virtual Reality will be researched for further improvement of FMCDS platform.

5. ACKNOWLEGEMENT

This work is supported by National Natural Science Foundation of China (60421002).

REFERENCES

ANSI/ISA-95.00.01-2000 (2000) Enterprise-Control System Integration - Part1: Models and Terminology.

Bassett, M. H., Dave, P., Doyle Iii, F. J., Kudva, G. K., Pekny, J. F., Reklaitis, G. V., Subrahmanyam, S., Miller, D. L., and Zentner, M. G. (1996). Perspectives on Model Based Integration of Process Operations. *Computers and Chemical Engineering*, **20**, 821-844.

Fang, H. F., Feng, Y. P., G. Rong. (2006). Simulation Platform in the Virtual Factory Laboratory System. *12th IFAC Symposium on Information Control Problems In Manufacturing*, *Saint-Etienne*, *France*.

Feng, Y. P. and G. Rong. (2005). Virtual Plant Laboratory System of Process Industries for Education. 1*6th Triennial World Congress of the International Federation of Automatic Control*, *Prague*,*Czech Republic*, **32**, 4-8.

Forrester, J. (1958). *Industrial Dynamics*. Portland, OR, USA., Productivity Press.

Li, W., C. W. Hui and Li, A. (2005). Integrating CDU, FCC and product blending models into refinery planning. *Computers and Chemical Engineering*, **29**, 2010-2028.

Narasimhan, S. and C. Jordache (2000). *Data Reconciliation & Gross Error Detection*. Houston, Gulf Publishing Company.

Pei, R. L. and G. Rong. (2005). Flowsheet Simulation Platform of Intelligent Plant in Oil Refinery. *Control and Instruments in Chemical Industry*. **32**. 43-46 (in chinese).

Rabelo, L., M. Helal, et al. (2003). A Hybrid Approach to Manufacturing Enterprise Simulation. Winter Simulation Conference Proceedings.

Soares, R. P. and A. R. Secchi. (2004). Modifications, Simplifications, and Efficiency Tests for The CAPE-OPEN Numerical Open Interfaces. *Computers and Chemical Engineering*. **28**. 1611-1621.

Witten, I. H. and E. Frank. (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. Elsevier Pte Ltd.

Zeigler, B. P. (2004). Discrete Event Abstraction:An Emerging Paradigm For Modeling Complex Adaptive Systems. http://www.acims.arizona.edu/PUBLICATIONS/PDF/HolandBook.pdf.