Autonomous Control of Space Reactor Systems

Final Report

Report No. DE-FG07-04ID14589/UTNE-06
NEER Grant Number: DE-FG07-04ID14589

Belle R. Upadhyaya
(Principal Investigator)
K. Zhao
S.R.P. Perillo
Xiaojia Xu
M.G. Na

The University of Tennessee
Nuclear Engineering Department
209 Pasqua Engineering Building
Knoxville, TN 37996-2300
E-mail: bupadhya@utk.edu

DOE Program Manager for NEER
Idaho Operations Office
Nancy A. Elizondo

November 2007
EXECUTIVE SUMMARY

Autonomous and semi-autonomous control is a key element of space reactor design in order to meet the mission requirements of safety, reliability, survivability, and life expectancy. In terrestrial nuclear power plants, human operators are available to perform intelligent control functions that are necessary for both normal and abnormal operational conditions. However, for a space mission with uncertain environment, rare events, and communication delays, all the control functions must be achieved through a sophisticated control system with very limited human intervention from the earth. It should be noted that autonomous control strategies are also of importance in reactor systems that are land-based, with remote deployment capabilities.

The objective of this research project under the U.S. Department of Energy NEER grant is to develop an integrated autonomous control system for space fission reactors, incorporating control mode selection, self-tuning, automated learning, on-line fault monitoring and failure anticipation, fault-tolerant, and supervisory control. The development and implementation of these technologies are illustrated with application to a liquid-metal cooled reactor with static power conversion. The results of this research are generic so that the technology can be easily adapted to different space power systems and to next generation reactors. The project involves completion of the following three major tasks during Phases 1, 2, and 3:

- Development of space reactor simulation models for control design, using static power conversion. This task will integrate the reactor system dynamics and the power conversion module, and study the performance of the system under various transients.

- Development of autonomous control strategies for one of the reactor systems. This will address the functional requirements of space reactor control and the implementation using advanced control methods, specifically the model-predictive controller (MPC).

- Integration of control, space reactor monitoring system, and interfacing measurement and reactor system modules for high reliability operation. Demonstration of the method using a laboratory flow control loop.

Some of the highlights of the project include the following accomplishments:

- Development and testing of the SP-100 reactor system dynamics and the power conversion module (thermo-electric conversion) nodal model as an integrated system.
• Development and application of the Model Predictive Control (MPC) algorithm to the SP-100 system. The SP-100 reactor simulation model has been used in this task. The MPC algorithm has been evaluated against traditional controllers, for various transient conditions.

• Development of a fault detection and isolation module for monitoring incipient faults in various field devices. This method uses the Principal Component Analysis (PCA) and data-driven models for fault monitoring.

• Design and completion of a laboratory multivariate water level control test loop, for evaluating the various control strategies and the fault-tolerant control method.

• On-line demonstration of the classical proportional-integral controller and the MPC using the laboratory control test loop.

• Development of an autonomous control framework, with control mode reconfiguration and hierarchical control strategy.

• Presentation of seven papers in national and international conferences, and publication of four manuscripts in refereed journals.

The project Final Report describes the development of the SP-100 space reactor simulation model in MATLAB-SIMULINK, development and application of the MPC controller and the fault-tolerant control strategy, development of a laboratory multivariate control test loop, data acquisition under classical control action, and on-line implementation of the MPC controller in the control loop. A detailed bibliographic reference is given. A listing of the computer codes developed under this project is also provided.

**DISCLAIMER**

Any opinions, findings, and conclusions or recommendations expressed in this report are those of the authors, and do not necessarily reflect the views of the United States Department of Energy.
ACKNOWLEDGMENTS

This research was performed under a grant from the U.S. Department of Energy NEER Program (DE-FG07-04ID14589) with the University of Tennessee, Knoxville. We want to acknowledge the assistance provided by the NEER Program Manager and the Idaho Operations Office. We want to thank the Nuclear Science and Technology Division of ORNL for discussions related to autonomous control and space reactor systems.

The assistance provided by the University of Tennessee Nuclear Engineering Department support staff, in the fabrication and testing of the laboratory flow control loop, is greatly appreciated.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>1</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>3</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>7</td>
</tr>
<tr>
<td>1.1. Background and Project Objectives</td>
<td>8</td>
</tr>
<tr>
<td>1.2. Autonomous Control</td>
<td>8</td>
</tr>
<tr>
<td>1.3. Summary of Significant Accomplishments</td>
<td>10</td>
</tr>
<tr>
<td>2. FUNCTIONAL ANALYSIS OF SPACE REACTOR CONTROL</td>
<td>13</td>
</tr>
<tr>
<td>3. DEVELOPMENT OF A NODAL MODEL OF THE SP-100 REACTOR</td>
<td>19</td>
</tr>
<tr>
<td>AND POWER CONVERSION SYSTEM</td>
<td></td>
</tr>
<tr>
<td>3.1. Description of the SP-100 System</td>
<td>19</td>
</tr>
<tr>
<td>3.2. Reactor Core Model</td>
<td>21</td>
</tr>
<tr>
<td>3.3. Primary Coolant Heat Exchanger Model</td>
<td>31</td>
</tr>
<tr>
<td>3.4. Thermal Electric Model</td>
<td>38</td>
</tr>
<tr>
<td>3.5. Review of the SP-100 Electromagnetic Pump</td>
<td>52</td>
</tr>
<tr>
<td>3.6. Integration of the SP-100 System Model</td>
<td>55</td>
</tr>
<tr>
<td>3.7. Summary of Model Equations for Control</td>
<td>58</td>
</tr>
<tr>
<td>3.8. Coordinated Control of the SP-100 Space Reactor</td>
<td>66</td>
</tr>
<tr>
<td>4. MODEL PREDICTIVE CONTROL (MPC)</td>
<td>69</td>
</tr>
<tr>
<td>4.1. Introduction</td>
<td>69</td>
</tr>
<tr>
<td>4.2. The Model Predictive Control (MPC) Method</td>
<td>70</td>
</tr>
<tr>
<td>4.3. Optimization of the Objective Function by a Genetic Algorithm (GA)</td>
<td>75</td>
</tr>
<tr>
<td>4.4. Application of the MPC Method to the SP-100 Space Reactor</td>
<td>79</td>
</tr>
<tr>
<td>5. FAULT DETECTION AND ISOLATION (FDI) OF SENSORS AND FIELD DEVICES</td>
<td>87</td>
</tr>
<tr>
<td>5.1. Introduction</td>
<td>87</td>
</tr>
<tr>
<td>5.2. Principal Component Analysis (PCA) for Measurement Modeling</td>
<td>88</td>
</tr>
<tr>
<td>5.3. Selection of Number of Principal Components</td>
<td>91</td>
</tr>
<tr>
<td>5.4. Fault Detection of Sensors and Field Devices</td>
<td>91</td>
</tr>
<tr>
<td>5.5. Fault Isolation of Sensors and Field Devices</td>
<td>92</td>
</tr>
<tr>
<td>5.6. Application of PCA to the SP-100 Reactor System</td>
<td>94</td>
</tr>
</tbody>
</table>
11. CONCLUDING REMARKS AND FUTURE WORK 196
11.1. Summary and Concluding Remarks 196
11.2. Recommendations for Future Research 197

REFERENCES 198

BIBLIOGRAPHY 200

APPENDIX A 221
Fault Tolerant Control of the SP-100 Space Reactor (Manuscript)

APPENDIX B 248
Orifice Meter Calibration Procedure and Results

APPENDIX C 255
M-File of System Parameters of SP-100 Reactor

APPENDIX D 259
Fault Detection and Isolation (FDI) Algorithm for SP-100 System
1. INTRODUCTION

1.1. Background and Project Objectives

Autonomous or semi-autonomous control is a key element of space reactor design to meet the mission requirements in terms of safety, reliability, and life expectancy. In land-based nuclear power plants, human operators are available to perform intelligent control functions necessary for both normal operation and during accident situations. However, for a space mission with uncertain environment, rare events, and communication delays, all the control functions must be achieved through a sophisticated control system with very limited degrees of human intervention from earth. Autonomous control strategies are also important in reactor systems that are land-based.

The objective of the three-phase research is to develop an integrated autonomous control system for space fission reactors, incorporating control mode selection, self-tuning, on-line fault monitoring and failure anticipation, reconfigurable and hierarchical control. The development and implementation of these technologies are illustrated with application to a liquid-metal cooled reactor with a thermo-electric power conversion system. It is anticipated that the results of this research would be generic so that the technology could be easily adapted to different space power systems.

The following three areas have been developed as part of an autonomous control system:

- Development of space reactor simulation models for transient analysis and control design, using a thermo-electric power conversion system. This task integrates the reactor system dynamics and the power conversion module, and studies the performance of the system under various system perturbations.

- Development of an autonomous control strategy that incorporates reactor system monitoring, fault detection and isolation of sensors and field devices, model-based controllers, and communication within a hierarchical control framework. This will address the functional requirements of space reactor control and the implementation using advanced control methods.

- Integration of control, space reactor monitoring and diagnostics, and decision-making modules. This includes the interfacing of measurements and reactor system modules for high reliability operation. On-line implementation and demonstration of the control strategies in an experimental flow control loop that is developed as part of the project. This task was not one of the original tasks defined for the project.
The simulation model of the space reactor follows the previously developed model of the SP-100 reactor [1, 2]. This model has been further enhanced using the MATLAB-SIMULINK platform [3]. Fault detection and isolation (FDI) methods are based on characterizing the relationship among a set of state variables using model-based techniques and then tracking the patterns of deviations between the actual measurements and their predicted values. A model-predictive controller has been developed [4, 5] and demonstrated to perform effectively in tracking the system demand. The autonomous control strategy uses a hierarchical framework to communicate the functional nature of the devices and the distributed controllers to a supervisor. The control action must be capable of changing the control mode by proper reconfiguration of the available information.

1.2. Autonomous Control

The basic requirement of space reactor control system is to achieve a stable control at different operation modes and enable a stable transition between these operation modes. From the viewpoint of a space mission, the major operation modes needed for space reactor control include normal operation at different power levels, hot standby mode for system testing, scheduled mission task, or transition to long-term shutdown, and cold shutdown mode for maintenance purpose. Correspondingly, the two major transitional modes are start-up and shutdown. During the transitional modes, the reactor systems and equipment must be closely monitored and controlled to assure compliance with safety requirements.

To meet the mission requirements in terms of safety, reliability, survivability, and life expectancy, the space reactor control system should incorporate advanced features such as economic performance monitoring, fault tolerance, control loop performance monitoring, and safety monitoring. Figure 1.1 shows the advanced control features necessary for the autonomous control of space nuclear reactor systems.

To support a mission with extended period of operation, the space reactor control system needs to have the capability of performance monitoring and optimization. For instance, in deep-space missions, the reactor system is expected to have continuous, remote, and unattended operation for up to fifteen years. During such an extended period, many thermal and electric components may experience significant level of degradation. In order to save mission costs, the operation parameters should be adjusted to optimize operational performance according to current operating conditions and implemented by a supervisory control mechanism.
To realize unattended operation and address communication delay, fault tolerant control should be an additional important feature of the space reactor control system. If a fault occurs in a control system, a fault diagnosis system should be able to perform reliable and timely fault detection and isolation before serious consequences happen. This fault diagnostic information may be further used for both operation planning and controller reconfiguration. The design objective of this fault tolerant control system is to enable a stable control under anticipated conditions with such faults as sensors, actuators, and reactor components.

Because unknown environment and operating conditions may be encountered for a deep space mission, the space reactor control system should also have the capability of monitoring the control loop performances themselves. If there is significant degradation in controller performance, the operating controller may need to be retuned or an alternative control algorithm may be implemented. The integration of control performance monitoring and automatic tuning feature provides the space reactor control system with adaptive capability.

In order to ensure operational safety of space reactor, the space reactor control system needs to continuously monitor safety critical parameters to protect reactor components and systems. If safety critical variables have triggered alarm signals, these alarm signals must be processed, prioritized, and converted into a concise representation of the reactor operating status before sent back to earth. In addition, a computerized emergency operation procedure should be in place to support automatic decision-making and control during accident conditions.

The above control functions constitute a highly autonomous control system, which assumes the responsibilities for normal control action, abnormal event response, fault tolerance, and provide an interface with operators on earth for high-level decision-making. This autonomous control system can be implemented by a hierarchical paradigm, which is shown in Figure 1.1.

In this hierarchical control paradigm, the bottom level has a direct interface with the space reactor systems and is responsible for executing fault detection and isolation, fault tolerant control, and abnormal event control. The middle level is mainly to coordinate between the plant manager level and the execution level. The coordination includes the determination of what is the performance of current control structure and whether it is necessary to tune the controllers or switch the control laws. Significant uncertainties can be dealt with by designing adaptive control laws for unanticipated situations and producing control sequences based on current operating conditions. The top level is a plant manager, which manages plant level performance...
monitoring, plant condition evaluation, capability assessments, and task planning. In addition, the plant manager can send out space reactor operational data to and receive commands from the operators on earth.

Figure 1.1. Schematic of the autonomous/hierarchical controller functional architecture.

References [7-11] describe model-predictive control and fault detection and isolation in dynamic systems.

1.3. Summary of Significant Accomplishments

Several control algorithms, reactor models, and experimental implementation of control strategies have been developed during the project. The following is a summary of significant results and the dissemination of research and development.

- Development and testing of the SP-100 reactor system dynamics and the power conversion module (thermo-electric conversion) nodal model as an integrated system.

- Development and application of the Model Predictive Control (MPC) algorithm to the SP-100 system. The SP-100 reactor simulation model has been used in this task. The MPC algorithm has been evaluated against traditional controllers, for various transient conditions.
• Development of a fault detection and isolation module for monitoring incipient faults in various field devices. This method uses the Principal Component Analysis (PCA) and data-driven models for fault monitoring.

• Design and completion of a laboratory multivariate water level control test loop, for evaluating the various control strategies and the fault-tolerant control method.

• Development of a multivariate flow control loop and on-line demonstration of the classical proportional-integral controller and the model-predictive controller (MPC) using the laboratory test loop.

• Development of an autonomous control framework, with control mode reconfiguration and hierarchical control strategy.

• Presentation of seven papers in national and international conferences, and publication of four manuscripts in refereed journals. These are listed below:


The report is organized in 11 sections that describe the research and development results accomplished under this project. The report also provides a detailed list of bibliographic references. A copy of the recently submitted manuscript to IEEE Transactions on Control System Technology is appended to the report.
2. FUNCTIONAL ANALYSIS OF SPACE REACTOR CONTROL

The basic requirement of a space reactor control system is to achieve a stable control at different operation modes and enable a stable transition between these operation modes. From the viewpoint of a space mission, the major operation modes needed for space reactor control include normal operation at different power levels, hot standby mode for system testing, scheduled mission task, or transition to long-term shutdown, and cold shutdown mode for maintenance purpose. Correspondingly, the two major transitional modes are start-up and shutdown. During the transitional modes, the reactor systems and equipment must be closely monitored and controlled to assure compliance with safety requirements.

To meet the mission requirements in terms of safety, reliability, survivability, and life expectancy, the space reactor control system should incorporate advanced features such as economic performance monitoring, fault tolerance, control loop performance monitoring, and safety monitoring. Figure 2.1 shows the advanced control features necessary for the autonomous control of space nuclear reactor systems.

To support a mission with extended period of operation, the space reactor control system needs to have the capability of performance monitoring and optimization. For instance, in deep-space missions, the reactor system is expected to have continuous, remote, and unattended operation for up to fifteen years. During such an extended period, many thermal and electric components may experience significant level of degradation. In order to save mission costs, the operation parameters should be adjusted to optimize operational performance according to current operating conditions and implemented by a supervisory control mechanism.

In order to realize unattended operation and address communication delay, fault tolerant control should be an additional important feature of the space reactor control system. If a fault occurs in a control system, a fault diagnosis system should be able to perform reliable and timely fault detection and isolation before serious consequences happen. This fault diagnostic information may be further used for both operation planning and controller reconfiguration. The design objective of this fault tolerant control system is to enable a stable control under anticipated conditions with such faults as sensors, actuators, and reactor components.

Because unknown environment and operating conditions may be encountered for a deep space mission, the space reactor control system should also have the capability of monitoring the control loop performances themselves. If there is significant degradation in controller
performance, the operating controller may need to be retuned or an alternative control algorithm may be implemented. The integration of control performance monitoring and automatic tuning feature provides the space reactor control system with adaptive capability.

In order to ensure operational safety of space reactor, the space reactor control system needs to continuously monitor safety critical parameters to protect reactor components and systems. If safety critical variables have triggered alarm signals, these alarm signals must be processed, prioritized, and converted into a concise representation of the reactor operating status before sent back to earth. In addition, a computerized emergency operation procedure should be in place to support automatic decision-making and control during accident conditions.

The above control functions constitute a highly autonomous control system, which assumes the responsibilities for normal control action, abnormal even response, fault tolerance, and provides an interface with operators on earth for high-level decision-making. This autonomous control system can be implemented by a hierarchical paradigm, which is shown in Figure 2.2.

In this hierarchical control paradigm, the bottom level has a direct interface with the space reactor systems and is responsible for executing fault detection and isolation, fault tolerant
control, and abnormal event control. The middle level is mainly to coordinate between the plant manager level and the execution level. The coordination includes the determination of what is the performance of current control structure and whether it is necessary to tune the controllers or switch the control laws. Significant uncertainties can be dealt with by designing adaptive control laws for unanticipated situations and producing control sequences based on current operating conditions. The top level is a plant manager, which manages plant level performance monitoring, plant condition evaluation, capability assessments, and task planning. In addition, the plant manager can send out space reactor operational data to and receive commands from the operators on earth.

![Diagram of autonomous/hierarchical controller functional architecture](image)

Figure 2.2. Schematic of the autonomous/hierarchical controller functional architecture.

Although the hierarchical control paradigm defined in Figure 2.2 provides a philosophical concept of autonomous control, it is still necessary to define an appropriate technology for implementation. Model Predictive Control (MPC) is such an available technology. In this hierarchical structure, a Model Predictive Controller performs dynamic control to move the space reactor system from one constrained steady state to another optimal state demanded by the plant manager. The dynamic model can be derived either from first-principles or from the plant test data. The control coordinator needs to guide the MPC controllers on which models should be used and monitor performance of the current MPC control strategy.
Figure 2.3 shows the schematic of a model predictive controller and the interaction among model development, fault diagnosis, and fault tolerant control. Multiple steps are defined to fulfill the objectives of fault diagnosis, fault tolerant control, and alarm event control in a MPC controller. The first step is to read current values of process inputs (including manipulated variables and disturbance variables) and process outputs as well as the functional status of the sensors and actuators. The second step is to perform an adaptive state estimate using mature Kalman filter technique. The objective of incorporating a step of state estimation preceding controller design is to accommodate model uncertainty and unknown disturbance through a mechanism of output feedback. It is during this step that MPC has been endowed with adaptive features. After process state is estimated, at the third step, MPC controller needs to determine which manipulated variables should be manipulated and which controller variables should be controlled based on the functional status of the sensors and actuators. The fourth step is to perform an online local steady state optimization to drive the steady state inputs and outputs toward the targets commanded by the plant manager. This step is necessary because the control problem may be redefined due to disturbances or operating condition changes. The last step is to compute a set of adjustments needed for the manipulated variables to the desired steady states without violating constraints on process input and output.
The kernel of an MPC controller is the process model updated on-line through state estimation. The same model can be used for both fault diagnosis of field devices and model predictive control. During space reactor power operation, the model can be updated every minute and a dynamic optimization based MPC can be implemented on the same time scale. In the meantime, the model based fault diagnosis algorithm should be implemented every second based on the available model. If a fault is detected and isolated, a new control design will be immediately triggered. Because dynamic control and fault diagnosis are organically combined, the implementation scheme of MPC defined in Figure 2.3 has the capability of fault tolerance.

In this project, a detailed dynamic model has been developed under the MATLAB/Simulink environment to simulate a liquid-metal cooled fast reactor system. The basic algorithm of Model Predictive Control is extended to include the above defined autonomous control features such as on-line fault monitoring and failure anticipation, controller self-tuning, automated learning, and reconfigurable control. Data-based fault diagnosis methods
have been used to detect and isolate sensor and actuator faults. The disturbance rejection capability of model predictive control has been studied and demonstrated for achieving fault tolerant control.
3. DEVELOPMENT OF A NODAL MODEL OF THE SP-100 REACTOR AND POWER CONVERSION SYSTEM

3.1. Description of the SP-100 System

A lumped parameter simulation model is developed for the thermo-electric (TE) SP-100 system in this research [2] based on the early work of Seo [6]. The simulation model has a modular architecture based on the configuration of reactor components. The individual modules considered include a model of reactor control mechanism, a neutron kinetics model, a reactor core heat transfer model, a primary heat exchanger model, and a thermoelectric conversion model. In this section, each component model is presented in terms of its mathematical formulation of governing equations, the parameters used, and the simulation results. After all the component models are developed, the integrated SP-100 system model is assembled through an iterative algorithm. The developed model involves both nonlinear ordinary differential equations and partial differential equations. The code development is under the MATLAB environment without using Simulink. In a separate application, the MATLAB/Simulink software is used for modeling the SP-100 reactor.

SP-100 is a fast spectrum, Lithium-cooled fuel pin reactor coupled with thermoelectric converters (TE) with the waste heat removed through a heat pipe distribution system and space radiators. Because the fast spectrum lithium-cooled reactor coupled with thermoelectric (TE) converters was finally selected for engineering study in 1980s, this TE based SP-100 system has resulted in the most sophisticated engineering development and a well documentation in public literature. For this reason, TE SP-100 system is chosen to study the autonomous control of space reactor systems in this research. This also facilitates the collection of reactor design parameters.

The TE SP-100 system [1] is made up of a nuclear reactor, a heat transport system with associated pumps, and a thermoelectric device to convert heat to electricity, and a radiator configuration system to reject waste heat into space. Figure 3.1 shows a schematic of one loop of the reactor system. The liquid metal fast reactor was chosen for the TE SP-100 design because of its lightweight and small size. The reactor core is composed of small disks of highly enriched (93%) uranium nitride fuel contained in sealed tubes. The use of high enrichment uranium allows a reactor design of long core lifetime and high power density. The uranium nitride fuel was selected due to the mass savings, chemical stability, and physical robustness [1].
The core reflectors are made of beryllium oxide. Some movable safety rods are in place during launch and are withdrawn for operation when the reactor is prepared for start-up in orbit [6].

The heat generated by the reactor is transferred by liquid lithium pumped by sealed electromagnetic (EM) pumps with no moving parts. Lithium is used as the heat carrier because of its good thermal conductivity and low vapor pressure up to 1,350 K. The operation of the EM pump is based on the principle that a pumping force is produced when lithium flows through a plane that is spanned by an electric current vector and a magnetic flux density vector, which are normal to each other. The EM pump is self-actuating since the magnetic field is generated by a permanent magnet, and the electric current is generated by the TE cells.

The interface between the primary heat transport system and the energy conversion system is the primary heat exchanger. The EM pumps pump the primary coolant from the reactor core to the hot header of the primary heat exchanger. In the primary heat exchanger, the coolant flow is distributed into the individual flow channels of the heat exchanger. About 480 thermoelectric cells are mounted on the surface of each channel. After the heat is removed from the channel surface through the thermoelectric cells, the cold fluid converges in the cold header and flows out of the heat exchanger. In the TE SP-100 design, three reactor loops are equipped to transport the heat to the thermoelectric devices. The thermoelectric cells are installed on 12 primary heat exchangers. For each primary heat exchanger, 30 channels are divided to enhance the heat transfer from the primary fluid to the thermoelectric cells (Seo [6]).

The energy conversion system is based on the thermoelectric conversion mechanism. A temperature drop of about 500 K is maintained across the thermoelectric elements by the cooling effect of a second liquid lithium loop that transfers the waste heat from the converter to a heat-pipe radiator. The thermoelectric elements are semiconductors made of silicon/germanium-gallium phosphate. The heat-pipe radiator rejects the waste heat into space. One of the radiator design concepts is to utilize hundreds of high temperature heat pipes attached in parallel rows to reject the waste heat to the space environment.
In addition, the TE SP-100 design has an independent auxiliary cooling system that can generate about 1,000 Watts of electrical power, which is used to cool the reactor and provide power for basic controls even when loss of lithium accident occurs. In the TE SP-100 design, conventional spacecraft techniques are used to distribute, convert, and regulate the electric power to meet the user requirements (Demuth, 2003).

3.2. Reactor Core Model

3.2.1. Neutron Kinetics Model

The point reactor kinetics model with six delayed neutron groups is used to describe the dependence of nuclear reactor power on the reactivity change. This is given by the following equations (Seo, 1986):

\[
\begin{align*}
\frac{dP_n(t)}{dt} &= \frac{\rho(t) - \beta_{int}}{\Lambda} P_n(t) + \sum_{i=1}^{6} \lambda_i C_i(t) \\
\frac{dC_i(t)}{dt} &= \beta_i \frac{P_n(t)}{\Lambda} - \lambda_i C_i(t)
\end{align*}
\]

where
\( P_n(t) \) = instantaneous nuclear power (W).

\( \beta_{1-6} \) = fractions of the 6 groups of delayed neutrons.

\( \beta_{\text{tot}} \) = total fraction of delayed neutrons.

\( \lambda_{1-6} \) = decay constants of the precursors of the 6 groups of delayed neutrons.

\( \Lambda \) = average neutron generation time (1.55E-7 s).

\( \rho(t) \) = total reactivity at time \( t \).

\[ \rho(t) = \rho_{CD}(t) + \rho_{FB}(t) \]

\( \rho_{CD} \) = external reactivity inserted by the control drum.

\( \rho_{FB} \) = feedback reactivity.

The neutronic data used to solve the reactor kinetics equation are listed in Table 3.1 (Berkan, Upadhyaya, Kisner, 1990).

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precursor decay constant(1/s)</td>
<td>0.0127</td>
<td>0.0317</td>
<td>0.115</td>
<td>0.311</td>
<td>1.4</td>
<td>3.78</td>
<td>0.08259</td>
</tr>
<tr>
<td>Delayed neutron fraction</td>
<td>2.52E-4</td>
<td>1.47E-3</td>
<td>1.344E-3</td>
<td>2.941E-3</td>
<td>1.024E-3</td>
<td>2.37E-4</td>
<td>7.178E-3</td>
</tr>
<tr>
<td>Precursor mean lifetime(s)</td>
<td>78.74</td>
<td>31.5457</td>
<td>8.6957</td>
<td>3.2154</td>
<td>0.7143</td>
<td>0.2643</td>
<td>12.1077</td>
</tr>
</tbody>
</table>

### 3.2.2. Feedback Reactivity

The core model takes into account three mechanisms of reactivity feedback. The reactor fuel Doppler feedback reactivity, \( \rho_D \), represents the competitive effects of the absorption of fertile material such as U-238 and the fission of fissile material such as U-235 when fuel temperature changes. The reactor core expansion feedback reactivity, \( \rho_E \), represents the reactivity effect due to the expansion of reactor core materials including the reactor fuel, the fuel
cladding, and the reflector. The reactor core coolant expansion coefficient, $\rho_c$, takes into account the reactivity effect of neutron leakage when the temperature of reactor core coolant changes.

The total reactivity feedback is calculated as follows (Seo, 1986):

$$\rho_{FB}(t) = \rho_D(t) + \rho_E(t) + \rho_c(t)$$

where

$$\rho_D(t) = \alpha_D \ln \frac{T_f(t)}{T_f(0)}$$

$$\rho_E(t) = \alpha_f (T_f(t) - T_f(0)) + \alpha_{clad} (T_{clad}(t) - T_{clad}(0)) + \alpha_R (T_R(t) - T_R(0))$$

$$\rho_c(t) = \alpha_c (T_c(t) - T_c(0))$$

$\alpha_D$ = fuel Doppler temperature coefficient (+2.4E-7)

$\alpha_f$ = fuel expansion feedback coefficient (-1.22E-5/K).

$\alpha_{clad}$ = fuel cladding expansion feedback coefficient (1.20E-7/K).

$\alpha_R$ = reactor core reflector expansion feedback coefficient (-4.5E-7/K).

$\alpha_c$ = reactor core coolant expansion feedback coefficient (-3.25E-6/K).

$T_f$ = fuel temperature (K).

$T_{clad}$ = fuel cladding temperature (K).

$T_R$ = reactor core reflector temperature (K).

$T_c$ = reactor core coolant temperature (K).

### 3.2.3. Reactivity Control Mechanism

The external reactivity control in the developed model uses the mechanism of the stepper motor control drum system (Shtessel, 1998). The control drum shaft can be rotated from 0 to 180 degrees. The control voltage is transformed into a set of 27 V rectangular impulses with the frequency varying from 0 to 1.33 Hz. These impulses are then converted into discrete
movements of the shaft connected to the control drums. The position of the shaft of the stepper motor is given as follows (Shtessel, 1998):

\[
\frac{d^2 \theta}{dt^2} + 1.01 \frac{d\theta}{dt} = 0.0525 u
\]

(3.3)

where

\( u = \) the control voltage (V).

\( \theta = \) the shaft angle of the stepper motor.

The external reactivity inserted by the control drum can be fitted as a function of the shaft angle of the stepper motor, and is given by (Shtessel, 1998):

\[
\rho_{CD} = 6.89 \times 10^{-13} \theta^5 - 2.33 \times 10^{-10} \theta^4 + 3.28 \times 10^{-9} \theta^3 + 4.57 \times 10^{-6} \theta^2 - 5.88 \times 10^{-5} \theta
\]

(3.4)

\( \rho_{CD} = \) integral reactivity worth due to the control drum corresponding to the positioned angle.

3.2.4. Core Heat Transfer Model

A simplified reactor core heat transfer model is developed to calculate the fuel temperature, the cladding temperature, and the average core coolant temperature. The fuel temperature \( T_f \), the cladding temperature \( T_{clad} \), and the average core coolant temperature \( T_c \) are described by the following ordinary differential equations (Seo, 1986):

\[
\frac{dT_f(t)}{dt} = \left[ P_n - (T_f - T_{clad}) \cdot (UA)_f \right] / C_f
\]

(3.5a)

\[
\frac{dT_{clad}(t)}{dt} = \left[ (T_f - T_{clad}) \cdot (UA)_f - (T_{clad} - T_c) \cdot (UA)_{clad} \right] / C_{clad}
\]

(3.5b)

\[
\frac{dT_c(t)}{dt} = \left[ (T_{clad} - T_c) \cdot (UA)_{clad} - m_c C_p (T_{ex} - T_{in}) \right] / C_c
\]

(3.5c)
where

\[ T_f = \text{average fuel temperature (1376.0 K)}. \]

\[ T_{in} = \text{core inlet coolant temperature (1254.0 K)}. \]

\[ T_{clad} = \text{average cladding temperature (1288.0 K)}. \]

\[ \dot{m} = \text{reactor core coolant flow rate (14.1 kg/s)}. \]

\[ C_c = \text{heat capacity of the coolant}. \]

\[ C_f = \text{heat capacity of the fuel (1.465 E2 J/kg/K *109.0kg)}. \]

\[ C_{clad} = \text{heat capacity of the fuel cladding (0.276 J/kg/K*25.5 kg)}. \]

\[ (UA)_f = \text{effective heat transfer coefficient between the fuel and the cladding (2.2728E4 J/K)}. \]

\[ (UA)_{clad} = \text{effective heat transfer coefficient between the fuel cladding and the reactor core coolant (9.5757 E4 J/K)}. \]

A more accurate representation may be developed using the Mann’s model for fuel-to-coolant heat transfer. If it is assumed that the average core coolant temperature is an arithmetic average of the core inlet and core outlet temperature, the core exit temperature can be calculated as follows:

\[ T_{ex} = 2 T_c - T_{in} \]

3.2.5. Reactor Core Hydraulic Model

In SP-100 design, the reactor coolant enters the reactor vessel and flows up through the annular space between the reactor vessel and the core baffle until it reaches the upper plenum. At the upper plenum, the reactor coolant reverses and flows downward through the triangular fuel channels into the lower plenum. The pressure drop in the reactor core is a sum of the pressure drop in the annular space, the pressure drop in the triangular fuel channel, the pressure drop in the upper plenum, and the pressure drop in the lower plenum. Because the available hydraulic data is limited, the developed simulation model has only explicitly dealt with the pressure loss in the annular space and in the triangular fuel channel.
For the triangular fuel channel in the core region, the Chiu-Rohsenow-Todreas (CRT) model is used to calculate the pressure loss $\Delta P_{c1}$, which is given as follows (Walter, 1981):

$$\Delta P_{c1} = \frac{L}{D_e} f_s \frac{\rho_c \nu^2}{2} \left[ 1 + C \frac{A_e D_e}{A H} \frac{p^2}{\pi p^2 + H^2} \right]$$  \hspace{1cm} (3.6)

where

$$\nu = \frac{\dot{m}}{\rho_c A_t}$$

$A_t$ = flow area of the triangular fuel channel.

$\nu$ = flow velocity (2.2073 m/s).

$f_s$ = skin friction factor

$C$ = empirical coefficient in the CRT model (2200.0).

$L$ = length of the fuel channel (0.454 m).

$H$ = lead of the wire wrap (0.37 m).

$p$ = pitch of the fuel rods (0.007981 m).

$s$ = diameter of the wire wrap (0.00058100 m).

$d$ = fuel rod diameter (0.0074 m).

$A_r$ = projected area of one wire over one lead.

The projected area of one wire over one lead is calculated as follows:

$$A_r = \frac{1}{6} \left[ \frac{\pi}{4} (d + 2s)^2 - \frac{\pi d^2}{4} \right]$$  \hspace{1cm} (3.7)

The flow area of the triangular fuel channel without the wire is calculated as follows:

$$A = \frac{\sqrt{3}}{4} p^2 - \frac{\pi d^2}{8}$$  \hspace{1cm} (3.8)

The hydraulic diameter of the triangular fuel channel, $D_e$, is calculated as follows:

$$D_e = \frac{2\sqrt{3} p^2 - \pi (d^2 - s^2)}{\pi (d + s)}$$  \hspace{1cm} (3.9)
For the annular space, the pressure loss can be calculated as follows:

$$\Delta P = \frac{L}{D_a} f_s \rho_c v^2 \left[ \frac{2}{\pi} \right]$$  \hspace{1cm} (3.10)

$D_a$ = equivalent hydraulic diameter in the annular space.

The hydraulic diameter of the annular space, $D_a$, is calculated as follows:

$$D_a = \frac{\pi(D_2^2 - D_1^2)}{\pi(D_2 + D_1)}$$  \hspace{1cm} (3.11)

$D_1$ = equivalent baffle diameter (0.328m).

$D_2$ = equivalent reactor vessel inner diameter (0.358m).

The skin friction factor are calculated based on the flow regimes as follows (Seo, 1986):

If $Re < 2100$, then $f_s = \frac{64}{Re}$  \hspace{1cm} (3.12a)

If $2100 < Re < 3000$, then $f_s = -0.00147633 + 1.522E - 5 * Re$  \hspace{1cm} (3.12b)

If $3000 < Re < 3.0E6$, then $f_s = 0.0058 + 0.5/Re^{0.32}$  \hspace{1cm} (3.12c)

3.2.6. Results of Reactor Core Model

The developed reactor core model was tested by examining the transient behavior after external reactivity is inserted into the core. The inserted external reactivity is plotted in Figure 3.2a. A positive ramping reactivity is inserted into the core beginning at 1.0 second and ending at 9.0 second with a maximum reactivity of 20 cents. After this reactivity was maintained for 10 seconds, a negative ramping reactivity is inserted into the core such that the external reactivity becomes zero after 10 seconds. Figure 3.2b shows the reactivity changes after the external reactivity was inserted. It is interesting to notice that the negative feedback reactivity did not decrease until the external reactivity began to decrease. This can be explained by the time delay of the temperature responses. After a new steady state condition is reached at about the 350th second, the total reactivity as well as the feedback reactivity becomes zero again.

Figure 3.3 shows the response of the reactor thermal power after the external reactivity was inserted. The rate of change of reactor thermal power is in agreement with the change rate
in the total reactivity. When the total reactivity increases rapidly at the beginning, the reactor thermal power shows a sharp peak.

Figure 3.4 shows the fuel temperature and the fuel cladding temperature when the external reactivity is inserted. Figure 3.5 shows the core coolant exit temperature as well as the core average temperature after the external reactivity is inserted when the core coolant inlet temperature is kept constant at 1,250 K. The reactor fuel temperature follows the reactor thermal power changes in a sensitive manner. During the transient when the peak power reaches 3.0 MWt, the peak fuel temperature reaches 1,440 K with an increase of temperature up to about 65 K. On the contrary, the fuel cladding temperature increases by about 18 K and the core exit temperature increases by about 16 K.

Figure 3.6 shows the pressure drop in the core as a function of mass flow rate. The figure shows that the pressure drop is strongly dependent on the flow regimes. In the turbulent flow regime, the pressure drop varies with the mass flow rate, as is expected according to Equation (3.12c).
Figure 3.2(b). Reactivity insertion into the core.

Figure 3.3. Reactor thermal power after the reactivity insertion.
Figure 3.4. Fuel and cladding temperature after the reactivity insertion.

Figure 3.5. Reactor core coolant temperature after the reactivity insertion.
3.3. Primary Coolant Heat Exchanger Model

3.3.1. Hydraulic Model

Figure 3.7 shows a schematic of the primary heat exchanger model. The primary fluid flows into the hot header and is distributed into individual flow channels of the heat exchanger. Multiple thermoelectric cells are mounted on the surface of each channel. After the heat is removed from the channel surface through thermoelectric cells, the cold fluid converges in the cold header and flows out of the heat exchanger. In the SP-100 design, three reactor loops are configured to transport the heat generated in the core to the thermoelectric devices. The thermoelectric cells are installed on 12 primary heat exchangers. For each primary heat exchanger, 30 channels are used to enhance the heat transfer from the primary fluid to the thermoelectric cells.

The developed hydraulic model is used to calculate the flow distribution inside each channel of the heat exchanger, which is an input to the thermoelectric model. This mass flow distribution is determined by the momentum balance equations of the system.

For the i-th segment of the header or the i-th channel, the mass balance equation is given by

\[ M_i = M_{\text{tot}} - \left( \sum_{j=1}^{i-1} m_j + \frac{m_i}{2} \right) \text{ for } i = 1 \text{ to } N \]  

(3.13)
where

\[ M_{tot} = \text{total mass flow rate into the heat exchanger (14.1 kg/s).} \]

\[ M_i = \text{mass flow rate in the i-th header segment.} \]

\[ m_j = \text{mass flow rate in the i-th channel.} \]

The pressure drop in the i-th hot header segment, \( \Delta P_i^H \), is given by

\[
\Delta P_i^H = \frac{w}{D_i} f_i^H \rho_i V_i^2 \frac{\pi}{2}
\]

(3.14)

where

\[
V_i = \frac{4M_i}{\pi \rho_i D_i^2}
\]

(3.15)

\[ D_i = \text{the diameter of the i-th header segment (0.0302 m).} \]

\[ w = \text{the length of the i-th header segment (0.0254 m).} \]

Therefore, the pressure drop, given in Equation (3.14), can be simplified as follows:

\[
\Delta P_i^H = C_i^H M_i^2
\]

where

\[
C_i^H = \frac{8w f_i^H}{\pi^2 \rho_i D_i^5}
\]
Similarly, the pressure drop in the i-th channel, $\Delta P_{i}^{CH}$, is calculated as follows:

$$\Delta P_{i}^{CH} = C_{i}^{c} m_{i}^{2}$$ (3.16)

where

$$C_{i}^{c} = \frac{(w + t) f_{i}^{c}}{4 \rho_{i} (wt)^{3}}$$ (3.17)

$w$ = channel width (0.0254 m).
$t$ = channel height (0.0035 m).
$L$ = channel length (0.254 m).

If the fluid loop in Figure 3.8 is examined, a pressure balance equation can be derived as follows:

$$\Delta P_{i}^{CH} - \Delta P_{i}^{H} - \Delta P_{i+1}^{CH} - \Delta P_{i}^{C} = 0$$ (3.18)

where

$\Delta P_{i}^{C}$ = pressure drop of the i-th cold header segment.
By combining Equation (3.13) and the N equations given in Equation (3.18), the following 2N coupled matrix algebraic equations are obtained (Seo, 1986):

\[
\begin{bmatrix}
1 & 1 & 0 & \ldots & 0 \\
C^c_{1} m_1 & -C^c_{2} m_2 & 0 & \ldots & 0 \\
0 & C^c_{2} m_2 & -C^c_{3} m_3 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \ldots & C^c_{N-1} m_{N-1} & -C^c_{N} m_{N}
\end{bmatrix}
\begin{bmatrix}
m_1 \\
m_2 \\
m_3 \\
\vdots \\
m_N
\end{bmatrix}
= 
\begin{bmatrix}
M_{tot} \\
C^c_1 M^2_1 + C^{hi}_2 M^2_2 \\
C^c_2 M^2_2 + C^{hi}_3 M^2_3 \\
C^c_{N-1} M^2_{N-1} + C^{hi}_N M^2_N
\end{bmatrix}
\begin{bmatrix}
m_1 \\
m_2 \\
m_3 \\
\vdots \\
m_N
\end{bmatrix}
\]  
(3.19)

The 2N equations given in Equation (3.19) can be used to calculate the mass flow distributions in the N header segments, \( \{M_1, M_2, \ldots, M_N\} \), and the mass flow rate in the N channels, \( \{m_1, m_2, \ldots, m_N\} \).

### 3.3.2. Thermal Model of Primary Heat Exchanger

The thermal model of primary heat exchanger is fully coupled with the thermoelectric model. The heat removal term in the thermal model of the primary heat exchanger depends on the hot shoe temperature of the thermal-electric device, and the heat addition term used in the thermoelectric model depends on the fluid temperature of the primary heat exchanger. In the developed component model of the primary heat exchanger, it is assumed that the hot shoe temperature of the thermoelectric cell is known.

Although 480 TE cells are mounted on the surface of one channel of a heat exchanger, one node is used to describe the heat removal by TE cells for simplification. With this simplification, the energy balance equation for one channel of the primary heat exchanger model is given as follows:

\[
\dot{m} c_p (T_{hx,\text{in}} - T_{hx,\text{out}}) = (UA)_{hx} (T_{hx,\text{avg}} - T_{\text{hot}})
\]

where

\( T_{hx,\text{in}} = \) inlet temperature of the heat exchanger channel of interest.
\( T_{hx,\text{out}} \) = outlet temperature of the heat exchanger channel.

\( T_{hx,\text{avg}} \) = average temperature of the heat exchanger channel.

\( T_{\text{hot}} \) = hot shoe temperature of the corresponding TE cells.

\( (UA)_{hx} \) = overall thermal conductance between the hot shoe material and the bulk fluid in the heat exchanger channel.

The overall heat conductance calculated as follows:

\[
\frac{1}{(UA)_{hx}} = \frac{1}{\sum_{i=1}^{N} \frac{t_i}{k_i} + \frac{1}{h_{hx}}A_{hx}}
\]

where

\( A_{hx} \) = convection heat transfer area, which is assumed to be equal to the heat conduction area of the TE padding material.

\( t_i \) = thickness of the i-th layer of the TE padding material.

\( k_i \) = heat conductivity of the i-th layer of the TE padding materials.

\( h_{hx} \) = forced convection heat transfer coefficient.

The forced convection heat transfer coefficient between the fluid and the channel walls is calculated based on the following correlation (Seo, 1986):

\[
h_{hx} = \frac{Nu \cdot k}{d_e}
\]

where

\[
Nu = \frac{1}{\gamma(5.8 + 0.02Pe^{0.8})}
\]

\( \gamma \) = heat transfer coefficient of liquid metals flowing between two parallel plates with one of them being adiabatic to that with two plates of equal heat flux.
In the developed thermal model of the primary heat exchanger, five layers of padding materials, including the channel wall, are considered. The thickness and the thermal conductivity used are listed in Table 3.2.

<table>
<thead>
<tr>
<th>Layer</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Material name</td>
<td>Nb-Zr</td>
<td>Al2O3</td>
<td>Tungsten</td>
<td>Forsterite</td>
<td>Nickel</td>
</tr>
<tr>
<td>t (cm)</td>
<td>0.05</td>
<td>0.0254</td>
<td>0.127</td>
<td>0.0025</td>
<td>0.03</td>
</tr>
<tr>
<td>K (w/cm/K)</td>
<td>41.9</td>
<td>25.0</td>
<td>163.3</td>
<td>8.0</td>
<td>60.7</td>
</tr>
</tbody>
</table>

### 3.3.3. Results of Primary Heat Exchanger Model

Figure 3.9 shows the pressure distribution in the primary heat exchanger. The upper curve and the lower curve in this figure represent the pressure distribution along the hot headers and cold headers, respectively. The calculated results are based on a mass flow rate of 1.175 kg/s into the primary heat exchanger. The total pressure drop is 519.2 Pa.

Figure 3.10 shows the mass flow distribution along the channel when the heat exchanger inlet flow rate is 1.175 kg/s. The mass flow rate into the first channel is 0.0555 kg/s and the mass flow rate into the last channel is 0.0262 kg/s. There is significant uneven flow distribution among the channels, which will cause significant uneven temperature distribution on the surface of the primary heat exchangers.

Figure 3.11 shows the coolant and hot shoe temperatures along the channels when the heat flux to the TE module is fixed at a constant value of 11.5741 W. As expected, the TE hot shoe temperatures vary by about 15 K for different channels. The temperature difference between the coolant and the hot shoe is approximately 25 K; this is needed to maintain the specified heat flux flowing from the heat exchanger coolant to the thermoelectric cells.
Figure 3.8. Pressure distribution in the heat exchanger.

Figure 3.9. Mass flow rate in the channel.
3.4. Thermal Electric Model

3.4.1. Thermoelectricity Phenomenon

Seebeck Effect

When an electrically conductive material is subjected to a temperature gradient, an electric voltage is generated. This effect is called the Seebeck effect of an electrically conductive material. If a P-type semiconductor material and a N-type semiconductor are connected with the two junctions at different temperatures, the Seebeck electric potential $\Delta V$ will be induced across the unpaired terminal as a result of the difference between the internal potential with the two semiconductors. To characterize this instantaneous thermoelectric phenomenon, the Seebeck coefficient is defined as follows:

$$\alpha_{mn} = \alpha_p - \alpha_n = \lim_{\Delta T \to 0} \frac{\Delta V}{\Delta T}$$  \hspace{1cm} (3.24)

where

$\alpha_p$ = absolute Seebeck coefficient of the P-type semiconductor (see Appendix B).

$\alpha_n$ = absolute Seebeck coefficient of the N-type semiconductor (see Appendix B).
The absolute Seebeck coefficient is dependent on the material as well as the temperature. For a pair of P-type and N-type semiconductors, the electric potential difference, as a result of the temperature difference $\Delta T$ between the two junctions, may be calculated as follows:

$$\Delta V = -\int_0^L \alpha_p(T) \frac{\partial T_p}{\partial x} \, dx + \int_0^L \alpha_n(T) \frac{\partial T_n}{\partial x} \, dx$$

(3.25)

where

$L =$ length of the semiconductor materials.

It should be noticed that both the first and the second terms in Equation (3.25) are positive since $\alpha_p$ is positive and $\alpha_n$ is negative. For this reason, the Seebeck voltage can be approximately doubled, which explains why there is always a pair of semiconductor materials in any thermoelectric device.

If a thermoelectric circuit is configured, the electric current, due to the temperature difference induced Seebeck voltage can be calculated as follows:

$$I = \frac{\Delta V}{R}$$

(3.26)

where

$$R = R_L + R_p + R_N$$

(3.27)

where

$R_L =$ electric resistance of the external load.

$R_p =$ internal electric resistance of the P-type semiconductor.

$R_N =$ internal electric resistance of the N-type semiconductor.

The internal electric resistances of the P-type and the N-type semiconductors can be calculated as follows:

$$R_p = \int_0^L \rho_p(T) \frac{dx}{A_p}$$

(3.28a)
\[ R_n = \int_0^L \rho_n(T) \frac{dx}{A_n} \]  

(3.28b)

where \( \rho_p \) and \( \rho_n \) are the electric conductance of the P-type and the N-type semiconductors, respectively (See Appendix B).

**Peltier Effect**

When an electric current crosses an interface between two different electric conductors, heat will be absorbed or liberated. This phenomenon is called Peltier effect of an electric conductor. The reason why Peltier effect occurs is that the entropy of the electric charge carriers changes as they cross an interface between two different electric conductors.

If a circuit of two dissimilar semiconductor materials carries an electric current, one junction will absorb heat and the other junction will give off heat. The absorbed heat or the liberated heat is proportional to the electric current, with the proportionality factor given by the Peltier coefficient, is given by

\[ Q_{\text{peltier}} = \pi_{pn} I \]  

(3.29)

where

\[ I = \text{the electric current across the circuit.} \]

\[ \pi_{pn} = \pi_p - \pi_n \]  

(3.30)

\[ \pi_p = \text{absolute Peltier coefficient of the P-type material.} \]

\[ \pi_n = \text{absolute Peltier coefficient of the N-type material.} \]

According to Kelvin’s first law, it can be proved that the absolute Peltier coefficient, which is actually the entropy of the electric carrier for a unit current in the circuit, can be related to the absolute Seebeck coefficient as follows (Soo, 1968):

\[ \pi = \alpha T \]  

(3.31)

If the Seebeck coefficient is considered to be temperature dependent, the absorbed heat due to Peltier effect at the hot junction can be calculated as follows:
\[ Q_{\text{Peltier}} = I \cdot [\alpha_p(T_p) \cdot T_p(0,t) - \alpha_n(T) \cdot T_n(0,t)] \]  

(3.32a)

where \( Q_{\text{Peltier}} > 0 \) indicates heat absorption.

The liberated heat due to Peltier effect at the cold junction is calculated as follows:

\[ Q_{\text{Peltier}} = I \cdot [\alpha_p(T_p) \cdot T_p(L,t) - \alpha_n(T) \cdot T_n(L,t)] \]  

(3.32b)

**Thomson Effect**

When an electric current passes a homogeneous conductor over which there is a temperature difference, heat is absorbed or liberated. This phenomenon is called Thomson effect of electric conductor. The Thomson effect can be described by Thomson coefficient, which is defined as follows:

\[ \tau = \lim_{\Delta T \to 0} \frac{\Delta Q/I}{\Delta T} \]  

(3.33)

where the Thomson coefficient \( \tau \) satisfies the second Kelvin’s relation, given by

\[ \tau = T \frac{d\alpha}{dT} \]  

(3.34)

If the physical properties are temperature dependent, the absorbed or liberated heat due to the temperature difference \( \Delta T \) between the two junctions can be calculated as follows:

\[ Q_T = -I \cdot [\int_0^l \tau_p(T) \frac{\partial T_p}{\partial x} \, dx - \int_0^l \tau_n(T) \frac{\partial T_n}{\partial x} \, dx] \]  

(3.35)

where \( Q_T > 0 \) indicates heat generation.

**Joule Effect**

When an electric current passes in a conductor, the electric energy is converted into heat by Joule heating. The Joule heating is given by

\[ Q_J = I^2 R \]  

(3.36)

where
\[ R = R_p + R_N \]  

**Fourier Effect**

Heat is conducted from high temperature to low temperature. According to Fourier law, the heat flux is given by:

\[ q = -k \frac{dT}{dx} \]  

where

- \( q \) = heat flux due to heat conduction in the semiconductor materials.
- \( k \) = thermal conductivity of the semiconductor materials.

### 3.4.2. Thermal Electric Generator Model

In the SP-100 design, the electricity is generated by 3 loops \( \times \) 12 primary heat exchangers \( \times \) 30 channels \( \times \) 480 TE cells. Each TE cell consists of two semiconductors, one P-type semiconductor and one N-type semiconductor. The hot shoes of TE cells are mounted on the two surface plates of the 12 primary heat exchangers. The cold shoes of TE cells are connected to radiators through heat pipes. In the current version of the simulation model, the transient heat transfer from the cold shoes of TE cells to the radiators is ignored. Because there is a temperature gradient between the hot shoes and the cold shoes of TE cells, when heat is conducted from the hot shoe of a TE cell to its cold shoe, electric power will be generated due to the Seebeck effect.

The governing equations to describe the temperature distribution in the P-type semiconductor and the N-type semiconductor are given as follows (Soo, 1968):

\[
\frac{\partial}{\partial x} [k_p \frac{\partial T_p}{\partial x}] + \rho_p J^2 - \tau_p J \frac{\partial T_p}{\partial x} = w c_p \frac{\partial T_p}{\partial t} \tag{3.39a}
\]

\[
\frac{\partial}{\partial x} [k_n \frac{\partial T_n}{\partial x}] + \rho_n J^2 + \tau_n J \frac{\partial T_n}{\partial x} = w c_p \frac{\partial T_n}{\partial t} \tag{3.39b}
\]

where \( J \) always takes positive sign such that the Thomson effect is positive to the heat storage of P-type material and negative to the heat storage of N-type material.
Equations (3.39a) and (3.39b) are the energy balance equations for P-type semiconductor and N-type semiconductor, respectively. On the left side of the two equations, the first term corresponds to the heat addition due to thermal conduction, the second term corresponds to Joule heat, and the third term corresponds to Thomson effect. The term on the right side of the two equations corresponds to the heat storage. Obviously, only volumetric effects appear in the governing equations to describe the temperature distribution.

The boundary condition on the hot shoe of a pair of semiconductor elements is explicitly stated in this research as follows:

\[-k_p A_p \frac{\partial T_p (0,t)}{\partial x} + 0.5(\alpha_p T_p (0,t) - \alpha_n T_n (0,t))JA_p = q_h (t)A_p\]  
(3.40a)

\[-k_n A_n \frac{\partial T_n (0,t)}{\partial x} + 0.5(\alpha_p T_p (0,t) - \alpha_n T_n (0,t))JA_n = q_h (t)A_n\]  
(3.40b)

where

\[q_h (t) = \text{heat flux flowing into the hot shoe of TE cells.}\]

The boundary condition on the cold shoe of a pair of semiconductor elements is stated as follows:

\[-k_p A_p \frac{\partial T_p (L,t)}{\partial x} + 0.5(\alpha_p T_p (L,t) - \alpha_n T_n (L,t))JA_p = q_{c,p} (t)A_p\]  
(3.40c)

\[-k_n A_n \frac{\partial T_n (L,t)}{\partial x} + 0.5(\alpha_p T_p (L,t) - \alpha_n T_n (L,t))JA_n = q_{c,n} (t)A_n\]  
(3.40d)

where

\[q_{c,p} (t) = \text{heat flux flowing out of the cold shoe of TE cells due to radiation.}\]

\[q_{c,p} (t) = \varepsilon F \sigma (T_p (L,t)^4 - T_a^4)\]
\[q_{c,n} (t) = \varepsilon F \sigma (T_n (L,t)^4 - T_a^4)\]

\[\varepsilon = \text{emissivity (0.85).}\]

\[F = \text{ratio of the surface area of the radiator to that of the TE cell cross section (15).}\]

\[\sigma = \text{Stephan-Boltzman constant (5.67E-8 Wm-2 K-4).}\]

\[T_a = \text{ambient temperature in space (253 K).}\]
In order to close the equations given by Equations (3.39a) and (3.39b), the Seebeck voltage and the electric current are calculated using Equations (3.25) and (3.26), respectively.

3.4.3. Energy Balance Analysis

The generalized relationship between the thermal power flow and the electric power output has been derived in this research for the developed model of the thermoelectric converter. For the operating temperature of TE converters for SP-100 design, the temperature difference between the hot shoe and the cold shoe is as great as 500 K. Therefore, it is necessary to take into account the temperature dependence of the thermal properties and the electric properties for the semiconductors. Because the energy balance analysis in most research is based on constant hot shoe temperature and constant cold shoe temperature, as well as on physical properties, we examined the energy balance to verify the results of our developed thermoelectric model when temperature dependent physical properties are considered.

During both steady state and transient conditions, for a pair of TE cells, the total heat flowing into the junction of the hot shoe $Q_H$ can be written as follows:

$$Q_H = Q_{peltier}(T_h) + Q_L$$  \hspace{1cm} (3.41)

where

$Q_{peltier}(T_h)$ = the absorbed heat due to Peltier effect at the hot junction.

$Q_L$ = heat flowing into the semiconductors through heat conduction.

$$Q_L = -k_p A_p \frac{\partial T_p(0,t)}{\partial x} - k_n A_n \frac{\partial T_n(0,t)}{\partial x}$$  \hspace{1cm} (3.42)

The heat flowing out of the junction of the cold shoe $Q_C$ is as follows:

$$Q_C = Q_{peltier}(T_c) + Q_R$$  \hspace{1cm} (3.43)

where

$Q_{peltier}(T_c)$ = liberated heat due to Peltier effect at the cold junction.

$Q_R$ = heat flowing out of the semiconductors through heat conduction.
If the governing equations (3.39a) and (3.39b) are integrated, the following steady state balance equation can be derived for the heat conduction in the semiconductors:

\[ Q_R - Q_L = I^2 (R_p + R_N) + Q_T \]  

(3.45)

where the first term on the right side is the volumetric Joule heat generation and the second term is the volumetric Thomson heat generation.

If the TE generator system is considered as a whole, the energy balance equation can be derived as follows:

\[ Q_{IT} = Q_C + Q_e \]  

(3.46)

where

\[ Q_e = \text{the electric power.} \]

Equation (3.46) provides a universal representation of the energy balance for thermoelectric conversion devices whether the physical properties of the semiconductors are temperature dependent or not. In order to verify that our model is correct, a detailed proof of the energy balance equation is given in this section.

**Proof:** The difference between the absorbed Peltier heat on the hot shoe and the liberated Peltier heat on the cold shoe is given as follows:

\[ Q_{\text{peltier}}(T_h) - Q_{\text{peltier}}(T_c) = I \cdot [\alpha_{ph} T_p(0,t) - \alpha_{ph} T_p(L,t)] - I \cdot [\alpha_{nh} T_n(0,t) - \alpha_{nh} T_n(L,t)] \]

\[ = -\int_0^L d(\alpha_p T_p) + \int_0^L d(\alpha_n T_n) = I \cdot [-\int_0^L \alpha_p d(T_p) + \int_0^L \alpha_n d(T_n)] - I \cdot [\int_0^L T_p d(\alpha_p) - \int_0^L T_n d(\alpha_n)] \]

(3.47a)
According to Seebeck law, the Seebeck voltage is rewritten as follows:

\[ \Delta V = -[\int_0^L \alpha_p(T) \frac{\partial T_p}{\partial x} dx - \int_0^L \alpha_n(T) \frac{\partial T_n}{\partial x} dx] \]  

(3.47b)

Therefore, according to Ohm’s law, the term inside the first brackets on the right side of Equation (3.47a) can be written as follows:

\[ \Delta V \cdot I = I^2 \cdot (R_L + R_p + R_n) \]

(3.47c)

The term inside the second brackets on the right hand side of Equation (47a) is related to the Thomson heat, which is written as follows:

\[ Q_T = -I \cdot [\int_0^L \tau_n(T) \frac{\partial T_n}{\partial x} dx - \int_0^L \tau_p(T) \frac{\partial T_p}{\partial x} dx] \]

(3.47d)

since

\[ \tau = T \frac{d\alpha}{dT} . \]

According to the steady state governing equations of TE generators, the following relationship exists:

\[ Q_T - Q_L = I^2 (R_p + R_n) + Q_T \]

(3.47e)

If we add Equations (3.47a) and (3.47e), then we have

\[ Q_L + Q_{pelletier}(T_h) = Q_R + Q_{pelletier}(T_c) + Q_e \]

(3.47f)

This completes the proof of Equation (46).

The energy efficiency \( \eta_e \) of the TE system is defined as follows:
\[ \eta_e = \frac{Q_e}{Q_h} \]

In the case that the physical properties are independent of the temperature, the Thomson effect may be ignored. If we further assume that the P-type semiconductor and the N-type semiconductor have the same physical properties and the hot shoe temperature and the cold shoe temperature are constant, the resulting temperature distribution based on the governing equation with the specified boundary conditions will have the following solutions (Soo, 1968):

\[ T_p(x) = [T_{ph} - \frac{x}{L}(T_{ph} - T_{pc})] + \frac{J^2 \rho_p}{2\lambda_p}x(L - x) \]  \hspace{1cm} (3.48)

\[ T_n(x) = [T_{nh} - \frac{x}{L}(T_{nh} - T_{pc})] + \frac{J^2 \rho_n}{2\lambda_n}x(L - x) \]  \hspace{1cm} (3.49)

We can verify that the following relationships exist:

\[ Q_R - Q_L = I^2(R_p + R_n) \]  \hspace{1cm} (3.50)

In the meanwhile, the load voltage is given by:

\[ V_L = R_LI = \alpha_{pn} \left( \frac{T_{ph} + T_{nh}}{2} - \frac{T_{pc} + T_{nc}}{2} \right) - (R_n + R_p)I \]  \hspace{1cm} (3.51)

where \( R_L \) is the electric resistance of the load.

The electric current is given by:

\[ I = \frac{\alpha_{pn} \left( \frac{T_{ph} + T_{nh}}{2} - \frac{T_{pc} + T_{nc}}{2} \right)}{R_L + R_n + R_p} \]  \hspace{1cm} (3.52)

The heat absorption and the heat liberation at the two junctions due to the Peltier effect are as follows:

\[ Q_p(T_h) = \alpha_{pn}I \frac{T_{ph} + T_{nh}}{2} \]  \hspace{1cm} (3.53a)
\[ Q_p(T_c) = \alpha \frac{T_{pc} + T_{nc}}{2} \]  \hfill (53b)

It is easy to verify that the following energy conservation relationship does exist.

\[ Q_L + Q_p(T_h) = Q_R + Q_p(T_c) + I^2 R_L \]  \hfill (3.54)

In Seo’s report, there is a misunderstanding about the energy balance for TE cells, where the heat conduction into and out of the TE semiconductors are related as follows:

\[ Q_L = Q_k + 0.5(Q_J + Q_T) \]  \hfill (3.55a)
\[ Q_R = Q_k + 0.5(Q_J + Q_T) \]  \hfill (3.55b)

where

\[ Q_k \] = steady state heat flux due to the heat conduction in the semiconductor materials.

If we examine the derivation described in this section, it can be found that Equations (3.55a) and (3.55b) are correct only if Equations (3.49a) and (3.49b) hold. In general, when the physical properties are dependent on temperature, the heat flux into and out of the semiconductor materials will be a complicated function of the heat flux flowing inside the semiconductor materials. However, the energy balance equation given by Equation (3.46) is always true and the calculation of the TE efficiency does not need the information about where the Joule heat and the Thomson heat are distributed.

### 3.4.4. Results of TE Model

Figures 3.11 and 3.12 show the steady state responses of a single TE cell as a function of the external load based on the developed TE model. Figure 3.11 shows that the electric power generated by the TE cell is able to follow the power demand only up to a certain external load value. In other words, the electric power output does not increase as the electric load increases after the electric load is greater than a certain critical value. In Figure 3.11, this critical value is about 0.017 Ohms. It can be theoretically proved that the critical value occurs when the external load is equal to the total resistance of the P and N elements. The reason why the electric power
output has a peak value is that there are two competitive mechanisms to control the electric output. The positive mechanism is that the electric power increases due to the increase in Seebeck voltage as external load increases. The negative mechanism is that the thermal input to the TE generators decreases due to the decrease in the Peltier’s heat absorption when the electric current decreases as the external load increases.

In Figure 3.12, the calculated results show that the efficiency of TE generators has a peak value, which is almost independent of the external load but is dependent on the TE hot shoe temperature. As the hot shoe temperature increases, the TE input power, the Seebeck voltage, and the TE current all increase. After the external load reaches the critical value for a given TE hot shoe temperature, which is not sensitive to the TE hot shoe temperature, the decrease in the input thermal power and the increase in the Seebeck voltage approach asymptotic values. Therefore, the thermal efficiency will continue to decrease because of the continuing decrease in the electric current as the external load increases.

Figure 3.13 shows a 3-D plot of the electric power output as a function of the external load and the TE hot shoe temperature. This figure clearly shows that the TE hot shoe temperature has a strong influence on the thermal efficiency but has almost no effect on the critical load. The figure also shows that the electric power is able to follow the power demand only up to a critical load value. Figure 3.14 shows a 3-D plot of the electric power output as a function of the external load and the TE hot shoe temperature. This shows that the thermal efficiency increases as the TE hot shoe temperature increases. However, the rate of increase starts decreasing as the TE hot shoe temperature becomes greater.
Figure 3.11. Electric power versus external load.

Figure 3.12. Steady state response of TE generator as a function of external load.
Figure 3.13. Electric power output versus external load and TE hot shoe temperature.

Figure 3.14. TE efficiency versus external load and TE hot shoe temperature.
3.5. Review of the SP-100 Electromagnetic Pump

The Electromagnetic Pump (EMP) is a critical component of the SP-100 system. The function of the EM pump is to pump lithium coolant from the reactor core to the energy conversion system and overcome the hydraulic resistance of the primary fluid in the primary heat transport system. In the SP-100 system, three EM pumps are used to constitute three reactor loops and transport thermal energy from the core to the power conversion assemblies.

The EM pumps of SP-100 system are self-actuated. The electric current is generated using thermoelectric elements positioned between the high temperature primary coolant and low temperature secondary coolant. The pumping force is created by making the electric current cross a magnetic field. Therefore, as long as there is a temperature difference between the primary and secondary coolant, the pump will continue to operate as designed.

The SP-100 system uses a rectangular design, a schematic of which is shown in Figure 3.15. The electric current generated by the SiGe/Ga thermoelectric generators forms a loop in the plane of the paper. The direction of the magnetic field created by Hiperco-27 permanent magnet is from the East to the West. The lithium coolant in the two coolant ducts between the magnetic structures is subjected to a magnetic force and flows in the coolant ducts along the directions indicated in the figure.

The EM pumps of SP-100 systems can be designed to operate in controllable mode or uncontrollable mode. In the uncontrollable mode design, the pressure rise of the pumps is determined by the electric current provided by the TE generators and the magnetic flux density created by the Hiperco-27 permanent magnet. In the uncontrollable mode design, the magnetic flux created by the magnet should be adjustable.
A simplified EM pump model can be developed by assuming the coolant duct is rectangular with constant area, and the magnetic flux is constant. The input parameters of the EM pump model include the electric current, the magnetic field, and the geometric data of the coolant ducts.

Figure 3.16 shows the diagram of principle used for the EM pump simplification. The subjected force on the coolant in the pump duct, $F$, is given as follows:

$$F = \frac{BI_e a}{10^4}$$

where

- $B =$ magnetic flux density, Gauss.
- $I_e =$ effective electric current flowing through the coolant inside the coolant duct, amp.
- $a =$ effective width of the coolant duct, m.

The coolant pressure rise $\Delta P$ due to the pumping force is then given as follows:

$$\Delta P = \frac{F}{ab} = \frac{BI_e}{10^4 b}$$
As the electrically conductive coolant in the coolant duct moves in the magnetic field, an induced voltage, $E_i$, will be generated, which is given by:

$$E_i = \frac{BQ}{10^4b}$$

where

$Q = $ the volumetric flow rate of the coolant in the coolant duct.

Suppose that the total electric current traversing the coolant duct is $I$. This should be equal to the sum of the effective current passing through the coolant inside the coolant duct, $I_e$, and the current passing through the duct wall, $I_w$, if the current passing through the coolant outside the magnetic field boundary can be ignored. This is given by

$$I = I_e + I_w$$

Therefore, the total voltage drop across the coolant duct is given by

$$E = E_i + I_e R_e = I_w R_w$$

where

$R_e = $ electric resistance of liquid metal in the coolant duct.

$R_w = $ electric resistance of wall material of the coolant duct.

Therefore, the effective current passing through the coolant inside the coolant duct, $I_e$, can be calculated as follows:

$$I_e = \frac{1}{R_e + R_w}(I \cdot R_w - \frac{BQ}{10^4b})$$

Furthermore, the relationship between the pump pressure rise and the volumetric flow rate can be written as follows:
\[ \Delta P = \frac{B}{10^4 b} \frac{1}{R_e + R_w} (I \cdot R_w - \frac{BQ}{10^4 b}) \]

3.6. Integration of the SP-100 System Model

The developed simulation model must be able to reproduce the SP-100 main design parameters at the nominal operation condition before it can be used for transient analysis and control study. In this section, the initialization of the individual component models and the integration of the SP-100 system model are presented. A schematic of the integrated model is shown in Figure 3.17.
Figure 3.17. The Integrated model of SP-100 system.

Table 3.3 shows a comparison between simulation results and design parameters at full operation condition. Most of the simulation results are in good agreement with the SNPSAM design parameters. The big differences are in TE hot side and cold side temperatures. The reason for the differences is that a secondary lithium loop is added to the developed simulation model while TE cold shoes are directly connected to radiators in SNPSAM design. However, the calculated electric conversion efficiency is close to the value of SNPSAM. Therefore, the developed simulation model can still be used to provide a credible control design for SNPSAM space reactor system.

Table 3.3. Comparison between simulation results and design parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>SNPSAM Design</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal thermal power (kW)</td>
<td>2000.0</td>
<td>2000.3</td>
</tr>
<tr>
<td>Electric power output (kW)</td>
<td>112.0</td>
<td>112.6</td>
</tr>
<tr>
<td>System efficiency (%)</td>
<td>5.59</td>
<td>5.63</td>
</tr>
<tr>
<td>Core inlet temperature (K)</td>
<td>1254.0</td>
<td>1254.0</td>
</tr>
<tr>
<td>Core outlet temperature (K)</td>
<td>1284.0</td>
<td>1284.0</td>
</tr>
<tr>
<td>Fuel average temperature (K)</td>
<td>1376.0</td>
<td>1376.0</td>
</tr>
<tr>
<td>Cladding average temperature (K)</td>
<td>1288.0</td>
<td>1288.0</td>
</tr>
<tr>
<td>TE hot side temperature (K)</td>
<td>1237.0</td>
<td>1228.5</td>
</tr>
<tr>
<td>TE cold side temperature (K)</td>
<td>857.0</td>
<td>841.8</td>
</tr>
</tbody>
</table>
The developed model was used to simulate the step responses of the increase in the control drum angle, which is shown in Figure 3.18. The control drum angle has a step increase of 10 degrees at the 10th second. Correspondingly, a step increase in reactivity is inserted into the core with a maximum value of 13.95 cents. Due to the feedback arising from Doppler effects and reactor coolant expansion, the total reactivity becomes zero at the 150.0th second. This indicates that the space reactor is a self-regulated process.

Figure 3.18 also shows the transients of reactor thermal power and the TE electric power output. The reactor thermal power follows the total reactivity and a power peak can be observed because the reactor thermal power change rate rather than reactor thermal power is proportional to the total reactivity. The electric power output follows the reactor core outlet temperature. A time lag can be observed between the electric power output and the reactor thermal power.

Figure 3.18. Step response of the increase in control drum angle.

Figure 3.19 also shows the step response of step increase in external load. When the external load changes by 0.01 Ohms for one TE element, TE electric power output has a quick response. As compared with the disturbance on control drum angle, changing external load may have a load following capability without disturbing the reactor thermal power as well as the reactor inlet
and outlet temperature significantly. In addition, the TE electric power output responds to the change in external load in a much faster speed than the change in control drum angle.

Figure 3.19. Step response of the increase in external load.

3.7. Summary of Model Equations for Control

**Neutron Kinetics Model**

Point reactor kinetics model with six delayed neutron groups is used to describe the dependence of nuclear reactor power on reactivity change, and is given by:

\[
\frac{dP_n(t)}{dt} = \rho(t) - \beta_{se} P_n(t) + \sum_{i=1}^{6} \lambda_i C_i(t)
\]

\[
\frac{dC_i(t)}{dt} = \frac{\beta_i}{\Lambda} P_r(t) - \lambda_i C_i(t)
\]

where

\( P_r(t) = \) instantaneous nuclear power (MWt).
\( \beta_{-6} \) = fractions of the 6 groups of delayed neutrons.

\( \beta_{ss} \) = total fraction of delayed neutrons.

\( \lambda_{ni} \) = decay constants of 6 groups of delayed neutrons.

\( C_{i}(t) \) = concentration of the precursors of delayed neutrons.

\( \Lambda \) = average neutron generation time (s).

\( \rho(t) \) = total reactivity at time \( t \).

\( \rho(t) = \rho_{cn}(t) + \rho_{m}(t) \)

\( \rho_{cn} \) = external reactivity inserted by the control drum.

\( \rho_{m} \) = feedback reactivity.

The neutronic data used to solve the reactor kinetics equation are obtained from [8].

**Feedback Reactivity**

The core model takes into account three mechanisms of reactivity feedback. The reactor fuel Doppler feedback reactivity, \( \rho_{D} \), represents the competitive effects of the absorption of fertile material such as U-238 and the fission of fissile material such as U-235 when fuel temperature changes. The reactor core expansion feedback reactivity, \( \rho_{E} \), represents the reactivity effect due to the expansion of reactor core materials including fuel, fuel cladding, and reflector. The reactor core coolant expansion coefficient, \( \rho_{C} \), takes into account the reactivity effect of neutron leakage when the temperature of reactor core coolant changes.

The total reactivity feedback is calculated as follows:

\[
\rho_{FB}(t) = \rho_{D}(t) + \rho_{E}(t) + \rho_{C}(t)
\]

where

\[
\rho_{D}(t) = \alpha_{D} \ln \frac{T_{f}(t)}{T_{f}(0)}
\]

\[
\rho_{E}(t) = \alpha_{E}(T_{f}(t) - T_{f}(0)) + \alpha_{clad}(T_{clad}(t) - T_{clad}(0))
\]

\[
\rho_{C}(t) = \alpha_{C}(T_{c}(t) - T_{c}(0))
\]

where

\( \alpha_{D} \) = fuel Doppler temperature coefficient

\( \alpha_{E} \) = fuel expansion feedback coefficient.
\( \alpha_{\text{clad}} = \) fuel cladding expansion feedback coefficient.

\( \alpha_c = \) reactor core coolant expansion feedback coefficient.

\( T_r = \) fuel temperature (K).

\( T_{\text{clad}} = \) fuel cladding temperature (K).

\( T_c = \) reactor core coolant temperature (K).

**Reactivity Control Mechanism**

The external reactivity control in the developed model uses the mechanism of the stepper motor control drum system. The control drum shaft can be rotated from 0 to 180 degrees. The control voltage is transformed into a set of 27V rectangular pulses, with the frequency varying from 0 to 1.33 Hz. These pulses are then converted into discrete movements of the shaft connected to the control drums. The position of the shaft of the stepper motor is given as follows:

\[
\frac{d^2 \theta}{dt^2} + 1.01 \frac{d\theta}{dt} = 0.0525u
\]

where

\( u = \) the control voltage (V).

\( \theta = \) the shaft angle of the stepper motor.

The external reactivity inserted by the control drum can be fitted as a function of the shaft angle of the stepper motor, and is given by

\[
\rho_{\text{cd}} = 6.89 \times 10^{-13} \theta^5 - 2.33 \times 10^{-10} \theta^4 + 3.28 \times 10^{-9} \theta^3 + 4.57 \times 10^{-6} \theta^2 - 5.88 \times 10^{-5} \theta
\]

where

\( \rho_{\text{cd}} = \) integral reactivity worth due to the control drum corresponding to the positioned angle.

**Core Heat Transfer Model**

A simplified reactor core heat transfer model is developed to calculate fuel temperature, fuel cladding temperature, and average core coolant temperature. The fuel temperature \( T_r \), the
fuel cladding temperature $T_{\text{clad}}$, and the average core coolant temperature $T_c$ are described as follows.

\[
\frac{d}{dt} T_f(t) = \frac{P_f(t)}{m_c} - (T_f - T_{\text{clad}}) \cdot (UA_f) T_c(t) \cdot (U - T_{\text{clad}}) \cdot (UA)_{\text{clad}}
\]

\[
C_{\text{clad}} \frac{d}{dt} T_{\text{clad}}(t) = (T_f - T_{\text{clad}})(UA)_f - (T_{\text{clad}} - T_c) \cdot (UA)_{\text{clad}}
\]

\[
0.5 C_c \frac{d}{dt} T_c(t) = 0.5(T_{\text{clad}} - T_c) \cdot (UA)_{\text{clad}} - \dot{m}_c c_p (T_c - T_{\text{clad}})
\]

\[
0.5 C_c \frac{d}{dt} T_{\text{pin}}(t) = 0.5(T_{\text{clad}} - T_{\text{pin}}) \cdot (UA)_{\text{clad}} - \dot{m}_p c_p (T_{\text{pin}} - T_c)
\]

where

- $T_f$ = average fuel temperature (K).
- $T_{\text{clad}}$ = core inlet coolant temperature (K).
- $T_c$ = core outlet temperature (K).
- $T_{\text{clad}}$ = average cladding temperature (K).
- $\dot{m}$ = reactor core coolant flow rate (kg/s).
- $C_c$ = heat capacity of the coolant (J/K).
- $C_f$ = heat capacity of the fuel (J/K).
- $C_{\text{clad}}$ = heat capacity of the fuel cladding (J/K).
- $(UA)_f$ = overall thermal transfer between the fuel and the cladding (J/K).
- $(UA)_{\text{clad}}$ = overall thermal transfer between the fuel cladding and the reactor core coolant (J/K).

**Thermal Model of Primary Heat Exchanger**

The thermal model of the primary heat exchanger is fully coupled with the thermoelectric model. The heat removal term in the thermal model of the primary heat exchanger depends on the hot shoe temperature of the thermal-electric device, and the heat addition term used in the thermoelectric model depends on the fluid temperature of the primary heat exchanger.

Although there are 30 channels, each of which has 480 TE cells mounted on the channel surface, only one node is used to describe the heat removal for simplification. With this simplification, the energy balance equation for the primary heat exchanger is given as follows:
\[
\begin{align*}
C_{\text{pHx}} \frac{dT_{\text{heat}}(t)}{dt} &= \frac{1}{N_{\text{Hx}}} \dot{m}_p c_p (T_{\text{pout}} - T_{\text{pin}}) - N_{\text{TE}} (UA)_{p\text{TE}} \\
(T_{\text{pin}} - T_{\text{Hshoe}}) \\
C_{\text{sHx}} \frac{dT_{\text{sout}}(t)}{dt} &= \frac{1}{N_{\text{Hx}}} \dot{m}_s c_p (T_{\text{sin}} - T_{\text{sout}}) + N_{\text{TE}} (UA)_{s\text{TE}} \\
(T_{\text{Cshoe}} - T_{\text{sout}})
\end{align*}
\]

where

\( T_{\text{pout}} \) = outlet temperature of the heat exchanger on the primary side (K).

\( T_{\text{sout}} \) = outlet temperature of the heat exchanger on the secondary side (K).

\( T_{\text{Hshoe}} \) = hot shoe temperature of a representative TE cell (K).

\( T_{\text{Cshoe}} \) = cold shoe temperature of a representative TE cell (K).

\( C_{\text{pHx}} \) = heat capacity of the primary heat exchanger on the primary side (J/K).

\( C_{\text{sHx}} \) = heat capacity of the primary heat exchanger on the secondary side (J/K).

\( N_{\text{TE}} \) = number of TE cells of each heat exchanger.

\( N_{\text{Hx}} \) = number of heat exchangers.

\( \dot{m}_p \) = primary flow rate of each heat exchanger (kg/s).

\( \dot{m}_s \) = secondary flow rate of each heat exchanger (kg/s).

\( (UA)_{p\text{TE}} \) = overall thermal transfer of a TE cell between the hot shoe material and the bulk fluid on the primary side (J/K).

\( (UA)_{s\text{TE}} \) = overall thermal transfer of a TE cell between the cold shoe material and the bulk fluid on the secondary side (J/K).

The overall heat transfer is calculated as follows:

\[
\frac{1}{(UA)_{\text{TE}}} = \frac{1}{\left( \sum_{i=1}^{N} \frac{t_i}{k_i} + \frac{1}{h_{\text{hx}}} \right) A_{\text{TE}}}
\]

where

\( A_{\text{Hx}} \) = area of a TE cell for convective heat transfer.

\( t_i \) = thickness of the i-th layer of the TE padding material.

\( k_i \) = heat conductivity of the i-th layer of the TE padding materials.
The forced convection heat transfer coefficient between the fluid and the channel walls is calculated based on the following correlation:

\[ h_{\text{hx}} = \frac{\text{Nu} \cdot k}{d_e} \]

where

\[ \text{Nu} = \frac{1}{\gamma} (5.8 + 0.02 Pe^{0.8}) \]

\( \gamma \) = the ratio of heat transfer coefficient of liquid metals flowing between two parallel plates with one of them being adiabatic to that with two plates of equal heat flux.

**Thermal-Electric Model**

In the SP-100 design, electric power is generated by 3 loops × 12 primary heat exchangers × 30 channels × 480 TE cells. Each TE cell consists of two semiconductors, one P-type semiconductor and one N-type semiconductor. The hot shoes of TE cells are mounted on the two surface plates of the 12 primary heat exchangers. The cold shoes of TE cells are maintained at a lower temperature by cooling effects of the secondary lithium loop, which is connected to radiators through heat pipes. Because there is a temperature gradient between the hot shoes and the cold shoes of TE cells, when heat is conducted from the hot shoe of a TE cell to its cold shoe, electric power will be generated due to the Seebeck effect.

In the case that physical properties are independent of the temperature, Thomson’s effect can be ignored. If it is further assumed that the P-type semiconductors and the N-type semiconductors have the same physical properties and the hot shoe temperature and the cold shoe temperature are constant, the resulting temperature distribution based on the governing equation with the specified boundary conditions has the following solution.

\[
K_{\text{TE}} (T_{\text{Hshoe}} - T_{\text{Cshoe}}) + \alpha_{\text{pn}} T_{\text{Hshoe}} I - 0.5 I^2 (R_p + R_N) = (UA)_{\text{TE}} (T_{\text{pout}} - T_{\text{Hshoe}})
\]

\[
K_{\text{TE}} (T_{\text{Hshoe}} - T_{\text{Cshoe}}) + \alpha_{\text{pn}} T_{\text{Cshoe}} I + 0.5 I^2 (R_p + R_N) = (UA)_{\text{TE}} (T_{\text{Cshoe}} - T_{\text{Sout}})
\]
\[ I = \frac{\alpha_{pn} (T_{Hshoe} - T_{Cshoe})}{R_N + R_P + R_L} \]

where

\( K_{n} \) = effective thermal conductance of a TE element (J/K).

\( \alpha_{pn} \) = relative Seebeck coefficient (V/K).

\( I \) = electric current (amp).

\( R_L \) = electric resistance of the shunt resistor (Ohm).

\( R_P \) = internal electric resistance of the P-type semiconductor (Ohm).

\( R_N \) = internal electric resistance of the N-type semiconductor (Ohm).

The electric power output, \( P_e \), is given as follows:

\[ P_e = I^2 R_L \]

The thermal power flowing into the TE generator, \( q_{TE} \), is given as follows:

\[ q_{TE} = (UA)_{pTE} (T_{pout} - T_{Hshoe}) \]

The efficiency of thermal electric conversion, \( \eta \), is defined by

\[ \eta = \frac{P_e}{q_{TE}} \]

**Radiator Model**

A single node is used to describe the heat removal by the radiators. It is assumed that the characteristic temperature of the coolant in the radiator is equal to the outlet temperature. The time dependent temperature of radiator, \( T_{in} \), is given by:

\[ C_{prad} \frac{dT_{in}(t)}{dt} = \frac{1}{N_{Hx}} \dot{m}_s c_p (T_{sout} - T_{in}) - \varepsilon \sigma F_{rad} A_{TE} (T_{in}^4 - T_a^4) \]
where
\[ \varepsilon = \text{the emissivity}. \]
\[ F_{\text{rat}} = \text{the ratio of the surface area of the radiator to that of the TE cell cross section}. \]
\[ \sigma = \text{the Stephan-Boltzman constant (Wm}^{-2}\text{K}^{-4}). \]
\[ T_a = \text{the ambient temperature in space (K)}. \]

Figure 3.20 shows the step response of the increase in control drum angle and Figure 3.21 shows the step response of the increase in external load.

Figure 3.20. Step response of the increase in control drum angle.

Figure 3.21. Step response of the increase in external load.
3.8. Coordinated Control of the SP-100 Space Reactor

The basic control objective of SP-100 system is to follow the power demand with the best TE thermal-electric conversion efficiency without violating the physical constraints of the system. Based on the results of transient modeling, electric power output can be regulated either by manipulating the external load or by manipulating the reactivity. However, the two control strategies have their own limitations.

Figure 3.22 shows the performance of the proportional-integral (PI) control of the TE electric power by manipulating the external load only. When the electric power demand decreases from 100% full power to 50% full power, the PI controller is able to follow the power demand almost immediately by reducing the external load. However, at the 300th second when power demand increases by 60% full power, increasing the external load cannot follow the power demand. When external load increases to a certain value, the produced electric power cannot increase any longer. For this reason, the system becomes unstable.

Figure 3.23 shows the TE electric power output as a function of the TE hot shoe temperature and external load. It is clear from the figure that the TE electric power increases while the reactor power as well as the primary heat exchanger inlet temperature increase. However, as external load increases, the electric power output would decrease rather than increase if the external load exceeds a certain value.

Figure 3.22. Proportional-integral control of electric power based on manipulating the external load.
Figure 3.23. TE electric power versus external load and TE hot shoe temperature.

Figure 3.24 shows the performance of PI controller of the TE electric power by manipulating the reactivity. When the electric power demand decreases from 100% full power to 50% full power, the PI controller introduces negative reactivity by manipulating the control drum angle. As compared with Figure 3.3, it is observed that the electric power responds at a slower speed. This is because the thermal inertia of reactor coolant system results in a time lag for the response of electric power output. As the power demand increases by 60% full power once again, the electric power demand can be followed, though at a slower speed, in this case. According to the above analysis, an optimal controller of electric power is to have a coordinated control on the external load and the reactivity. Figure 3.25 shows the results of the coordinated PI control. In this control strategy, when the error signal of electric power is small, external load is manipulated to follow the power demand. If the error signal is large, the reactivity is manipulated to follow the power demand. Because the two control modes are coordinated, both the optimal speed of load following and system stability can be ensured. Figure 3.25 clearly shows the expected results. The coordinated control can ensure stability when the power demand increases up to 110% full power and has a quick response to changes in power demand.
Figure 3.24. Proportional-Integral (PI) control of electric power based on manipulating the reactivity.

Figure 3.25. Coordinated control of electric power.
4. MODEL PREDICTIVE CONTROL (MPC)

4.1. Introduction

In order to optimize the reactor power control performance, techniques for the optimal power control of nuclear reactors have been studied extensively in the past two decades (Cho and Grossman, 1983, Niar and Gopal, 1987, Lin et al., 1986, Park and Cho, 1993, Shtessel, 1998). But it is very difficult to design optimized controllers for nuclear systems because of variations in nuclear system parameters and modeling uncertainties, and in particular, for the long-term operation of the SP-100 reactor.

The model predictive control methodology has received increased attention as a powerful tool for the control of industrial process systems (Kwon and Pearson, 1977, Richalet et al., 1978, Garcia et al., 1989, Clarke and Scattolini, 1991, Kothare at al., 1996, Lee at al., 1997, Lee et al., 1998). The basic concept of the model predictive control is to solve an optimization problem for a finite future at current time. Once a future input trajectory has been chosen, only the first element of that trajectory is applied as the input to the plant and the calculation is repeated at each subsequent instant. This method has many advantages over the conventional infinite horizon control because it is possible to handle input and state (or output) constraints in a systematic manner during the design and implementation of the control. In particular, it is a suitable control strategy for nonlinear time varying systems because of the model predictive concept and recently, the problem of controlling uncertain dynamical systems has been of considerable interest to control engineers. The model predictive control method has been applied to a nuclear engineering problem (Na, 2001, Na et al., 2003).

In this work, a model predictive control (MPC) method is applied to design an automatic controller for thermoelectric (TE) generator power level for an SP-100 space reactor. The SP-100 reactor core dynamics is identified online by a recursive least squares method. Based on this identified reactor model, consisting of the control drum angle and the TE generator power, the future TE generator power is predicted. The objective function for the model predictive control is minimized by a genetic algorithm that is widely used for optimization problems. A lumped parameter simulation model of the SP-100 space reactor is used to verify the proposed controller for a space nuclear reactor.
4.2. The Model-Predictive Control (MPC) Method

4.2.1. The MPC Concept

Figure 4.1 shows the basic concept of the model predictive control method. For any assumed set of present and future control moves, the future behavior of the process outputs can be predicted over a prediction horizon $N$, and the $M$ present and future control moves ($M \leq N$) are computed to minimize a quadratic objective function. Although $M$ control moves are calculated, only the first control move is implemented. At the next time step, new values of the measured output are obtained, the control horizon is shifted forward by one step, and the same calculations are repeated. The purpose of taking new measurements at each time step is to compensate for unmeasured disturbances and model inaccuracies, both of which cause the measured system output to be different from the one predicted by the model. At every time instant, the model predictive control requires the on-line solution of an optimization problem to compute optimal control inputs over a fixed number of future time instants, known as the time horizon. The basic idea of model predictive control is to calculate a sequence of future control signals in such a way that it minimizes a multistage cost function defined over a prediction horizon.

![Figure 4.1. Basic concept of a model predictive control method (Garcia et al., 1989).](image)
Also, in order to achieve fast responses and prevent excessive control effort, the associated performance index for deriving an optimal control input is represented by the following quadratic objective function:

\[
J = \frac{1}{2} \sum_{j=1}^{N} Q \left[ \hat{y}(t+j|t) - w(t+j) \right]^2 + \frac{1}{2} \sum_{j=1}^{M} R [\Delta u(t+j-1)]^2 ,
\]

subject to constraints:

\[
\Delta u(t+j-1) = 0 \quad \text{for} \quad j > M ,
\]

\[
\begin{align*}
\min & u & \max \\
\Delta & \leq \Delta & \Delta & \leq \Delta
\end{align*}
\]

where \( Q \) and \( R \) weight the TE generator power (system output) error and control drum angle (control input) change between time steps at certain future time intervals, respectively, and \( w \) is a set point (desired TE generator power). The estimate \( \hat{y}(t+j|t) \) is an optimum \( j \)-step-ahead prediction of the system output (TE generator power) based on data up to time \( t \); that is, the expected value of the output at time \( t \) if the past input and output and the future control sequence are known. \( N \) and \( M \) are called the prediction horizon and the control horizon, respectively. The prediction horizon represents the limit of the instant in which it is desired for the output to follow the reference sequence. In order to obtain control inputs, the predicted outputs have to be first calculated as a function of past values of inputs and outputs and of future control signals. The constraint, \( \Delta u(t+j-1) = 0 \) for \( j > M \), means that there is no variation in the control signals after a certain interval \( M < N \), which is the control horizon concept. \( u_{\min} \) and \( u_{\max} \) are the minimum and maximum values of input, respectively, and \( \Delta u_{\max} \) is a maximum allowable control move per time step.

4.2.2. Output Prediction

The process to be controlled is described by the following Controlled Auto-Regressive and Integrated Moving Average (CARIMA) model, which is widely used as a mathematical model for controller design:

\[
A(q^{-1}) y(t) = B(q^{-1}) \Delta u(t-1) + C(q^{-1}) \xi(t) ,
\]

where \( y \) is an output (TE generator power), \( u \) is a control input (control drum angle to regulate external reactivity), \( \xi \) is a stochastic random noise sequence with zero mean value and finite
variance, $q^{-1}$ is the backward shift operator, e.g., $q^{-1}y(t) = y(t-1)$, and $\Delta$ is defined as $\Delta = 1 - q^{-1}$. In Equation (4.2), $A(q^{-1})$ and $C(q^{-1})$ are monic polynomials as a function of the backward shift operator $q^{-1}$, and $B(q^{-1})$ is a polynomial of order $nB$.

In order to optimize the objective function, the output must be predicted first. The process output at time $t + j$ can be predicted from the measurement values of the output and input up to time step $t$. The optimal prediction is derived by solving a Diophantine equation whose solution can be found by an efficient recursive algorithm. In this derivation, the most usual case of $C(q^{-1}) = 1$ will be considered. By multiplying Equation (4.2) by $\Delta E_j(q^{-1})$ from the left and applying the Diophantine Equation (4.4) to the CARIMA model of Equation (4.2), the $j$-step-ahead output prediction of a process has the following form.

$$y(t + j) - E_j(q^{-1})\xi(t + j) = F_j(q^{-1})y(t) + E_j(q^{-1})B(q^{-1})\Delta u(t + j - 1), \quad (4.3)$$

where $E_j(q^{-1})$ and $F_j(q^{-1})$ are polynomials satisfying

$$1 = E_j(q^{-1})A(q^{-1}) + q^{-j}F_j(q^{-1}), \quad (4.4)$$

$$E_j(q^{-1}) = e_{j,0} + e_{j,1}q^{-1} + \cdots + e_{j,j}q^{-j}, \quad (4.5)$$

$$F_j(q^{-1}) = f_{j,0} + f_{j,1}q^{-1} + f_{j,2}q^{-2} + \cdots + f_{j,nA}q^{-nA}, \quad (4.6)$$

There exist unique polynomials $E_j(q^{-1})$ and $F_j(q^{-1})$ of orders $j$ and $nA$, respectively, such that $e_{j,0} = 1$. By taking the expectation operator and considering that $E\{\xi(t)\} = 0$, the optimal $j$-step-ahead prediction of $\hat{y}(t + j | t)$ satisfies

$$\hat{y}(t + j | t) = F_j(q^{-1})y(t) + G_j(q^{-1})\Delta u(t + j - 1), \quad (4.7)$$

where

$$G_j(q^{-1}) = E_j(q^{-1})B(q^{-1}),$$

$$\hat{y}(t + j | t) = E\{y(t + j) - E_j(q^{-1})\xi(t + j) | t\} = E\{y(t + j) | t\}. $$
\( \hat{y}(t+j|t) \) denotes an estimated value of the output at time step \( t+j \) based on all the data up to time step \( t \). The output prediction can be easily extended to the nonzero mean noise case by adding a term \( E_j(q^{-1})E\{\xi(t)\} \) to the output prediction \( \hat{y}(t+j|t) \).

By dividing the matrix polynomial, \( G_j(q^{-1}) \), into two terms as in the following equation

\[
G_j(q^{-1}) = \tilde{G}_j(q^{-1}) + q^{-j}\tilde{G}_j(q^{-1}) \quad \text{with} \quad \deg(\tilde{G}_j(q^{-1})) < j,
\]

the prediction equation, Equation (3.7), can now be rewritten as

\[
\hat{y}(t+j|t) = \tilde{G}_j(q^{-1})\Delta u(t+j-1) + \tilde{G}_j(q^{-1})\Delta u(t-1) + F_j(q^{-1})y(t), \quad (4.8)
\]

where \( \deg(\cdot) \) denotes the order of a polynomial. The last two terms of the right hand side of Equation (4.8) consist of past values of the process input and output variables and correspond to the response of the process if the control input signals are kept constant. On the other hand, the first term of the right hand side consists of future values of the control input signal and corresponds to the response obtained when the initial conditions are zero, that is, \( y(t-j) = 0 \) and \( \Delta u(t-j-1) = 0 \) for \( j > 0 \) (Camacho and Bordons, 1999).

A set of \( N \) \( j \)-step-ahead output predictions can be expressed as

\[
\hat{y} = \tilde{G}\Delta u + f, \quad (4.9)
\]

where

\[
\hat{y} = [\hat{y}(t+1|t) \quad \hat{y}(t+2|t) \quad \cdots \quad \hat{y}(t+j|t) \quad \cdots \quad \hat{y}(t+N|t)]^T,
\]

\[
\Delta u = [\Delta u(t) \quad \Delta u(t+1) \quad \cdots \quad \Delta u(t+j) \quad \cdots \quad \Delta u(t+N-1)]^T,
\]

\[
f = [f_1 \quad f_2 \quad \cdots \quad f_j \quad \cdots \quad f_N]^T,
\]

\[
f_j = \tilde{G}_j(q^{-1})\Delta u(t-1) + F_j(q^{-1})y(t),
\]

\[
\tilde{G} = \begin{bmatrix}
g_0 & 0 & \cdots & 0 & \cdots & 0 \\
g_1 & g_0 & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
g_{j-1} & g_{j-2} & \cdots & g_0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
g_{N-1} & g_{N-2} & \cdots & \cdots & \cdots & g_0
\end{bmatrix}.
\]
\[
\bar{G}_j(q^{-1}) = \sum_{i=0}^{j-1} g_i q^{-i}.
\]

If all initial conditions are zero, the response \( f \) is zero. If a unit step is applied to the first input at time \( t \), that is, \( \Delta u = [1 \ 0 \ \cdots \ 0]^T \), the expected output sequence \( [\hat{y}(t+1) \ \hat{y}(t+2) \ \cdots \ \hat{y}(t+N)]^T \) is equal to the first column of the matrix \( \bar{G} \). That is, the first column of the matrix \( \bar{G} \) can be calculated as the step response of the plant when a unit step is applied to the first control signal.

Since the control signal is kept constant after the first \( M \) control moves (that is, \( \Delta u(t+j-1) = 0 \) for \( j > M \)) due to the model predictive control concept, the set of predictions affecting the objective function can be expressed as

\[
\hat{y} = \bar{G}_s \Delta u_s + f, \quad (4.10)
\]

where

\[
\bar{G}_s = \begin{bmatrix}
g_0 & 0 & \cdots & 0 \\
g_1 & g_0 & \cdots & 0 \\
& \ddots & \ddots & \ddots \\
& & \ddots & g_0 \\
g_{N-1} & g_{N-2} & \cdots & g_{N-M}
\end{bmatrix},
\]

\[
\Delta u_s = [\Delta u(t) \ \Delta u(t+1) \ \cdots \ \Delta u(t + M - 1)]^T.
\]

The objective function of Equation (3.1), including the summation form (\( \Sigma \)), can be rewritten in the following matrix-vector form:

\[
J = \frac{1}{2} (\hat{y} - w)^T \bar{Q}(\hat{y} - w) + \frac{1}{2} \Delta u_s^T \bar{R} \Delta u_s, \quad (4.11)
\]

subject to constraints

\[
\begin{aligned}
\Delta u(t+j-1) = 0 \quad \text{for} \quad j > M, \\
u_{\min} \leq u(t) \leq u_{\max}, \\
|\Delta u(t)| \leq \Delta u_{\max}.
\end{aligned}
\]

(4.12)

where \( \bar{Q} = \text{diag} \{Q, \cdots, Q\} \) is a diagonal matrix consisting of \( N \) diagonal elements, \( Q \), which usually has a value of 1 (\( Q = 1 \)) and \( \bar{R} = \text{diag} \{R, \cdots, R\} \) is a diagonal matrix consisting of \( M \) diagonal elements, \( R \), which is called an input-weighting factor.
4.2.3. Recursive Parameter Estimation

The process model is estimated recursively at every time step to reflect time-varying conditions of the plant including fuel burn-up, parameter changes due to long-term operation, and others. Equation (4.7) can be expressed as the following inner product of the parameter vector $\hat{\theta}(t)$ and the measurement vector $\phi(t)$:

$$\hat{y}(t+1) = F_1(q^{-1})y(t) + G_1(q^{-1})\Delta u(t) = \hat{\theta}^T(t) \cdot \phi(t), \quad (4.13)$$

where

$$\hat{\theta}^T(t) = \left[ \hat{a}_1(t) \quad \hat{a}_2(t) \quad \cdots \quad \hat{a}_n(t) \quad \hat{b}_0(t) \quad \hat{b}_1(t) \quad \cdots \quad \hat{b}_{nB}(t) \right],$$

$$\phi^T(t) = \left[ -y(t) \quad -y(t-1) \quad \cdots \quad -y(t-nA+1) \quad \Delta u(t) \quad \Delta u(t-1) \quad \cdots \quad \Delta u(t-nB) \right].$$

The parameter vector $\hat{\theta}(t)$ is estimated using a recursive least-squares method as follows:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \Sigma(t)\phi(t-1)\left[ y(t) - \hat{\theta}^T(t-1) \cdot \phi(t-1) \right], \quad (4.14)$$

$$\Sigma(t) = \frac{1}{\lambda(t)} \left[ \Sigma(t-1) - \frac{\Sigma(t-1)\phi(t-1)\phi^T(t-1)\Sigma(t-1)}{\lambda(t) + \phi^T(t-1)\Sigma(t-1)\phi(t-1)} \right], \quad (4.15)$$

where the covariance matrix $\Sigma(0) > 0$ and $0 < \lambda(t) \leq 1$. A forgetting factor $\lambda(t)$ is usually used to account for an exponential decay of the past data in tracking a slow drift in parameters and is calculated from the following equation:

$$\lambda(t) = \lambda_0 \lambda(t-1) + (1 - \lambda_0) \quad \text{with} \quad \lambda_0 \leq 1 \quad \text{and} \quad \lambda(0) \leq 1. \quad (4.16)$$

4.3. Optimization of the Objective Function by a Genetic Algorithm (GA)

The objective function of Equation (3.11) can be solved by linear matrix inequality (LMI) techniques. In this work, a genetic algorithm is used to minimize the objective function with some constraints. The genetic algorithm has been known to be proper in solving multiple objective functions. Compared to the conventional optimization methods that move from one point to another, genetic algorithms start from many points simultaneously climbing many peaks in parallel. Accordingly, genetic algorithms are less susceptible to being stuck at local minima compared to conventional search methods (Goldberg, 1989; Mitchell, 1996).
In the genetic algorithm, the term chromosome refers to a candidate solution that minimizes an objective function. As the generation proceeds, populations of chromosomes are iteratively altered by biological mechanisms inspired by natural evolution such as selection, crossover, and mutation. The genetic algorithms require a fitness function that assigns a score to each chromosome (candidate solution) in the current population, and maximize the fitness function value. The fitness function evaluates the extent to which each candidate solution is suitable for specified objectives. The genetic algorithm starts with an initial population of chromosomes, which represent possible solutions of the optimization problem. The fitness function is computed for each chromosome. New generations are produced by the genetic operators that are known as selection, crossover, and mutation. The algorithm stops after the maximum allowed time has elapsed.

In this research, the optimization problem is to solve $M$ present and future control signals to minimize the objective function. Therefore, a chromosome is represented by $s_l$ whose elements consist of present and future control inputs and has the following structure:

$$s_l = [u_l(t) \quad u_l(t+1) \quad \cdots \quad u_l(t+M-1)]. \quad (4.17)$$

The subscript $l$ indicates that it is related to the $l$-th chromosome ($l = 1, \ldots, L$ where $L$ is the number of chromosomes). Chromosomes will constitute the initial population, the crossover probability $P_c$, and the mutation probability $P_m$. The genetic algorithm proceeds according to the following steps (Sarimveis and Bafas, 2003):

Step 1 (initial population): Set the number of iterations $iter = 1$. Generate an initial population consisting of $L$ chromosomes of Equation (4.17). The values are allocated randomly, but they should satisfy both input and input move constraints of Equation (4.12). For this purpose, we use a simple procedure as follows:

(a) Read the measured value of the input variable at the previous time point $t-1$, which has already been implemented.

(b) Select the current input value using the following equations:

$$u_l(t) = u(t-1) + r_l \cdot \Delta u_{\text{max}}. \quad (4.18)$$

If $u_l(t) \geq u_{\text{max}}$, set $u_l(t) = u_{\text{max}}. \quad (4.19)$

If $u_l(t) \leq u_{\text{min}}$, set $u_l(t) = u_{\text{min}}. \quad (4.20)$
(c) Select the rest of the input moves using the following equations:

\[ u_i(t+i) = u_i(t+i-1) + r_i \cdot \Delta u_{\text{max}}, \quad 1 \leq i \leq M - 1. \]  \hspace{1cm} (4.21)

If \( u_i(t+i) \geq u_{\text{max}} \), set \( u_i(t+i) = u_{\text{max}}, \) \( 1 \leq i \leq M - 1. \)  \hspace{1cm} (4.22)

If \( u_i(t+i) \leq u_{\text{min}} \), set \( u_i(t+i) = u_{\text{min}}, \) \( 1 \leq i \leq M - 1. \)  \hspace{1cm} (4.23)

In the above equations, \( r_i \) is a random number in the range of \([-1,1]\). A new random number \( r_i \) is generated whenever Equation (4.18) or Equation (4.21) is used.

Step 2 (fitness function evaluation): Evaluate the objective function of Equation (4.11) for all the chosen chromosomes. Then invert the objective function values and find the total fitness of the population as follows:

\[ F = \sum_{l=1}^{L} \frac{1}{J_l(t)}, \]  \hspace{1cm} (4.24)

where \( J_l(t) \) is the objective function value for the \( l \)-th chromosome and the inversion of \( J_l(t) \) is a fitness value of the \( l \)-th chromosome. Then, calculate the normalized fitness value of each chromosome, meaning that the selection of probability \( p_l \) calculated by

\[ p_l = \frac{(1/J_l(t))}{F}, \quad 1 \leq l \leq L. \]  \hspace{1cm} (4.25)

Step 3 (selection operation): Calculate the cumulative probability \( q_l \) for each chromosome using the following equation:

\[ q_l = \sum_{n=1}^{l} p_n, \quad 1 \leq l \leq L. \]  \hspace{1cm} (4.26)

For \( l = 1, \ldots, L \), generate a random number \( r_2 \) between 0 and 1. Select the chromosome for which \( q_{l-1} \leq r_2 \leq q_l \). At this point of the algorithm a new population of chromosomes has been generated. The chromosomes with high fitness value have more chance to be selected.

Step 4 (crossover operation): For each chromosome \( s_i \), generate a random number \( r_3 \) between 0 and 1. If \( r_3 \) is lower than \( p_c \), this particular chromosome will undergo the process of crossover, otherwise it will remain unchanged. Mate the selected chromosomes and for each selected pair
generate a random integer number \( z \) between 0 and \( M - 1 \). The crossing point is the position indicated by the random number. Two new chromosomes are produced by interchanging all the members of the parents following the crossing point. Graphically, the crossover operation can be represented as shown below, assuming that the crossover operation is applied to the parent chromosomes \( s_i \) and \( s_{i+1} \):

\[
s_i = \begin{bmatrix}
 u_i(t) & u_i(t+1) & \cdots & u_i(t+z-1) & u_i(t+z) & \cdots & u_i(t+M-1)
\end{bmatrix}
\]

\[
s_{i+1} = \begin{bmatrix}
 u_{i+1}(t) & u_{i+1}(t+1) & \cdots & u_{i+1}(t+z-1) & u_{i+1}(t+z) & \cdots & u_{i+1}(t+M-1)
\end{bmatrix}
\]

↓ crossover operation

\[
s_i^{\text{new}} = \begin{bmatrix}
 u_i(t) & u_i(t+1) & \cdots & u_i(t+z-1) & u_i(t+z) & \cdots & u_i(t+M-1)
\end{bmatrix}
\]

\[
s_{i+1}^{\text{new}} = \begin{bmatrix}
 u_{i+1}(t) & u_{i+1}(t+1) & \cdots & u_{i+1}(t+z-1) & u_{i+1}(t+z) & \cdots & u_{i+1}(t+M-1)
\end{bmatrix}
\]

The above operation might produce an infeasible offspring if the input values at the cross point do not satisfy the input move constraints. This situation is avoided by the following correction mechanism for an input variable, which modifies the values of the input parameters after the cross position so that the input move constraints are satisfied. At first, for one of the produced chromosomes \( s_i^{\text{new}} \),

If \( u_{i+1}(t+z) - u_i(t+z-1) > \Delta u_{\text{max}} \),

then

\[
\Delta = u_{i+1}(t+z) - u_i(t+z-1) - \Delta u_{\text{max}},
\]

\[
u_{i+1}(t+z+i) = u_{i+1}(t+z+i) - \Delta, \quad 0 \leq i \leq M - 1 - z.
\]

If

\[
u_{i+1}(t+z) - u_i(t+z-1) < -\Delta u_{\text{max}},
\]

then

\[
\Delta = u_i(t+z-1) - u_{i+1}(t+z) - \Delta u_{\text{max}},
\]

\[
u_{i+1}(t+z+i) = u_{i+1}(t+z+i) + \Delta, \quad 0 \leq i \leq M - 1 - z.
\]

A similar set can be written for the chromosome \( s_{i+1}^{\text{new}} \).
Step 5 (mutation operation): For every member of each chromosome $s_j$, $u_i(t+i)$, generate a random number $r_4$ between 0 and 1. If $r_4$ is lower than $p_m$, this particular member of the chromosome will undergo the process of mutation, otherwise it will remain unchanged. For the selected members define upper and lower bounds as follows:

$$b_u = \min(\Delta u_{\text{max}} + u(t-1), \Delta u_{\text{max}} + u_i(t+i+1), u_{\text{max}})$$
$$b_l = \max(-\Delta u_{\text{max}} + u(t-1), -\Delta u_{\text{max}} + u_i(t+i+1), u_{\text{min}})$$

if $i = 0$, \hspace{1cm} (4.33)

$$b_u = \min(\Delta u_{\text{max}} + u_i(t+i-1), \Delta u_{\text{max}} + u_i(t+i+1), u_{\text{max}})$$
$$b_l = \max(-\Delta u_{\text{max}} + u_i(t+i-1), -\Delta u_{\text{max}} + u_i(t+i+1), u_{\text{min}})$$

if $0 < i < M-1$, \hspace{1cm} (3.34)

$$b_u = \min(\Delta u_{\text{max}} + u_i(t+i-1), u_{\text{max}})$$
$$b_l = \max(-\Delta u_{\text{max}} + u_i(t+i-1), u_{\text{min}})$$

if $i = M-1$. \hspace{1cm} (4.35)

The above bounds define the region of values of $u_i(t+i)$ which will produce a feasible solution. This definition is followed by the generation of a random binary number $b$. Based on the value of $b$, $u_i(t+i)$ is modified by the following equations:

$$u_i(t+i) = u_i(t+i) + (b_u - u_i(t+i))(1 - r_5^{1-(\text{iter}/\text{iter}_{\text{max}})})$$

if $b = 0$, \hspace{1cm} (4.36)

$$u_i(t+i) = u_i(t+i) - (u_i(t+i) - b_l)(1 - r_5^{1-(\text{iter}/\text{iter}_{\text{max}})})$$

if $b = 1$. \hspace{1cm} (4.37)

where $r_5$ is a random number between 0 and 1, $\text{iter}$ is the number of iterations performed so far, and $\text{iter}_{\text{max}}$ is the expected final number of iterations.

Step 6 (repeat or stop): If the maximum allowed time has not expired, set $\text{iter} = \text{iter} + 1$ and return Step 2. Otherwise, stop the algorithm and select the chromosome that produced the lowest value of the objective function throughout the entire procedure.

### 4.4. Application of the MPC Method to the SP-100 Space Reactor

Figure 4.2 shows the schematic block diagram of the model predictive controller combined with a parameter estimation algorithm. In this work, the MPC power controller was applied to the SP-100 reactor model. The lumped parameter model for a SP-100 space reactor and the proposed control algorithm are written in MATLAB.
At first, a SP-100 space reactor power controller was designed by using the model predictive control optimized by a genetic algorithm.

The model predictive controller for the power level control is subject to constraints as follows:

\[ \Delta u(t + j - 1) = 0 \quad \text{for} \quad j > M, \]

\[ 0^\circ \leq u(t) \leq 180^\circ, \]

\[ |\Delta u(t)| \leq 1.4 T^\circ, \]

where \( T \) is a sampling time. The external reactivity control uses the mechanism of the stepper motor control drum system (Shtessel, 1998). The control drum angle of the stepper motor shaft can be rotated from 0 to 180 degrees. The maximum angular velocity of the drums is \( 1.4^\circ / \text{sec} \). The optimal control input could be obtained by solving the minimization objective function of Equation (4.11) using a genetic algorithm.

Most of the computation time is involved in the calculation of the reactor dynamics, and that of the controller is insignificant. The sampling time is 1 second. One-step calculation time of the controller is about 0.08 sec on a 3.0 GHz PC. Therefore, it is possible to accomplish real time performance even under low computing environments.
Figure 4.3 shows the simulation results of the proposed controller according to the input-weighting factor. The desired power is 100% initially and decreases to 50% by ramp from 100sec and increases to 100% by ramp from 300sec. Also, it decreases from 100% to 50% by step at 500sec and increases from 50% to 100% by step at 700sec. It is shown that the TE generator power and the nuclear reactor power level follow their desired value very well although their response speed depends on the input weighting factor which is a tuning variable. The input-weighting factor is expressed by the following equation:

\[ R = \omega \left( \frac{y_o}{u_o} \right)^2 \]

where \( y_o \) is a rated value of the TE generator power and \( u_o \) is a rated value of control drum angle.

Figure 4.4 shows the performance of the proposed model predictive controller. Figures 4.4(g) and 4.4(h) show some parameters related to \( A(q^{-1}) \) and \( B(q^{-1}) \) that are estimated recursively at every time step. It is assumed that all the parameters of the polynomials \( A(q^{-1}) \) and \( B(q^{-1}) \) are 0.1 at the beginning of simulation. In this work, these assumed values are estimated by the parameter estimation algorithm for initial 50 sec; this means 50 time steps since the sampling period is 1 second. This is done by exciting the core dynamics with initial small random movements of the stepper motor shaft that changes the control drum angle. Subsequently, the parameters are continuously adapted according to the changing operating conditions of the drum angle and the reactor power. Figures 4.4(g) and 4.4(h) show that the reactor dynamics changes according to the power level, the control drum angle, and other effects. Figure 4.4(i) shows the trend of the best fitness function value, which is affected by the magnitudes of the estimated output error and the control input move.
Figure 4.3. Performance of an MPC controller according to an input weighting factor.
(a) TE power and thermal reactor power

(b) control drum angle

(c) reactivity

(d) external resistance

(e) TE efficiency

(f) temperatures

(g) $A(q^{-1})$ polynomial

(h) $B(q^{-1})$ polynomial
(i) best fitness value of the genetic algorithm

Figure 4.4. Performance of the proposed MPC controller \( (\omega = 300) \).

In addition, a conventional proportional-integral (PI) controller was designed to compare the performance of the power level response with the proposed model predictive controller optimized by the genetic algorithm. Figure 4.5 shows the performance of a well-tuned PI controller. Its response is a little slow. As shown in Figure 4.5, the PI controller performs less effectively than the proposed model predictive controller. Also, if nonlinear characteristics are strong because of nuclear fuel burn-up and other effects, that are not considered in this work, it is expected that the proposed model predictive controller will have much better performance than the PI controller, because the MPC is optimized at each time step.

Figure 4.6 shows the performance of a proposed MPC controller according to a parameter change. It is assumed that the fuel expansion feedback coefficient changes as shown in Figure 4.6(d). The proposed controller well controls the nuclear reactor even under the significant change of the parameter.

A constrained minimization version of the model-predictive control and its application to the SP-100 reactor control is given in Appendix A.
(a) TE power and thermal reactor power  
(b) control drum angle  
(c) reactivity  
(d) temperature  

Figure 4.5. Performance of a well-tuned proportional-integral (PI) controller.
Figure 4.6. Performance of the proposed MPC controller ($\omega = 300$) under the condition of changes in the fuel expansion feedback coefficient, as shown in Fig. 4.6(d).
5. FAULT DETECTION AND ISOLATION (FDI) OF SENSORS AND FIELD DEVICES

5.1. Introduction

A successful implementation of hierarchical control with control mode reconfiguration depends on the reliability of sensor measurements and functional readiness of devices such as control actuators, pumps, and other process units. The University of Tennessee [9-11] has developed efficient data-based techniques for validating sensors and devices, and to distinguish between changes in the devices and in the process itself.

The characterization of the physical relationship among a set of process variables can be established using data-driven modeling techniques. These include linear principal component analysis [10], nonlinear group method of data handling [11], and others. In order to be effective for anomaly tracking, it is necessary to have a complete set of operational data. In the absence of such data, physics simulation may be combined with available plant data for building these predictive models. During the monitoring phase the models are used to predict the process variables of interest and establish the space of residuals or model errors. The time behavior, patterns, and other signatures derived from the model and the measurements are used to detect and isolate sensor and device anomalies.

In the hierarchical control scheme, this information is then fed to the supervisor and used to make decisions about the adequacy of the control actions and the need to reconfigure the control mode. This is the essence of autonomous control design. Principle Component Analysis (PCA) based approach is proposed as the baseline algorithm for steady state fault diagnosis because of its explicit representation of fault detectability and identifiability. It was applied to fault diagnosis of steam generators (Lu, 2002) and the a typical pressurized water reactor (Kaistha, 2001). In this project, the Principal Component Analysis (PCA) has been applied to monitor the steady state performance of the SP-100 reactor, using the MATLAB-Simulink model developed in Section 3.
5.2. Principal Component Analysis (PCA) for Measurement Modeling

The Principal Component Analysis (PCA) was originally developed by Pearson, 1901, as a statistical method of dimensional reduction while preserving the variation of data. The original data can be represented in a lower dimensional space without significant loss of the variability. From the modeling point of view, PCA transforms correlated variables into uncorrelated ones and determines the linear combinations with large and low variability (Flury, 1988).

Before the original data are transformed into a lower dimensional space, they are mean centered because only the variability of the data is of interest. The data are standardized with unit variance so that equal weights are given to all the variables as far as their variability is concerned.

For a measurement vector \( x \in \mathbb{R}^n \), in general, it can be represented in \( n \)-dimensional space as follows:

\[
x = \sum_{i=1}^{n} t_i p_i
\]

where

\( p_i \) = the basis vectors of the \( n \)-dimensional space, which are orthonormal; \( t_i \) = the component of the vector \( x \) when projected onto the basis vector \( p_i \). Since the basis vector are orthonormal, \( t_i \) may be written as

\[
t_i = x^T p_i
\]

If the true dimensionality of the measurement vector is \( l \), where \( l < n \), then the projection can be separated into two parts:

\[
x = \sum_{i=1}^{l} t_i p_i + \sum_{j=l+1}^{n} \tilde{t}_j \tilde{p}_j
\]

where the second term represents the random vectors obtained by projecting the vector \( x \) onto the remaining \( (n-l) \) dimensional space.
The objective of the PCA algorithm is to determine the true number of components directly from the measured data such that the second part of Equation (5.3) behaves purely random. This is equivalent to determining \( l \) principal components such that the squared error of the approximation, given by Equation (4.4) is minimized.

\[
\varepsilon^2 = E[\varepsilon^T \varepsilon] \quad (5.4)
\]

where \( \varepsilon \) is the measurement error. \( \varepsilon = x - \sum_{i=1}^{l} t_i p_i = \sum_{j=l+1}^{n} t_j p_j. \)

Given a data matrix \( X \) associated with \( n \) measured variables and \( m \) measurements when the mean values are removed, the first principal component is obtained by finding a basis vector \( p_1 \) such that the score vector \( t_1 \) of the original data along this direction has maximized variance, which is given by:

\[
\max \{ E[\text{var}(t_1)] \} = \max \frac{p_1^T \Sigma p_1}{p_1^T p_1} \quad (5.5)
\]

with the constraints \( p_1^T p_1 = 1 \), where \( \Sigma = X^T X \).

This solution of Equation (5.5) is the eigenvector of \( \Sigma \) corresponding to its maximum eigenvalue, that is, \( \Sigma p_1 = \lambda_1 p_1 \), where

\( \lambda_1 = \) the largest eigenvalue of the covariance matrix.

Using the same method, we can obtain the other \( j^{th} \) principal component by solving the eigenvalue problem of the scatter matrix \( \Sigma \), which is given by:

\[
\Sigma p_j = \lambda_j p_j \quad j = 1, 2, \ldots l
\]

This PCA algorithm has chosen the eigenvectors of the covariance matrix of the measured data as the basis vectors for projection. The implication of the projection in this manner is that the variation of the measured data can be separated into the variation in the principal component.
subspace and the variation remaining in the residual space. The PCA decomposition of the
original measurement data is given by

\[ X = TP^T + \tilde{P}\tilde{P}^T \]  

(5.6)

where

\[
P = \begin{bmatrix}
p_{11} & p_{12} & \cdots & p_{1l} \\
p_{21} & p_{22} & \cdots & p_{2l} \\
\vdots & \vdots & \ddots & \vdots \\
p_{n1} & p_{n2} & \cdots & p_{nl}
\end{bmatrix} \]

and

\[
T = \begin{bmatrix}
t_{11} & t_{12} & \cdots & t_{1l} \\
t_{21} & t_{22} & \cdots & t_{2l} \\
\vdots & \vdots & \ddots & \vdots \\
t_{m1} & t_{m2} & \cdots & t_{ml}
\end{bmatrix}
\]

\( P \) = the loading matrix whose columns span the principal component space (PCS) and consist of
the eigenvectors corresponding to the larger \( l \) eigenvalues of the matrix \( \Sigma \).

\( \tilde{P} \) = the loading matrix whose columns span the residual space (RS) and consist of the
eigenvectors corresponding to the smaller \( n-l \) eigenvalues of matrix \( \Sigma \).

For measurement vector \( x \), the PCA algorithm estimates the true value by a projection onto the
PCS and subtract out the random component, which is given by

\[ \hat{x} = PP^T x \]  

(5.7)

The estimation error of the approximation is given by

\[ \varepsilon = x - \hat{x} = \tilde{P}\tilde{P}^T x \]  

(5.8)

The expectation of the squared error is given by

\[ \varepsilon^2 = E[\varepsilon^T \varepsilon] = E[x^T \tilde{P}\tilde{P}^T \tilde{P}\tilde{P}^T x] = E[\tilde{T}^T \tilde{T}] = \sum_{k=l+1}^n \lambda_k \]  

(5.9)
The above equation shows that the mean-squared error is equal to the sum of the least significant eigenvalues of the covariance matrix. This also indicates that the choice of the number of principal components is to neglect the eigenvectors as the basis vectors for expansion that correspond to the least significant eigenvalues of the covariance matrix of the original measured data.

5.3. Selection of Number of Principal Components

Many approaches have been proposed to determine the number of principal components for different applications. Cumulative Percent Variance, Scree Plot, Average Eigenvalue, and Cross Validation are discussed by several earlier investigators. The cumulative percent variance method selects the number of principal components by setting a threshold of cumulative percent variance. The Scree plot method is based on the plot of the fraction of variance explained by each principal component. The plot orders the principal components from the one that explains the largest amount of variation to the one that explains the least amount of variation. This method considers the beginning point of the Scree as the most reasonable number of principal components. The average eigenvalue method assumes that all the components whose corresponding eigenvalues are less than the average value should be discarded. When cross validation method is used to determine the number of principal components, the original data are randomly divided into N-blocks, the cross validation error is computed as the residual sum of squares (RSS) for one block of data with the PCA model built using the other blocks of data. The number of principal components is chosen to be the one beyond which the TSS begins to increase. In this project, the Scree plot method is used to determine the number of principal components.

5.4. Fault Detection of Sensors and Field Devices

Fault detection can be performed by monitoring the change of the correlation structure of the measured data. Because the variation of data is separated in the principal component space and the residual space, two statistics are defined to measure the variation in the two spaces, respectively. If a new observation exceeds the effective region in the PC space defined by the normal operation data, a change in operation regime can be detected. If a significant residual is observed in the residual space, a special event, either due to disturbance changes or due to changes in the relationship between variables, can be detected.
The Q-statistic is the measurement of distance from an observation to the PCA model or new vector space formed by the selected PCs.

\[ Q = e^T \varepsilon \]  \hspace{1cm} (5.19)

where \( e = \tilde{P} \tilde{P}^T x \).

Under the standard assumptions of a multivariate normal distribution for \( \varepsilon \), \( \varepsilon \sim N(0,1) \), control limits at confidence level \( \alpha \) may be obtained using the \( \chi^2 \) distribution as

\[ Q_\alpha = \theta_1 \left[ c_\alpha \sqrt{\frac{2\Theta_0}{\theta_1}} + 1 + \frac{\Theta_2 h_0(h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \]  \hspace{1cm} (5.21)

where

\[ \theta_i = \sum_{i=1}^{n} \left( \frac{\lambda_i}{m - 1} \right)^i, \ i = 1, 2, 3; \ m \text{ is the number of measurements}. \]

\[ h_0 = 1 - \frac{2\Theta_1 \Theta_3}{3\Theta_2^2} \]

\( c_\alpha \) is the standard normal deviate (equal to 2.57 for \( \alpha = 0.01 \)).

5.5. Fault Isolation of Sensors and Field Devices

The task of fault isolation (or identification) is to determine what the most affected variables are, once a fault happens. Fault identification is useful because it can help operators focus their attention on a reduced number of variables.

5.5.1. Fault Directions

Let \( f_i \) represent the direction in the residual space for the \( i \)-th fault such that the samples corresponding to the fault have the maximum projection on \( f_i \). In other words, if \( E_i \) denotes the residuals for samples corresponding to the \( i \)-th fault, the optimization problem is
\[ J = \max_{f_i} f_i^T E_i^T E_i f_i \]

Subject to the constraint \( f_i^T f_i = 1 \).

Using the Lagrangian multiplier and differentiating \( J \) with respect to \( f_i \) and setting the derivative to zero for maximization, gives

\[ 2E_i^T E_i f_i - 2\sigma f_i = 0 \quad \text{or} \quad E_i^T E_i f_i = \sigma f_i \quad (5.22) \]

Then the fault direction \( f_i \) is the eigenvector of \( E_i^T E_i \) corresponding to the largest eigenvalue.

### 5.5.2. Fault Isolation Index

Let \( F = [f_1, f_2, \ldots, f_R] \), where \( f_1, f_2, \ldots, f_R \) are \((n \times 1)\) column vectors, denote the fault directions for the various fault scenarios that are observed in the database. In case the fault is of a particular type \( j \), then the projection of the residuals on \( f_j \) would be very high. Fault isolation is accomplished by calculating the projections onto \( F \) and classifying the fault as the one with the maximum projection norm. For example, a fault isolation index for the \( i \)-th fault is defined as

\[ FI_i = 1 - \frac{Q_i}{Q}, \quad (5.23) \]

where

\[ Q_i = e(1 - f_i f_i^T)(1 - f_i f_i^T)^T e^T \quad \text{and} \quad Q = ee^T. \]

In the above equations, \( Q_i \) denotes the distance of the sample from the origin after subtracting the projection of the residuals on the fault direction \( f_i \). It represents sum of the squares of the residuals remaining after removing the contribution from the \( i \)-th fault direction. The fault isolation index quantifies the fraction of \( Q \) that is due to \( f_i \). If the \( j \)-th fault scenario occurs, \( FI_i \) varies from 1 to \( R \). This results in the isolation of the fault from the various existing scenarios. The fault matrix \( F \) can be extracted from historic data as described in the previous section.
5.6. Application of PCA to the SP-100 Reactor System

The application of reconstruction based PCA approach to SP-100 is presented in this section.

5.6.1. Data Generation and Model Development

The normal operation data were generated by adjusting the reactor control drum angle from 125.93° to 115.93°, to simulate reactor power change using the simulation model developed in Chapter 2. The step increment is 2°, and six groups of data are generated, which cover 100% to around 80% of the plant thermal power capacity. At each power level 200 samples were recorded. Table 5.1 lists the 8 measured variables used to develop the PCA model for SP-100 system. Before the simulated data are used to build a model, Gaussian noise with mean zero, and variance, equal to ±0.2% of the mean value of the whole data set, is added to the data to mimic the measurement noise of the corresponding sensors.

Table 5.1. Measured variables used to develop PCA model

<table>
<thead>
<tr>
<th>Variable number</th>
<th>Symbol</th>
<th>Variable Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tpout</td>
<td>Outlet Coolant Temperature for the primary loop</td>
</tr>
<tr>
<td>2</td>
<td>Tpin</td>
<td>Inlet Coolant Temperature for the primary loop</td>
</tr>
<tr>
<td>3</td>
<td>Tsout</td>
<td>Outlet coolant temperature for the secondary loop</td>
</tr>
<tr>
<td>4</td>
<td>Tsin</td>
<td>Inlet coolant temperature for the secondary loop</td>
</tr>
<tr>
<td>5</td>
<td>Thot</td>
<td>Hot leg temperature</td>
</tr>
<tr>
<td>6</td>
<td>Tcol</td>
<td>Cold leg temperature</td>
</tr>
<tr>
<td>7</td>
<td>Pt</td>
<td>Reactor thermal power</td>
</tr>
<tr>
<td>8</td>
<td>Pe</td>
<td>Reactor electric power</td>
</tr>
</tbody>
</table>
The standard deviation of the measurement noise added to each variable is listed in the Table 5.2.

Table 5.2. Magnitude of measurement noise

<table>
<thead>
<tr>
<th>Variables</th>
<th>Tpout (K)</th>
<th>Tsout (K)</th>
<th>Tpin (K)</th>
<th>Tpout (K)</th>
<th>Thot (K)</th>
<th>Tcol (K)</th>
<th>Pt (kW)</th>
<th>Pe (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise standard deviation</td>
<td>2.4346</td>
<td>1.5928</td>
<td>2.3747</td>
<td>1.5360</td>
<td>2.3353</td>
<td>1.6274</td>
<td>3.5497</td>
<td>0.1866</td>
</tr>
</tbody>
</table>

A PCA model is built using the data for the nominal operation case. The nominal operation data matrix is preprocessed by auto-scaling the columns in the data matrix to zero mean and unit variance. This puts all the measurements with their different units on a common unit variance scale.

5.6.2. The PCA model

Figure 5.1 shows the fractions of the variance contained in the data explained by the 8 eigenvectors. Taking just the first eigenvector as the principle component can explain more than 99% of the total variation in the data. From the figure, we also can observe that the measured variables are very collinear.
Figure 5.1. Fractions of the variance explained by different PC components.

The eigenvector to define the model space is as follow:

-0.3537
-0.3531
-0.3537
-0.3530
-0.3537
-0.3532
-0.3540
-0.3539

Figure 5.2 shows the true measurement values and the PCA predicted values. We can see that values PCA model prediction can keep the main information of the data and also filter some noise.
5.6.3. Fault detection using the PCA approach

From above PCA model, we can calculate that $Q_\alpha = 0.1063$ when $\alpha = 0.01$. The $Q^2$ statistics for the normal measurements with temperature drift fault is plotted in Figure 5.3. The blue line is the Q-statistics limits corresponding to 99% confidence level. It illustrates that the fault free data are well below the limit line. Figure 5.3 shows the Q-statistics based fault detection for the conditions of a drift fault in the measurement of the reactor outlet temperature.

This PCA model can only deal with steady state condition or a slow dynamic process. The algorithm to perform PCA based fault detection is only applicable to steady state condition. The choice of the confidence level affects the false alarm. In a real application, the confidence level needs to be adjusted according to the operation requirements.
5.6.4. Fault identification

Assume that there is a 5%-2% drift fault existing in one sensor, using the method introduced in Section 5.4.2, we can calculate the fault isolation indices which represent the sum-of-squares of residuals remaining after removing the contribution from the \( i \)-th fault direction. The fault isolation indices for the detected fault scenarios are plotted in Figures 5.4 - 5.6. For the particular fault in each case, the fault index is close to unity, and for the others it is much smaller. If the drift magnitude is larger, it is easier to isolate the sensor with measurement fault from those without fault. This is demonstrated in Figure 5.5. If the drift magnitude is smaller, the measurement noise affects the distinction between the sensors with fault and those without fault (Figure 5.6). If it is known there is only one fault exits in the system, then we still can isolate the fault using the fault indices by checking which one is closer to unity. All the fault cases are detected correctly using the fault isolation approach described earlier. This demonstrates the effectiveness of the proposed method for fault detection and isolation in complex systems with several interacting units and feedback control loops.

Figure 5.3. Q-statistics for Temperature measurement drift fault.
Figure 5.4. Fault isolation index with 1.5% measurement drift fault.

Figure 5.5. Fault isolation index with 2% measurement drift fault.
Figure 5.6. Fault isolation index with 0.5% measurement drift fault.
6. DEVELOPMENT OF A LABORATORY MULTIVARIATE CONTROL FLOW LOOP

6.1. Design Features of the Two-Tank Flow Control Loop

An experimental flow loop with two tanks, control valves, and a variety of sensors, has been designed and constructed and is used to evaluate the performance of the classical proportional-integral controller and the model-predictive controller, along with and the demonstration of the fault-tolerant control strategy. Figure 6.1 is a schematic of this loop with low-pressure water circulation that is facilitated by a fractional horsepower motor-driven pump.

The loop has a centrifugal pump, one stainless steel water reservoir, two acrylic tanks, four flow meters, two level transmitters, four motor-operated control valves, and three manual valves. The piping is made of PVC schedule 80, with diameters varying from 1/2” to 3/4” and with enough flexibility to accommodate minor design changes. A data acquisition system is used to monitor and control the loop by varying the position of two control valves until a stationary flow throughout the system as well as the water level in each of the acrylic tanks is attained. A manual valve in between tanks provides the capability to work with both acrylic tanks at the same level.

A bypass valve is provided to divert any excess water back to the water reservoir. The maximum water flow is estimated to be around 27 GPM. Though water temperature will be monitored, temperature control is not addressed in this work but changes can be implemented for this purpose. A 4ft deep, 7ft long and 7ft tall steel frame supports the equipment, up to a maximum load of 800 pounds. Figure 6.2 shows a simplified schematic of the structure. The major equipment and parts are listed in Table 6.1.
Figure 6.1. Schematic of the two-tank experimental control loop.
Figure 6.2. Two-dimensional and 3-D views of the steel support structure.
Table 6.1. List of Equipment and Major Parts

<table>
<thead>
<tr>
<th>ITEM</th>
<th>DESCRIPTION</th>
<th>QUANTITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rosemount 3051S Differential Pressure Transmitter. Part Number: 3051S 1CD 2 A 2 E12 A 1A B4</td>
<td>04</td>
</tr>
<tr>
<td></td>
<td>Input Range: -250 to 250 inH₂O Output: 4 – 20 mA</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Rosemount 1195 Integral Orifice Primary Element. Part Number: 1195 S 010 P1 S 0150 C</td>
<td>02</td>
</tr>
<tr>
<td>3</td>
<td>Omega Turbine Flow Meters FTB-953</td>
<td>02</td>
</tr>
<tr>
<td>4</td>
<td>Worcester Control Valves – Series 75</td>
<td>04</td>
</tr>
<tr>
<td>5</td>
<td>Ball Valve</td>
<td>01</td>
</tr>
<tr>
<td>6</td>
<td>Globe Valves</td>
<td>02</td>
</tr>
<tr>
<td>7</td>
<td>Acrylic Tanks and Fittings</td>
<td>02</td>
</tr>
<tr>
<td>8</td>
<td>Steel Tank</td>
<td>01</td>
</tr>
<tr>
<td>9</td>
<td>Aluminum Sheets</td>
<td>02</td>
</tr>
<tr>
<td>10</td>
<td>Shortec Centrifugal Pump CMPP12T (36GPM max.)</td>
<td>01</td>
</tr>
<tr>
<td>11</td>
<td>Steel frame, bolts, nuts, brackets, wheels, etc</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>Electrical Connections (wires, cables, buttons, etc)</td>
<td>-</td>
</tr>
<tr>
<td>13</td>
<td>Computer and Data Acquisition Card</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>LabView, MATLAB</td>
<td>2</td>
</tr>
</tbody>
</table>
6.2. General Description of the Experimental Control Loop

The two-tank control loop is built on a wheeled table-like steel frame structure. This structure holds all sensors, piping, pump, sump tank, aluminum table top, cables, control valves, manual valves, and two tanks and can be easily moved around. Since eighty percent of the piping used to build the loop are CPVC and union connections were strategically distributed; any maintenance or setup modification can easily be carried out. Figure 6.3 shows the main structure right at the beginning of construction and in Figure 6.4 shows the detail of the bypass hole drilling operation.

Figure 6.3. Main structure.  Figure 6.4. Bypass hole drilling.

6.3. Description of Components

A description of each of the major loop components is given in this section. Some components were bought or built and some were donated by our sponsors.

6.3.1. Water Tanks

For level control experiments, two similar acrylic tanks are installed in this loop, referred to as Tank 1 and Tank 2, respectively. Their dimensions are: 5.75 inch diameter and 36 inch long. A 27-gallon stainless steel tank is installed underneath the table top to provide the
necessary water storage for the circuit. Figure 6.5 shows a detail of the two acrylic tanks used in this project.

![Figure 6.5. Details of water tanks.](image)

### 6.3.2. Sensors for Process Measurement

There are 11 sensors installed in the loop: 4 differential pressure sensors, 4 thermocouples, and 3 turbine meters. Also installed are 2 primary orifice plates. Each sensor is discussed below.

**A. Pressure Transmitters**

Four Rosemount™ differential pressure sensors are currently installed. Two of them are used to measure the water level in each tank, and two other pressure sensors are connected to orifice meters and are used to measure the water flow going into each of the level tanks. Information about the pressure sensors is shown in Table 6.2 and Figure 6.6 shows the details of two of the sensors. To calibrate the pressure sensors range, a software package called **AMS Suite** from Emerson® was used. This software package can provide us an easy way to calibrate, zero-trim, schedule maintenance, and keep record of every calibration performed. It comes with an RS-232 modem and cable to plug into the computer serial port. On the other end a pair of probes is used to connect the computer to the sensor terminals.
Table 6.2. Information about the pressure sensors

<table>
<thead>
<tr>
<th>Sensor ID</th>
<th>Function</th>
<th>Output</th>
<th>Calibration Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tank 1</td>
<td>Measures Level in Tank 1</td>
<td>4 – 20mA</td>
<td>0 – 900 mmH₂O</td>
</tr>
<tr>
<td>Tank 2</td>
<td>Measures Level in Tank 2</td>
<td>1 – 5 Volts</td>
<td>0 – 900 mmH₂O</td>
</tr>
<tr>
<td>Flow 1</td>
<td>Tank 1 Inlet flow</td>
<td>4 – 20mA</td>
<td>0 - 6303 mmH₂O</td>
</tr>
<tr>
<td>Flow 2</td>
<td>Tank 2 Inlet flow</td>
<td>4 – 20mA</td>
<td>0 - 6303 mmH₂O</td>
</tr>
</tbody>
</table>

Figure 6.6. Pressure sensors used in the loop.

B. Flow Meters

Two different types of flow rate sensors are used in the loop: turbine and orifice meters. The orifice meters are provided by Rosemount™ and the turbine meters are manufactured by Omega™. There are 3 turbine meters installed in the loop: one at each tank outlet and one in the bypass; all of them are factory calibrated. A signal conditioner transforms the frequency output (0-1000Hz) to voltage (0-5V or 0-10V). The volumetric flow is calculated using the calibration tables provided by the manufacturer for both signal conditioner and turbine flow meter. In Figure 6.7 a turbine meter and the signal conditioner are shown. Figure 6.8 shows some details of the orifice plate and pressure lines. Table 6.3 is an example of a calibration table provided for one of the turbine flow meters by the manufacturer, showing the relationship between frequency and
voltage and volume and volumetric flow rate. The orifice meter calibration procedure and results are shown in Appendix B. Figures 6.9 and 6.10 show the calibration data points and the predicted values using polynomial fits for both voltage and volumetric flow.

Figure 6.7. Turbine and signal conditioner.  
Figure 6.8. Orifice plate.

Table 6.3. Calibration table provided by manufacturer for the turbine flow meter

<table>
<thead>
<tr>
<th>% of F/S Input Frequency</th>
<th>Input Frequency (Hz)</th>
<th>Calculated Output (VDC)</th>
<th>Output (VDC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0.0</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>25%</td>
<td>250.0</td>
<td>2.500</td>
<td>2.502</td>
</tr>
<tr>
<td>50%</td>
<td>500.0</td>
<td>5.00</td>
<td>5.000</td>
</tr>
<tr>
<td>75%</td>
<td>750.0</td>
<td>7.5</td>
<td>7.500</td>
</tr>
<tr>
<td>100%</td>
<td>1000.0</td>
<td>10.00</td>
<td>10.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Rate GPM</th>
<th>Time [s]</th>
<th>Approximate Frequency [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>27.808007</td>
<td>9.083</td>
<td>813.60784</td>
</tr>
<tr>
<td>21.849580</td>
<td>11.560</td>
<td>639.87889</td>
</tr>
<tr>
<td>14.621356</td>
<td>17.275</td>
<td>427.38061</td>
</tr>
<tr>
<td>8.182758</td>
<td>30.869</td>
<td>238.94523</td>
</tr>
<tr>
<td>2.400992</td>
<td>105.205</td>
<td>70.25331</td>
</tr>
</tbody>
</table>
C. Thermocouples

To monitor the temperature, four type-J thermocouples are installed: one inside each level tank, one in the reservoir tank, and one to monitor the pump temperature. None of these sensors are calibrated and a standard calibration curve is used.

D. Water Circulation Pump

A centrifugal Noryl-27 GPM pump is used to circulate the water through the loop and is installed underneath the circuit. It has a 0.5 HP motor, 1.5 inch inlet – 1.5 inch outlet piping connections, and rated at 110-230V. Details of the pump and of a thermocouple used to monitor its temperature are shown in Figure 6.11.
E. Control Valves

In order to control the flow in the loop, three motor-operated control valves are used: one on each tank inlet and one at tank-1 outlet. Each of the control valves has two components: an electric actuator and a 0.5 inch ball valve. The electric actuators were manufactured by Worcester Controls™ and one of them is shown in Figure 6.12. These actuators are responsible for opening and closing the ball valves to regulate the flow. The actuators are 120VAC powered, with input and output of 2-10 VDC, and can be manually or remotely operated, with a complete stroke time of about 15 seconds.
F. Data Acquisition Software

A National Instruments® (NI) data acquisition card installed in the computer converts the analog signals coming from the various sensors and components to digital signals through A/D converters. These signals are processed by a Virtual Instrument (VI) program developed using NI LabView™ package and are presented in engineering units. The data acquisition card, model PCI-6259 is a 16-bit converter with 32 analog inputs, and a maximum sampling rate of $1.2 \times 10^6$ Hz. Figure 6.13 shows the front panel of one of two monitor screens available to the user. This version shows the trend for each of the variables monitored or controlled. Figure 6.14 is a more intuitive version showing the location of every variable in the loop. The loop has a PI controller through interface with a MATLAB program.
Figure 6.13. Virtual Instrument (VI) front panel, showing signal trends.

Figure 6.14. Virtual Instrument (VI) front panel, showing all the sensors and devices.
7. MODEL PREDICTIVE CONTROLLER APPLIED TO A TWO-TANK FLOW CONTROL SYSTEM USING PHYSICS MODELS

7.1. Background of Optimal Control

Linear Quadratic Regulator (LQR) has been initially studied for the optimal control of multivariable systems. In the design of LQR, a quadratic objective function is minimized. For a discrete linear process, the state space representation is as follows:

\[
x_{k+1} = Ax_k + Bu_k \\
y_k = Cx_k
\]

(7.1)

where

\[
\begin{align*}
u_k &= \text{process inputs.} \\
y_k &= \text{process output.} \\
x_k &= \text{process states.}
\end{align*}
\]

The objective function of LQR control is to minimize the weighted squared input and state deviations with state and input weight matrices Q and R, which is given by:

\[
J = \sum_{j=1}^{\infty} \| x_{k+j} \|_Q^2 + \| u_{k+j} \|_R^2
\]

(7.2)

It can be shown that the solution to the LQR problem is a proportional controller, with a gain matrix K computed from the solution of a matrix Ricatti equation:

\[
u_k = -Kx_k
\]

(7.3)

One of the significant properties of the LQR algorithm is that the infinite prediction horizon allows to stabilize any reasonable linear system as long as the objective function weight matrices Q and R are positive definite.

Although LQR theory is a powerful tool to control an unconstrained linear plant, it has very limited applications in process industry where complex, nonlinear, constrained multivariable systems must be dealt with. In fact, a successful controller has to maintain a
process close to some constraints while without violating them in order to achieve high economic performance. In addition, a process system typically has time dependent operation characteristics. It is difficult in the framework of LQR theory to express the performance criteria.

Model Predictive Control is a more general model based control methodology than LQR. In MPC, any desired objective function can be used to solve a generalized dynamic control optimization problem. Process input and output constraints are directly included in the formulated optimization problem. In addition, all types of models can be handled in the same manner without changing the MPC algorithm.

Model Predictive Control (MPC) is also appropriate for a hierarchical implementation proposed in this research. The top of the hierarchical structure is a plant-wide optimizer to determine optimal steady-state settings for each unit in the space reactor system. These optimal steady-state settings may be sent to local optimizers at each unit such as the reactor core and the thermal electric generator that run more frequently or consider a more detailed model. The unit optimizer computes an optimal economic steady state and passes this to the dynamic constraint control system for implementation. The dynamic constraint control moves the plant from one constrained steady state to another while minimizing constraint violations. In the MPC methodology, the combination of PID algorithms, Lead-Lag (L/L) blocks and High/Low select logic, which are implemented in conventional control, is replaced by a single MPC controller.

7.2. Theory of Model Predictive Control

For a discrete-time linear system, the state space representation is given by

\[ x_{k+1} = Ax_k + Bu_k \]
\[ y_k = Cx_k \]
\[
\begin{bmatrix}
x_{k+1} \\
x_{k+2} \\
x_{k+3} \\
\vdots \\
x_{k+n_y}
\end{bmatrix} = \begin{bmatrix} A \\
A^2 \\
A^3 \\
\vdots \\
A^{n_y}
\end{bmatrix} x_k + \begin{bmatrix} B \\
AB \\
A^2B \\
\vdots \\
A^{n_y-1}B
\end{bmatrix} \begin{bmatrix} 0 \\
B \\
AB \\
\vdots \\
A^{n_y-2}B
\end{bmatrix} \begin{bmatrix} 0 \\
B \\
AB \\
\vdots \\
A^{n_y-3}B
\end{bmatrix} \begin{bmatrix} \vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots
\end{bmatrix} \begin{bmatrix} u_k \\
u_{k+1} \\
u_{k+2} \\
\vdots \\
u_{k+n_y-1}
\end{bmatrix}
\]

(7.4)

\[
\begin{bmatrix}
y_{k+1} \\
y_{k+2} \\
y_{k+3} \\
\vdots \\
y_{k+n_y}
\end{bmatrix} = \begin{bmatrix} CA \\
CA^2 \\
CA^3 \\
\vdots \\
CA^{n_y}
\end{bmatrix} x_k + \begin{bmatrix} CB \\
CAB \\
CA^2B \\
\vdots \\
CA^{n_y-1}B
\end{bmatrix} \begin{bmatrix} 0 \\
CB \\
CAB \\
\vdots \\
CA^{n_y-2}B
\end{bmatrix} \begin{bmatrix} 0 \\
CB \\
CAB \\
\vdots \\
CA^{n_y-3}B
\end{bmatrix} \begin{bmatrix} \vdots \\
\vdots \\
\vdots \\
\vdots \\
\vdots
\end{bmatrix} \begin{bmatrix} u_k \\
u_{k+1} \\
u_{k+2} \\
\vdots \\
u_{k+n_y-1}
\end{bmatrix}
\]

(7.5)

\[
\begin{align*}
x_k &= P_{xx} x_k + H_x u_k \\
y_k &= P x_k + H u_k
\end{align*}
\]

(7.6)

In MPC, the performance function has the form

\[
J = \sum_{i=1}^{n_y} \left\| W_y (r_{k+i} - y_{k+i}) \right\|_2^2 + \lambda \sum_{i=0}^{n_y-1} \left\| W_u \Delta u_{k+i} \right\|_2^2
\]

(7.7)

The performance function has two components, the weighted sum squares of the predicted tracking errors over the output horizon \(n_y\) with a weight matrix \(W_y\) and the weighted sum squares of the control changes over the control horizon \(n_u\) with a weight matrix \(W_u\). The scalar parameter \(\lambda\) balances the penalty on the predicted tracking errors and the control changes. However, the two weight matrices, \(W_y\) and \(W_u\), are positive diagonal matrices. The balance between the tracking performance of each loop and its corresponding input activity can be realized by tuning the diagonal elements of these two matrices.
In the definition of MPC optimization problem, the number of control inputs and the length of the control horizon are somewhat independent. Therefore, the control problem can be stated to minimize the objective function with respect to the control move, which is written as follows:

$$\min J = \sum_{i=1}^{n} ||W_y (r_{k+i} - y_{k+i}) ||_2^2 + \lambda \sum_{i=0}^{n-1} || W_u \Delta u_{k+i} ||_2^2$$

(7.8)

The optimization algorithm is constantly solved over a receding horizon. In other words, the time window over which to solve the optimization problems is constantly moving at the same speed. In the MPC algorithm, the constraints on the input, output, as well as process states can also systematically taken into account. For example consider the following constraints.

$$u_L < u < u_U$$
$$\Delta u_L < \Delta u < \Delta u_U$$
$$x_L < x < x_U$$
$$y_L < y < y_U$$

(7.9)

In order to solve a real world control problem, the model predictive control algorithm needs to achieve three requirements:

- Include a disturbance model.
- Use a model giving offset free prediction.
- Ensure the minimum of the cost function is consistent with zero tracking errors and a modified performance index defined in the MPC algorithm.

One solution to achieve the requirements is to form an algorithm to give estimates of $x_{ss}$ and $u_{ss}$ which are consistent with zero tracking errors and robust to model uncertainty, and redefine the cost functions as follows:

$$\min J = (x - x_{ss})^T Q (x - x_{ss}) + (u - u_{ss})^T R (u - u_{ss})$$

(7.10)

Compared to the conventional cost function, $J = x^T Q x + u^T R u$, the new cost function is consistent with zero tracking errors even if there is model uncertainty.
Based on the LQR theory, the above minimization will give a state feedback control with the following form:

\[
u_k - u_{ss} = -K(x - x_{ss}) = -[H_x^T QH_x + R]^{-1} H_x^T QP_{ss}(x - x_{ss})
\]

(7.11)

In the derived control law, it is necessary to ensure unbiased estimates of the steady state values by using an appropriate disturbance observer. The original state space model with additive model uncertainty is as follows:

\[
\begin{align*}
x_{k+1} &= Ax_k + Bu_k \\
y_k &= Cx_k + d_k
\end{align*}
\]

(7.12)

where \(d_k\) represents model uncertainty taking into account of disturbance and model mismatch.

The standard Kalman filtering method may be used to estimate the true state \(x\) and model uncertainty \(d\), which is as follows:

\[
\begin{pmatrix}
x_{k+1} \\
d_{k+1}
\end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} x_{k+1} \\ d_{k+1} \end{pmatrix} + \begin{pmatrix} B \\ 0 \end{pmatrix} u_k - L \begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} x_k \\ d_k \end{pmatrix} - y_k
\]

(7.13)

where \(L\) is the observer gain.

Once the model uncertainty \(d\) and the output reference \(r\) are known, the steady state values of \(x_{ss}\) and \(u_{ss}\) can be estimated as follows:

\[
\begin{align*}
x_{ss} &= Ax_{ss} + Bu_{ss} \\
y_{ss} &= Cx_{ss} + d
\end{align*}
\]

(7.14)

Therefore, the control law can be written as follows:

\[
u_k = -K(x - x_{ss}) + u_{ss}
\]

(7.15)

If the control law is given as above, when a steady state is reached the observer output must be equal to the process output, that is,
\[ x_k = Ax_k + Bu_k \]
\[ Cx_k + d = y_k \]

(7.16)

Again,

\[ x_{ss} = Ax_{ss} + Bu_{ss} \]
\[ y_{ss} = Cx_{ss} + d \]

For steady state condition, \( x_k = x_{ss} \) and \( u_k = u_{ss} \). Therefore \( y_k = y_{ss} \).

7.3. Two-Tank Flow Control Loop

In this research, a two tank experimental setup is designed to demonstrate the integration of model predictive control algorithm for the purpose of autonomous control. Specifically, the experimental setup has the following objectives:

- Investigate the extent to which a decentralized controller is capable of controlling the process as the interactions increase by increasing the open width of Valve 3.
- Implement partial or full decoupling to reduce process interactions.
- Study multivariable model based control.
- Investigate the performance of system identification for system control and fault diagnosis.
- Study robust model based control by introducing disturbance through adjusting the width of valve 1 and valve 2 in a sine manner.
- Demonstrate model based fault diagnosis methods.
Figure 7.1 shows a schematic of the two tank experimental setup. It consists of three tanks of Plexiglas Tubes; one constant speed centrifugal pump; two control valves; one connection valve between two tanks; two drainage valves; and a bypass valve. In order to facilitate system identification and fault diagnosis, flow meters are installed to measure the mass flow rate of the two control valves, the connection valve, and the two drainage valves. In addition, two pressure transmitters are installed to measure the water level of the two tanks for level control.

The two-tank system follows the following dynamic equations:

\[
A \frac{dH_1}{dt} = Q_1 - C_1 a \sqrt{2gH_1} - C_2 b \sqrt{2g(H_1 - H_2)} 
\]  \hspace{1cm} (7.17a)

\[
A \frac{dH_2}{dt} = Q_2 + C_2 b \sqrt{2g(H_1 - H_2)} - C_2 a \sqrt{2gH_2} 
\]  \hspace{1cm} (7.17b)

where
Q1 and Q2 = the volumetric flow rate through the control valve 1 and control valve 2, respectively.
H1 and H2 = the water level in the tank 1 and tank 2, respectively.
A = the cross section area of the two tanks.
g = the gravitational acceleration.
a = the cross section of the drainage pipe.
b = the cross section of the connection pipe.

7.4. Results of MPC for the Two Tank System

Figure 7.2 shows the results of model predictive control for the two tank level system. The comparison between Figure 2(a) and Figure 7.2(b) shows that as the prediction horizon increases, the closed loop dynamics slow down. Figure 7.2(c) shows the effects of varying the weights of output. As more weight is given to the tank-1 level, the closed loop performance of the tank level-2 degrades.

The results in Figure 7.2 also show that MPC algorithm can provide an excellent solution to multivariable control. The advantage of MPC over traditional decoupling PID control is that the decoupling can be realized by solving an optimization problem with unified performance function. The desired closed loop performance can be achieved by tuning the length of control horizon and prediction horizon. In addition, the constraints of process input and output can be handled with great flexibility.
Q1=1.0E-3 m³/s  Q2=1.0E-3 m³/s  PredictionHorizon=10  Control Horizon=2  W.Output=[1.0, 1.0]

MV(1).Min=0.0  MV(1).Max=Q1*2  MV(2).Min=0.0  MV(2).Max=Q2*2  MV(1).RateMin=-Q1*0.35  MV(1).RateMax=Q1*0.35  MV(2).RateMin=-Q2*0.35  MV(2).RateMax=Q2*0.35

Figure 7.2(a). MPC Control of Two Tank Level System

Q1=1.0E-3 m³/s  Q2=1.0E-3 m³/s  PredictionHorizon=100  Control Horizon=2  W.Output=[1.0, 1.0]

MV(1).Min=0.0  MV(1).Max=Q1*2  MV(2).Min=0.0  MV(2).Max=Q2*2  MV(1).RateMin=-Q1*0.35  MV(1).RateMax=Q1*0.35  MV(2).RateMin=-Q2*0.35  MV(2).RateMax=Q2*0.35

Figure 7.2(b). MPC Control of Two Tank Level System
7.5. Multivariate PI Control of the Two Tank System

Coming back to the basic controller design strategy, we use a proportional-integral (PI) controller to test the performance of the MPC. The experiment was performed on the SIMULINK panel of MATLAB software. Figure 7.3 shows the Simulink model of the PI controller. Tank-1 reference level and tank-2 reference level are the set point water level values of tank-1 and tank-2. The water level in tank-1 changes from 0.5 m to 0.3 m at the 30th second, and the water level in tank-2 changes from 0.5 m to 0.8 m at the 30th second. G11 and G12 are the 1st order approximations of the response of tank-1 level and tank-2 level triggered by unit step change of tank-1 reference level, and G21 and G22 are the 1st order approximations of the response of tank-1 level and tank-2 level triggered by unit step change of tank-2 reference level. Fcn1 and Fcn2 are the non-linear parts of the dynamic equations.

\[
Fcn1 = \left( -C_a \sqrt{2gH_1} - C_d b \sqrt{2g(H_1 - H_2)} \right) / A, \tag{7.18a}
\]
\[ Fcn2 = \left( C_d b \sqrt{2g(H_1 - H_2)} - C_2 a \sqrt{2gH_2} \right) / A. \] (7.18b)

Figure 7.3. Simulink model of the PI controller.

Figure 7.4 shows the results of PI control for the two tank level system. The comparison between Figure 7.4(a), Figure 7.4(b), and Figure 7.4(c) shows that as the Kp and Ki change in G11, G12, G21 and G22, the closed loop dynamics will change a lot. It is hard to get the optimal Kp and Ki to make the system get the best performance. In another word, it is hard to tune the system. There are eight parameters to tune. No matter which parameter is changed, the response of water level will change.
G11=0.1+0.5/s, G12=0.3+0.074s,

Figure 7.4(a). PI Control for the two tank level system.
$G_{11} = G_{22} = 0.06 + 0.5/s$, $G_{12} = G_{21} = 0.3 + 0.074s$,

Figure 7.4(b). PI Control for the two tank level system.
G11=G22=0.06+0.5/s, G12=G21=0.3+0.3/s,

Figure 7.4(c). PI Control for the two tank level system.

### 7.6. Comparison of Multivariable PI Control and MPC

From above examples we can get that the MPC has more advantages than PID controller.

1. MPC handles difficult process dynamics with ease. For PID controller, we need the 1st order approximation of the dynamic system.
2. MPC delivers higher performance: robustness against model errors and non-linearities; handles interactions and constraints.
(3) MPC acts smoother on the process than PID controller.
(4) MPC reduces the tuning effort. It has separate tuning for both load upsets and set-point changes. For PID controller, there are four controllers need to be tuned in this example. And they are coupled. It is hard to balance the parameters for each controller.
8. DEMONSTRATION OF PROPORTIONAL-INTEGRAL CONTROLLERS IN THE EXPERIMENTAL LOOP

Operation of the flow control loop, data acquisition from process sensors, and the implementation of the PI controller are described in this section.

8.1. Data Acquisition Hardware

Two National Instrument® (NI) data acquisition cards installed in the computer converts the analog signals coming from the various sensors and components to digital signals through A/D converters. These signals are processed by a Virtual Instrument (VI) program developed using NI LabView™ package and presented in engineering units. The acquisition cards are, respectively, models PCI-6259 and PCI-6259E. They are 16-bit, 32 analog inputs, with up to 1.2 MS/s. These cards have two 1-10V analog outputs each. These outputs can be used to open or close the control valves and change the inlet flow into each of the tanks. Three SCB-68 patch panels are used with the acquisition cards. These patch panels are used for interfacing the acquisition software and control with the loop and the acquisition cards. The programs used for control and acquisition and control are described next.

8.2. Data Acquisition and Control Virtual Instrumentation (VI)

There are two main virtual instrumentation or VIs currently being used for control and data acquisition. The control VI is used to control the loop, based on the experiment to be carried out. This VI has several different options to choose from embedded into it, seen in Table 8.1, giving the researcher complete control over the loop, but as stated previously, installing a fourth control valve in tank 1 outlet, will further increase the loop flexibility and our control over the loop dynamics. Among other options, the researcher can choose to have manual or automatic set-point change in a ramp-like fashion, as well as controlling the rate at which the set-point can be changed in mm-H₂O/minute. One can choose between manual and automatic control, using the PI or the MPC control. Finally, the researcher can choose from manual or automatic control for either one tank or both tanks at the same time. When using PI control, if a fine-tuning is required, it can be accomplished by choosing the option located on the main panel, shown in
Figure 8.1. Some options are common to all controls, but for the most part, the set of options available for each control (manual, PI or MPC) is type dependent, i.e., for each type of control there is a relevant set of options to choose from that will pop-up on the front panel.

Table 8.1. Loop Control Options

<table>
<thead>
<tr>
<th>Control Option</th>
<th>Tank 1</th>
<th>Tank 2</th>
<th>Both Tanks</th>
<th>Setpoint Profile</th>
<th>Tuning</th>
<th>Fault Insertion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
</tr>
<tr>
<td>PI</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>MPC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>N/A</td>
<td>Yes</td>
</tr>
</tbody>
</table>

At start up, the control VI will call the data acquisition VI and Matlab, the last to be used with MPC. The data acquisition VI is responsible for displaying the current loop status on the screen and saving the data to disk whenever needed. The acquisition rate can be change to accommodate the kind of experiment being run, i.e., for fast or slow changing dynamics. The data saved in ASCII file, which can be read in any software able to handle this kind of data. Figure 8.2 shows the front panel of this VI and it shows the trend for each of the variables monitored or controlled. It is friendlier more intuitive data acquisition user-interface version showing the location of every variable in the loop.

Though working together, hardware-wise these VIs are decoupled from each other. This decoupling is necessary because of the differences involved in the VIs sampling and control rates. In general the control scan rate is in the order of tenths of a second, whereas the data acquisition scan rate is typically around 1 sample per second. When you want to save the data at a different scan rate than you are controlling the loop with, problems can arise in two different ways. First, you end up with large data files with unnecessary data variables outputs, that don't vary much. Secondly, some variables need sometime to settle between channel sweeping; otherwise their readings may become unstable.
Also, we are using two different data acquisition cards (DAQ) to acquire data and control the loop. Using one VI to acquire the data and other VI to control the loop caused both VIs to try to access the same channel and/or access the same DAQ at the same time. With this I/O conflict the VIs were not able to run in parallel controlling and acquiring data at different scan rates.
8.3. Tests and Experiments

The first test was to perform a valve-opening-to-flow calibration curve. But, during a working session, a modification to the connection between the two tanks was proposed, since the present setup did not correspond to the mathematical model available, and more importantly, the connection was not appropriate, since with both tanks connected at low flow, the water flowing from one tank goes back to the sump-tank even before going into the other tank. So a small modification was introduced. The original piping connecting both tanks was kept but the manual valve will be kept permanently closed and a ½" internal diameter clear hose linking both tanks was installed with a ball valve in between. Since it drastically modified the loop dynamics, a new valve-opening-to-flow calibration curve will have to be performed to account for the modification. These results are presented below.
8.4. Valve Opening-to-Flow Calibration

To learn how the loop dynamics works and in order to have a valve-position-to-flow mapping, a few tests were carried out. It is important to notice that during all tests both tanks were disconnected, i.e. the manual valve connecting both tanks was kept closed.

The inlet flow control valve (CV1) was gradually opened until the tank level reached its maximum level (900 mmH₂O), or as close to the maximum level as possible. The valve position was changed in a step-like fashion only after the level in tank reached steady state, as can be seen in Figure 8.3, where the upper graph indicates the signal sent to the control valve and the lower graph indicates the control valve position output. It is clear, from Figures 8.4 and 8.5, the outlet and inlet flow and level profiles are similar to the control valve opening position (upper and lower left plots) and the error between inlet and outlet flows after steady state was reached (lower right plot) is well within acceptable range (around 0.003 liters/s, or 0.05 GPM, or 2% F.S.). However, the control valve shows a hysteresis effect when the direction changes from open to close, though one closing point measurement was skipped, and this difference can be easily compensated by software.

![Control Valve 1 Input Voltage](image1)

![Control Valve 1 Output Voltage](image2)

Figure 8.3. Control Valve 2 Input and Output Position Profile.
Figure 8.4. Tank 1 Outlet Flow Profile, Inlet Flow Profile and Net Flow Difference.
In Figure 8.6, a linear fit was made using valve opening as independent the variable and flow rate as the dependent variable (upper plot) and the residuals shown in the lower plot. Figure 8.7 shows the same relationship as the valve was gradually closed, despite the hysteresis present due to the control valve mechanical limitations.
Figure 8.6. Inlet Flow Versus Control Valve-1 Position Showing Linear Relationship.

Figure 8.7. Inlet Flow Versus Control Valve-2 Position Showing Linear Relationship.
Similar tests were repeated using control valve 2 (CV2) and the same results were achieved. These results are shown in Figures 8.8 – 8.12.

Figure 8.8. Control Valve-2 Input and Output Position Profile.
Figure 8.9. Tank-2 Outlet Flow Profile, Inlet Flow Profile and Net Flow Difference.
Figure 8.10. Tank-2 Level Profile.
Figure 8.11. Inlet Flow Versus Control Valve 2 Position Showing Linear Relationship.
Also, a mismatch was observed between inlet and outlet flows. This problem was fixed by performing a quick calibration, and now both tank lines present a maximum flow difference between respective inlet and outlet of around 0.002 liter/second (~0.03 GPM). But it was observed that an overall flow difference, that is \([\text{inlet-outlet flow}]_{\text{TANK-1}}-\ [\text{inlet-outlet flow}]_{\text{TANK-2}}\) of around 0.02 liters/second (~0.3 GPM); although both flow measurements are close, the loop lines cannot be considered 100% symmetric.

8.5. Artificial Time Delay Test

This purpose of this experiment is to help us understand how to handle time-delayed systems. The next sections will describe how this experiment was done and discuss the results.

8.5.1. Determination of the Delay-Free System Dynamic Parameters

To open or close the control valve (CV), an analog input signal voltage is sent from the digital-to-analog converter installed in the computer. The signal ranges from 0-10V (0V closed, 4.5V open), but since the tank is only 900mm high, the opening range is limited to 0 - 4.5V,
enough to almost completely fill the tank, keeping the flow outlet valve opened. So, to determine the system "CV-Tank" dynamic parameters, a step input ranging from 0 to 4.5V was applied to the CV input. Figure 8.13 shows the corresponding level variation to such an input. It is clear that this system has no delay, since the controller, tank and outlet valve are very close to each other.

Using Ziegler-Nichols Step Response Method, a delay-free system model was determined:

\[
\frac{Y(s)}{X(s)} = G(s) = \frac{k}{sT + 1} = \frac{175.7}{140s + 1}
\]  

(8.1)

Where 140s is the time constant (0.632 of the steady state value for a step input, in seconds) and 175.7mmH2O/V is the steady state final value.

8.5.2. Determination of the Delayed System Dynamic Parameters

As stated before, the controller, tank and outlet valve components are very close to each other, so the system has no time-delay. To overcome this problem, since we are interested in time-delayed systems, an artificial delay can be added to the system. This delay can be added either to the controller or to the signal used as input to the PI controller, i.e., the level measurement (see Ref [27-31]).
For this study, a 5 seconds time-delay will be considered, meaning the tank will be artificially pushed over ~5 meters from its present position relatively to the control valve (considering a 0.95 m/s velocity). With this, Equation (8.1) becomes

\[ G_d(s) = \frac{k}{sT + 1} = \frac{175.7e^{-5s}}{140s + 1} \]  

(8.2)

where \( d \) indicates delayed system.

Now it is necessary to choose the stabilizing PI gains for the plant given in Equation (8.1), where the plant parameters are \( k=175.7, L=140 \) and \( T=5 \) seconds. To do so we first have to obtain the range for \( K_p \) values by using the following relationship (Datta, 2000):

\[-\frac{1}{k} < K_p < \frac{T}{kL} \sqrt{\frac{\alpha_1^2 + L^2}{T^2}}\]
\[ \tan(\alpha) = -\frac{T}{L} \alpha \]  
\hspace{1cm} (8.3)

in the interval \( \left( \frac{\pi}{2}, \pi \right) \).

Solving equation (3): \( \alpha = -0.035714 \)

And the \( K_p \) range becomes: \(-0.001265 < K_p < 0.023065\)

Now, \( K_i \) value range can be determined according to the following relationship (Datta, 2000):

\[ 0 < K_i < \frac{T}{kL^2} z \sqrt{z^2 + \frac{L^2}{T^2}} \]

where

\[ \frac{1}{\tan(z)} = \frac{T}{L} z \]  
\hspace{1cm} (8.4)

in the interval \( \left( 0, \frac{\pi}{2} \right) \).

Solving Equation (8.4): \( z = 0.1874 \)

And the \( K_i \) range becomes: \( 0 < K_i < 0.000253 \)

**8.6. Results of Controller Implementation**

After installing the time-delay free PI controller and manually fine tuning it, \( K_p \) and \( K_i \) final values were: \( K_p = 0.011 \) and \( K_i = 0.00026 \) and Figure (8.14) shows the PI controller response to changes in Tank 1 level set-point.
After manually tuning, the time-delay PI controller gains were found to be $K_p=0.011$ and $K_i=0.00013$ and the results are shown in Figure 8.15.

It is seen that in both cases, delay-free and delayed systems, some overshoot is present and some further improvement can be achieved by selecting more appropriate PI values. This way, the VI was changed to use what is called Real Time Systems Package, from National Instruments. This package allows us to use advanced PI control sub-VI's to provide signal filtering, deadband, auto-tuning and signal conversion, among other features, like Fuzzy Control and Neural Network control. One of the most useful features in this package is the "PI Autotuning" sub-VI, which can be used once a stable system is determined using initial PI values. The auto-tuning method used is the relay method. The results for both delay-free and delayed system using auto-tuning are shown in Figure 8.16 and 8.17, respectively, where it can be seen that the overshoot in both cases was dramatically lowered. However, it is also clear there is room for more improvement in the PI controller output.
Figure 8.15. Level Set-point and PI Level Control.
Figure 8.16. Delay-Free Level Setpoint and PI Level Control with Auto-tuned PI Values.

Figure 8.17. Level Set-point and PI Level Control with Auto-tuned PI Values.
8.7. Remarks on the Loop Control Design

For small level amplitude values the overshooting reached 40% of the final value whereas for higher level amplitudes the overshoot lowered to ~8%, which is comparable to the same system but without any delay where the maximum overshoot at low levels was around ~6% and for higher levels was less than 3% and for clarity, these results, as well as the PI values used in this experiment are summarized in Table 8.4. There are a few ways to try to enhance the time-delay PI controllers, which are listed below.

1) Using piece-wise PI value ranges for both low and high level amplitudes.

2) Using piece-wise gain values for both low and high level amplitudes.
Table 8.4. Overshoot Percentage for Delayed and Delay-Free Systems.

<table>
<thead>
<tr>
<th></th>
<th>No Time Delay</th>
<th>With Time Delay</th>
<th>With Time Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Manual Tune</td>
<td>Auto-tune</td>
</tr>
<tr>
<td><strong>Low Level</strong></td>
<td>6%</td>
<td>40%</td>
<td>16%</td>
</tr>
<tr>
<td><strong>High Level</strong></td>
<td>&lt;3%</td>
<td>8%</td>
<td>4%</td>
</tr>
<tr>
<td><strong>Initial Values</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Initial Values</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Auto-tune</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>$K_p$</strong></td>
<td>0.011</td>
<td>0.011</td>
<td>0.0049</td>
</tr>
<tr>
<td><strong>$K_i$</strong></td>
<td>0.00026</td>
<td>0.00013</td>
<td>0.7402</td>
</tr>
</tbody>
</table>

8.8. Control Loop Fault Types

There are several faults that can lead to significant error in the MPC prediction model. Directly or not, the following are the components that can contribute to a specific MPC as input and/or output variables:

- Tank Water Levels $= f($Inlet Flow Rate, Outlet MOVs Position$)$.
- Inlet Flow Rates $= f($Sump tank water level, Bypass valve position, Inlet MOVs Position$)$.
- Outlet Flow Rates $= f($Tank Water Level, Outlet MOVs Position$)$.
- Controller Outputs $= f($Bypass valve position, Tank Water Levels, Outlet Flow Rates$)$.

When the predicted Inlet Flow Rates has significant discrepancy compared with the measured inlet flows, some possible faults may be responsible for this discrepancy:

a) inlet flow meters themselves
b) bypass valve position, and
c) inlet MOV valves.

Whenever any of these devices is faulty, the predicted model will not match closely with the measured inlet flow rates.
One way to emulate such faults, the following well-controlled tests can be designed:

a) Water Level Controller (controller fault).
b) Water level sensor is drifting (sensor fault).
c) Outlet turbine flow meter is drifting (sensor fault).
d) Outlet MOV positioning device is drifting (actuator).
e) Bypass valve position (actuator fault).
f) Inlet MOV positioning device is drifting (actuator fault).
g) Flow meter is drifting (sensor fault).

The Bypass valve can be used in the loop system to simulate changes in the pump speed.

8.9. Controlled Fault Implementation

Running only one tank, the first fault implementation performed was a simple bias value added to tank 1 level sensor. The first test was made using a 5 mmH2O bias and the results are shown in Figures 8.18 and 8.19 (with a zoomed in look). It can be seen that the PI controller was misled by the wrong level.

Next, a 50 mm-H2O positive trend fault was introduced to the level sensor. While level set-point was kept fixed, the trend occurred over a 2-hour period and the results are shown in Figures 8.20 and 8.21. Though a little hard to see, the positive trend caused the PI to counter-react by lowering the inlet flow rate, consequently lowering the water level in tank. Using the same fault trend, a 9-hour period experiment was performed. The results are shown in Figures 8.22 and 8.23.
Figure 8.18. Level Sensor with 5 mm-H2O Fault.

Figure 8.19 Level Sensor with 5 mmH2O Fault (zoomed in).
Figure 8.20. 50 mm-H2O Positive Level Trend Fault.

Figure 8.21. Inlet Flow rate for a 50 mm-H2O Positive Level Trend Fault.
Figure 8.22. 50 mm-H2O Positive Level Trend Fault.

Figure 8.23. Inlet Flow rate for a 50 mm-H2O Positive Level Trend Fault.
Another experiment consisted of adding a negative bias trend of 50 mm-H₂O over a 2-hour time period to tank 1 sensor level while connected to fault-free tank number 2, with both tanks having fixed set-point levels at 600 mm-H₂O. Figures 8.24 and 8.25 show the level changes in the tanks as the fault in tank 1 level sensor changed. From Figure 8.25, it appears the PI control had problems trying to keep the water level at the set point while having a bias trend in tank 1; this may be because of the fault severity in tank 1 over such a short period of time, or some fine tuning needed in tank 2 PI controller. Figures 8.26 and 8.27 show the inlet flow rate in tank-1 increased to compensate for the negative trend, as opposed to tank-2 flow rate.

Figure 8.24. Water Level in Tank 1 with Bias Trend.
Tank 2 Level – 50 mmH2O negative trend in tank 1 over a 2-hour period

Figure 8.25. Water Level in Fault–Free Tank 2.

Tank 1 Inlet Flow

Figure 8.26. Tank 1 Inlet Flow rate.
The control valve outputs are shown in Figures 8.28 and 8.29, whose response trends follow the water level changes in both tanks.
Figure 8.28. Control Valve 1 Voltage Output.

Figure 8.29. Control Valve 2 Voltage Output.
To verify how the PI controllers would react to slow changing, low amplitude faults, 2 more experiments were done. The first experiment consisted of adding a negative bias trend of 20 mmH2O over a 17-hour time period to tank 1 sensor level while connected to fault-free tank number 2. In the second experiment, a positive bias trend of 30 mmH2O over a 24-hour period was added to tank number 2. In both cases, initial set-point levels were fixed at 600 mmH2O. Figures 8.30 and 8.31 show the level changes in the tanks as the fault in tank 1 level sensor changed. In Figure 8.31 it is possible to see that, after fine-tuning tank 2 PI controller, the level in the tank was kept stable at around 600 mmH2O for the entire experiment, though this was a slow changing fault.

Figures 8.32 and 8.33 shows tank 1 inlet flow rate increasing, while there is a slight decrease in tank 2, to compensate for the negative trend in the first tank.

![Figure 8.30. Water Level in Tank 1 with 17-hour Bias Trend.](image)
Figure 8.31. Water Level in Fault–Free Tank 2.

Figure 8.32. Tank 1 Inlet Flow rate.
Figure 8.33. Tank 2 Inlet Flow rate.

Finally, the results of the second experiment are shown in Figures 8.34 – 8.37 where, among other things, it is possible to see an increase in tank 1 inlet flow rate, and a decrease in tank 2 to compensate for the negative trend in tank 2.
Tank 2 Level - 30 mmH2O positive trend over a 24-hour period

Figure 8.34. Water Level in Tank 2 with 24-hour Bias Trend.

Tank 1 Level

Figure 8.35. Water Level in Fault–Free Tank 1.
Figure 8.36. Tank 1 Inlet Flow rate.

Figure 8.37. Tank 2 Inlet Flow rate.
9. MATLAB/SIMULINK IMPLEMENTATION OF MODEL PREDICTIVE AND FAULT-TOLERANT CONTROLLERS FOR THE SP-100 SYSTEM

9.1. MPC MATLAB Toolbox

MATLAB 7.0 has a Model Predictive Control toolbox, which facilitates a convenient design of the controller. It is a collection of software routines that helps the user to design, analyze, and implement an advanced industrial automation algorithm. It provides a convenient graphical user interface as well as a flexible command syntax that supports customization.

The model Predictive Control Toolbox provides a special controller block, if a Simulink graphical tool is used to model the plant. First, linearize the nonlinear Simulink model. Then use the linearized model to build a Model Predictive Control Toolbox controller, and evaluate its ability to control the nonlinear model.

A Model Predictive Control Toolbox design requires a plant model, which defines the mathematical relationship between the plant inputs and outputs. The Model predictive Control Toolbox requires the model to be linear and time invariant (LTI). There are three ways to define such a model:

- Create a transfer function, state space, or zero/pole/gain model using methods provided by the Control System Toolbox.
- Derive the model from plant data using, for example, methods provided by the System Identification Toolbox.
- Derive the model by linearizing the Simulink model.

Most real systems are nonlinear. If the user wants to simulate Model Predictive Control Toolbox control of a nonlinear system, the plant must be modeled in Simulink. The SP-100 reactor system is a very complex nonlinear system. Often, it is not easy to create a transfer function or a state space model. A Simulink model facilitates the easy inclusion of nonlinear dynamics. It is easier to generate the LTI model by linearizing it, and to control the system defined by the nonlinear Simulink model.
9.1.1. Linearization

Although a controller designed using the Model Predictive Control Toolbox can regulate a nonlinear plant, the model used to compute the control step must be linear. The accuracy of this approximation is a key issue affecting the controller performance. The usual approach is to linearize the nonlinear plant at a specified operating point. The Simulink environment provides two ways to accomplish this:

- Linearization using Simulink Functions—‘linmod’.
- Linearization using Simulink Control Design.

9.1.2. Linearization using Simulink Functions

A Simulink block can be linearized using standard Simulink functions. The variables to be retained in the linearized model must be connected to an ‘inport’ or an ‘outport’ block.

Figure 9.1 shows a scheme of SP-100. In this project, the goal is to regulate the electric power using the control drum angle as the control action; hence, there is only one ‘inport’ block designating the input signal (control drum angle), and one ‘outport’ block designating the output (electric power output).

![Simulink model of nonlinear SP-100 reactor system.](image)

This model is named sp100. The ‘linmod’ command linearizes it as follows:

```
[a,b,c,d]=linmod('sp100')
```
By default, ‘linmod’ uses the initial conditions defined in the model as the operating point. Here the operating point corresponds to the values at the rated power of 2MWt. The LTI model is now created. For the MPC Toolbox, a state space model is needed; it can be created by using the following command:

$$\text{Plant}=\text{ss}(a,b,c,d)$$

### 9.1.3. Linearization using Simulink Control Design

Simulink Control Design is an optional product that supports model linearization.

![Simulink model of nonlinear SP-100 reactor](image)

**Figure 9.2. Simulink model of nonlinear SP-100 reactor**

To linearize the system represented by Figure 9.2, first define the input and output signals to be retained in the linear approximation. Next, create a linearization project within the Simulink Control and Estimation Tools Manager. Using this toolbox, the Simulink model can be linearized about the default operating point (inputs values equal to zero), or can be linearized at a specified operating point. Simulink Control Design can search for a new steady state operating point.

### 9.1.4. Linearization Result

Figure 9.3 shows the step response of the electric power output. The blue line is the linearization model plot, and the red line is the step response generated by the Simulink model. They are almost the same; thus, the linearized model is adequate in representing the above nonlinear Simulink model.
9.2. The Design Tool

The Model Predictive Control Toolbox design tool is a graphical user interface for control design. The initial view of the MPC Toolbox is shown in Figure 9.4. In this user-interface, the linearized plant model has been loaded. The input is the control drum angle, which is set as the manipulated variable. The output is the electric power, which is set as the measured output. By doing this, the control drum angle can be adjusted by the measurement of the electric power output. Figures 9.5 (a) and (b) show the simulation results with the default controller setting. These often work well, but there is still a small overshoot in the input. The procedure for tuning a controller for better performance is discussed in the following section.

Figure 9.3. Step Response of electric power output.
Figure 9.4. Model predictive control toolbox design tool initial view.

Figure 9.5a. Plant output with default controller setting.
9.3. Controller Setting

When selecting the MPC1 controller in the tree in the left window, the screen shown in Figure 9.6 should appear. The controller setting includes the following topics:

- Model and horizons
- Defining manipulated variable constraints
- Weight tuning
- Disturbance modeling and estimation.
In ‘Model and Horizons’ tab, the user can change the prediction horizon, control horizon and select the Blocking (an alternative to penalty weighting) option to get more control over the way in which the controller's moves are allocated.

In the ‘Constraints’ tab, the user can set the manipulated variable constraints. Physical devices have limited ranges and rates of change. In this project, the control drum angle's maximum rate of change is ±1.4 degree per second. Without these limits, the controller might attempt an unrealistic adjustment just like that shown in Figure 9.5b.

In the ‘Weight Tuning’ tab, the user can change the input weights, output weights, and overall rate weights. The accuracy with which each output must track its set point is dictated by setting the output weights. A large weight on a particular output causes the controller to minimize deviations in that output. The use of a nonzero input weight forces the corresponding input back towards its nominal value. For overall rate weights, the smaller weight forces the
controller to make smaller, more cautious adjustments, but takes a slower approach to reach the set point.

In the ‘Estimation’ tab, we can set the disturbance of both input and output. The Model Predictive Control Toolbox allows users to tailor the controller's disturbance response.

### 9.4. Results of MPC Controller Implementation

Three tests with different set of MPC parameters were mentioned in this section.

**a) Test 1**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction horizon</td>
<td>10</td>
</tr>
<tr>
<td>Control horizon</td>
<td>2</td>
</tr>
<tr>
<td>Overall weight</td>
<td>0.3</td>
</tr>
<tr>
<td>Input weight</td>
<td>0.3</td>
</tr>
<tr>
<td>Output weight</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 9.7(a) shows the controller performance when the electric power is set to increase by 4 units at the 10th second. Figure 9.7(b) shows how the control drum angle is adjusted with the increase and decrease constraint of 1 degree/sec. The controller with the above parameters settings has a very slow response. The set point reaches at around 100 seconds.
Figure 9.7a. Plant electric power output.

Figure 9.7b. Plant input (drum angle).

b) Test 2

| Prediction horizon | 10 |
By just increasing the overall weight from 0.3 to 0.8, we can get the controller performance shown in Figures 9.8 (a) and (b). This controller has a very quick response, but there is a large overshoot in the input, which will make the SP-100 system less stable.

<table>
<thead>
<tr>
<th>Control horizon</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall weight</td>
<td>0.8</td>
</tr>
<tr>
<td>Input weight</td>
<td>0.3</td>
</tr>
<tr>
<td>Output weight</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 9.8a. Plant electric power output.
c) Test 3

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Prediction horizon</td>
<td>10</td>
</tr>
<tr>
<td>Control horizon</td>
<td>2</td>
</tr>
<tr>
<td>Overall weight</td>
<td>0.5</td>
</tr>
<tr>
<td>Input weight</td>
<td>0.3</td>
</tr>
<tr>
<td>Output weight</td>
<td>1</td>
</tr>
</tbody>
</table>

By changing the overall weight to 0.5, and setting the initial values of the input and output to be the values at which the SP-100 model is linearized, the controller performance shown in Figures 9.9 (a) and (b) is achieved. This controller has a better performance than case (b). The control drum angle changes more smoothly, and the electric power output reaches the set point at around 25 seconds. This is a much better controller. It is saved as mpc_sp100, and is used to control the nonlinear plant.
9.5. Controller Performance for the Nonlinear Plant

The MPC controller has a good performance using the linearized model. This is connected to the nonlinear Simulink model to control the SP-100 system. Figure 9.10 shows the
SP-100 with an added MPC Controller block. The MPC controller is the one that is designed using the linearized model. The set point for the MPC block ref input can be any value or trend of electric power.

Figure 9.10. Simulink model for SP-100 controller.

Figure 9.11 shows the comparison of the electric power output between the set values and the actual values. The electric power output increases by 10 kW with a ramp 1kW/sec at the 10th second, and then decreases by 20 kW, with a ramp 1kW/sec at the 80th second. The actual values match the set values very well. The response of the control drum angle is shown in Figure 9.12. There is a small overshoot at the end of each change. The reactor thermal power, the fuel temperature, cladding temperature, and the core coolant temperature changes are shown in Figures 9.13 and 9.14. The reactor thermal power change is similar to the control drum angle change, which has a large overshoot at the end of each change. This is because the temperatures change is much smoother and slower, which makes the feedback reactivity change slower; but, the reactivity insertion has a quicker response, so it needs quite a while for the feedback reactivity to compensate the external reactivity insertion. Therefore, the reactor thermal power change has a larger overshoot.
Figure 9.11. The electric power output.

Figure 9.12. Control drum angle response.
9.6. Summary

The MPC Toolbox makes it a lot easier and more convenient to design a MPC controller for the SP-100 system. It provides a convenient graphical user interface (DUI). All the changes and comparison can be down under the same interface. It shortens the time for coding and tuning the
program. The MPC Toolbox also provides a Simulink block the users can use to control a plant modeled in Simulink. In this chapter, we are using MPC Simulink block to control the Simulink model of SP-100. The procedures for implementing MPC to SP-100 Simulink system is shown in the below flow chart figure 9.15.

![Flow chart of MPC implementation](image)

The advantage for using the MPC block to control a Simulink system is: at each time step, the nonlinear system will be linearized based on the last step inputs and outputs. So the linearized system will be updated at each step. It will give more precise predictions than the empirical model developed basing on one group of observations.

### 9.7. Fault Tolerant Control

When a single failure occurs, the FDI system will point out the nature of the failure. In this chapter, the control strategies about how to handle the single failure are discussed.

#### 9.7.1. System failures

Generally there are three types of failures may occur in a control system: actuator failure, ‘internal’ failure and sensor failure.
1. Actuator failure: the range of the actuator is reducing, or the movement of the actuator is getting slower or faster, or the actuator is stuck in a special position. Actuator failures are hardware failures. They would not change the dynamics of the system. For our study case, the SP-100 space reactor system we designed is a single-input and single-output system. The input is the control drum angle. Here we assume the control drum was stuck in a special position, and could not move further more. Also, we need a secondary actuator which can change to compensate the loss of major actuator failure. The TE power output can be adjusted by changing the value of control drum angle, coolant flow rate or shunt resistor. Here we use shunt resistor as the backup control actuator. When the control drum actuator failure is detected, the control system will suspend the MPC controller which is used to control the control drum angle, and startup the secondary PI controller which is used to control the shunt resistor.

2. ‘Internal’ failure: some parts of the plant fail, with the consequence of significant changes of the plant dynamics and gains. In this project, the case we studied is the calibration for control drum angle corresponding to the insertion of external reactivity has changed.

3. Sensor failure: some measurements become unavailable, or incorrect, or unusually noisy. For some sensor failures, they may not cause any trouble for the control system, but for the sensor, whose measurement will go to the control algorithm, its failure will cause the control action to deviate from the desire result.

In this section, each of the three types of failures will be considered separately, and the corresponding control strategies are designed to compensate the loss caused by each failure.

9.8. Strategies for Fault Tolerant Control

9.8.1. Actuator Failure

It is assumed that the control drum motion is stuck at a given position. It can only change its angle from $0^\circ$ to $130^\circ$, and the corresponding electric power output is 0 kW to 128 kW. It is impossible to reach the power demand beyond this range by only controlling the control drum angle. A reconfigurable control strategy is developed to solve this kind of problem. A sub-
controller is started-up if the electric output remains unchanged for 30 seconds and there is still a big gap between the target value and the real value. This sub-controller is a PI controller, which regulates the electric power output to the desired power by controlling the shunt resistor value.

The performance of the reconfigurable control is illustrated in Figure 9.16. Figure 9.16a shows performance of the SP-100 space reactor system under the actuator failure. We can see that when the reactor desired power is changing from 112.53 kW to 120 kW, it is still within the adjustable range of the control drum angle. So just adjusting the control drum angle, we can reach the aim point. At the 300th second, the desired electric power changes from 120 kW to 132 kW, which is out of the range that control drum angle can cover. Thus, when the control drum angle is stuck at its extreme position, the TE power output will remain at 128 kW, although there is still a big gap between the desired value and the actual value. Figure 9.16b shows the system performance with the reconfigurable control. When the actual electric power remains unchanged for 30 seconds, and there is still a big difference between the target value and actual value, an alarm signal is sent to a subsystem. A PI controller is activated to control the shunt resistor to trace the target value. From this figure, we can see the sub-controller is enabled around 250th second, and the actual electric power output will reach the desired the value in around 100 second. Figure 9.16c shows the change of shunt resistor. A small change of the shunt resistor will cause of big change of the TE power output. It is not a very stable and easy manipulative way to control the TE power output. Within a limited range, as a secondary control method, its performance is acceptable.
Figure 9.16a. System performance without reconfigurable control.

9.16b. System performance with reconfigurable control.
9.16c. Change of the shunt resistor.

Figure 9.16. Performance of the reconfigurable control

9.8.2. 'Internal' Failure

In this work, the assumed 'internal' failure is the change in the gain of the external reactivity of the control drum angle, starting at the 400th second point. For the normal condition, 5° change of the control drum angle approximately inserts or takes out 0.1 cent of reactivity. As the 'internal' failure occurs, 5° change of the control drum angle approximately inserts or takes out 0.09 cent of reactivity. Figure 9.17 shows the comparison of the electric power output with and without the MPC. Note that without MPC the electric power is much lower than the desired power. With MPC, the condition is improved, although it does not fully meet the desired values.
9.8.3. Sensor Failure

With the above knowledge, a PCA model representing the normal operating conditions is developed to detect the sensor failure. The objective of sensor failure detection is to diagnose sensor degradation at the incipient stage, with a very small probability of making a wrong decision. When a measurement fault is detected in the TE power sensor, an alarm signal is sent to the system. The MPC controller will switch its control input from the difference between the reference value and measurement value to the difference between the reference value and the predictive value. The predictive value of the TE power output is estimated by the empirical model, which is built with the normal operation data. Figure 9.18a depicts the step response of the empirical model and the step response of the non-linear Simulink model. It can be seen that the empirical model can provide a very good prediction of the system. The performance of the SP-100 system when the TE power sensor failure occurs at the 300th second is illustrated in Figure 9.18b. Without any interference, the actual electric power will run further away from the target values because the degrading fault is getting larger. Figure 9.18c shows the FDI detects the sensor failure at 450th second, and sends alarm signal to the SP-100 system. So the input signal of the MPC controller will change to track the difference between the estimated value and
reference value. Since the empirical model to make the estimation is good enough to present the system performance, the control results are much better improved.

(a) Comparison between the step response of the empirical model and the Simulink model

(b) Performance of the SP-100 system under sensor failure
9.8.4. Summary

In this chapter the control strategies are designed to handle different single failures. A reconfigurable control is designed to solve the actuator failure. A secondary PI controller is used to compensate the loss of the actuator failure. For a well-tuned MPC controller, it has the capability of reducing the effect of dynamics change caused by 'Internal' failure. To solve the sensor failure problem, a model is needed to estimate the measurement values during the process. When the sensor failure is detected, instead of using the measurement value, the model's estimate value will go to the control algorithm.

By far the control strategies we designed can solve the problems caused by single specified failure. For the actual situation, one single failure may cause consequent failures. To realize autonomous control, the control strategies to handle multi-failure need to be considered in the future research work.
10. IMPLEMENTATION OF MPC IN THE EXPERIMENTAL LOOP
USING EMPIRICAL INPUT-OUTPUT MODELS

10.1. General Description

The two-tank flow rate control loop is built to test the control strategies which have been applied to the SP-100 space reactor system.

The schematic of the two-tank loop is shown in Figure 10.1. CV-1 and CV-2 are the two control valves. By sending different control voltages, the control valves open or close, so that the flow rate of the water flows to tank-1 and tank-2 change as well. The other valves are manual valves. They are kept fully open during the normal operation condition. The maximum flow rate can reach 0.2 L/sec.

![Figure 10.1. Schematic of two-tank flow rate control loop.](image)

16 different state variables are measured and recorded. They are: water level in tank 1, water level in tank 2, inlet flow rate to tank1, inlet flow rate to tank2, outlet flow rate from tank 1, outlet flow rate from tank 2, bypass flow rate, the water temperatures in tank1, tank 2, sump tank
and pump, the voltages from control valves 1, 2 and 3. Usually, the change of temperatures is very small. It is not an important factor that will affect the water levels, so we will not consider the effect of temperatures.

Figure 10.2 shows the photograph of the experimental loop at the University of Tennessee. The yellow-colored devices are the control valves. There are three control valves in the real experimental loop. The third control valve is used to control the flow rate of the water flows out of tank 2. It is kept fully open during the current experimental design. It will work as a redundancy control method in the future research work about fault tolerant control. The blue parts at the right bottom corner are the pressure meters. The pressure signals will be converted to voltage signals and sent to the computer. There are also two orifice meters to measure the inlet flow rates and two turbine meters to measure the outlet flow rates.

Figure 10.2. The experimental loop.

Two data acquisition cards are used to acquire loop process data. One set of data is sent to the LabVIEW VI simulation model to monitor the current operation situation. One set of data is sent to a notepad file for future data analysis. The front panel of the VI is shown in Figure 10.3.
When the two-tank flow rate control loop has been well tuned, and all the software and hardware are working properly, the PI controllers were first implemented to the loop. The PI controllers will drive the water levels to the desired positions. The data under normal operation condition will be acquired, which will cover the water level range from 300 mm to 800 mm.

The whole experiments that will do to the loop are listed in the flow chart as Figure 10.4. First we acquire the normal operation data; then use the data to build an empirical model; design the MPC based on the empirical model, then implement MPC online to make sure the MPC will work fine on the non-linear real-time loop. The above procedures guarantee the two-tank loop with the basic control functions, that the water levels will arrive at the desired points. Then a fault diagnosis system is needed to monitor the operation situation of the two-tank loop and send alarms to the operator if there is any abnormal situation happens. The fault diagnosis system will base on the Principal component analysis (PCA) technique. The data under normal operation condition are acquired to develop the PCA model. Then the control strategies to handle the abnormal situations are needed. So when the abnormal data go to the control system, the fault
diagnosis system will point out the nature of the fault, and the fault tolerant control system will take some actions to compensate the loss caused by the fault, and make the whole system still satisfy the water level requirements.

In this chapter, we are focusing on the implementation of MPC controllers to the real flow rate control loop.

10.2. Control Problem Description

For model predictive control, the more precise the model is, the better the performance of the controller will be. We have built a Simulink model to simulate the dynamics of the two-tank experiment loop, but because there are a lot of simplifications in the dynamic dominating equations, the Simulink model could not represent the dynamics of the real two-tank loop. In the design of MPC controller for the two-tank loop, we are using the empirical model which is developed with the data acquired from the running loop under normal operation condition.
Three aspects need to be considered to achieve a good performance of MPC control system:

1. The improvement of empirical model. The original data set could not be used to develop an empirical model directly. There will be big vibrations in the data when the water level is low. When the water level is high, the high pressure caused by the height of water column will affect the inlet flow rate, so the dynamic of the system will be slightly changed. It is very important to choose the correct data region. Here we choose the data corresponding to the water level between 400 mm to 700 mm. Also different orders and different combination of the orders of the Auto Regression Moving Average with eXogenous inputs (ARMAX) model need to be tested to find the best fit.

2. The parameters setting for MPC. There are several parameters need to be adjusted in the MPC Toolbox, such as the control horizon, prediction horizon, input weight, output weight, overall weight, control constraints, et al. When a MPC controller with specified parameters is developed, it needs to be implemented on-line to the real loop to check the real performance of the MPC controller. Several MPC controllers with different designs need to be tested on the real loop, until we can find one which satisfies all the control design requirements.

3. The correct way to implement MPC online. There were several papers about how to online implement MPC. Special code needs to be run to hook up MATLAB with LabView. The MPC controller is running under MATLAB environment, and the control movement generated by the MPC controller will go to LabView to drive the motors of control valves, and the new values of the current system state variables will go back to MPC controller to generate the next step control decision.

The final goal is to implement MPC control for the two tank loop. Before working on the MPC controller for the two-tank loop, we started with a simpler case—MPC implementation on single tank. The experience of implementing MPC to single tank loop is very valuable for the implementation of MPC to Two-tank loop.
10.3. MPC Implementation in Single Tank Loop

Figure 10.5 shows the data acquired to build the empirical model. We changed the water level from 300mm to 800 mm, and the changing step is 50mm. The PI controller is used to direct the water level to the desired values. We can also acquire data with manual control by changing the control voltage sending to the control valve. But in this way it will take longer time for the water level to reach the steady state, and some of the data may just be the same value. So in this way the data may not carry a lot of information about system dynamics. The PI controller we used is not well tuned. At the beginning of each water level change, there will be a big jump of the inlet flow rate.

![Graph of water level and flow rate](image)

Figure 10.5. The data acquired for single tank loop with PI controller.

The entire data set could not be directly used to develop the empirical model. We just choose the data within the water level region from 400mm to 700 mm to build an empirical model. The MATLAB command ‘armax’ is used to generate the empirical model. ARMAX stands for Auto Regression Moving Average with eXogenous inputs. The model orders need to be specified.
Different combinations of model orders are tried to generate the empirical model. The best fit is with model order [10, 4, 1, 1]. The empirical model can fit the data up to 99%.

That SISO model is defined as follows:

$$A(q)y(t) = B(q)u(t) + C(q)e(t)$$

Where, 

$$A(q) = 1 - 1.946 q^{-1} + 1.096 q^{-2} - 0.2593 q^{-3} + 0.04471 q^{-4} + 0.06736 q^{-5} - 0.03803 q^{-6} + 0.04998 q^{-7} - 0.01116 q^{-8} - 0.004916 q^{-9} + 0.002253 q^{-10}$$

$$B(q) = 72.64 q^{-1} - 74.6 q^{-2} + 20.19 q^{-3} - 18.23 q^{-4}$$

$$C(q) = 1 - 0.9966 q^{-1}$$

and [10 4 1 1] stands for the order of A(q), order of B(q), order of C(q) and the step size.

The performance of MPC controller for the single tank loop is shown in Figure 10.6. The constraints of the MPC controller are:

$$\begin{align*}
0 < u < 0.2 & \quad L/sec \\
-0.01 < \Delta u < 0.01 & \quad L/sec \\
0 < y < 900 & \quad mm
\end{align*}$$

A dead-band is set for the flow rate perturbation. If the change of water level from current value to next step value is smaller than 5mm, there is no change of the inlet flow rate. This ensures that the control valve does not move frequently to follow any small changes in the water level.
10.4. MPC implementation in two tank loop

Because the two-tank loop is very nonlinear, it is not easy to develop a linear empirical model to represent the dynamic of the two-tank loop. Figure 10.7 shows the data that were acquired for the two-tank loop.
Figure 10.7 Data acquisition for two-tank loop.

Only a part of the data was selected to develop an empirical model.

The MATLAB command ‘pem’ can be used to develop a Multi-Input and Multi-Output (MIMO) model. After trying several orders of the MIMO model, when the order is 5, the MIMO model fits the data best. Figure 10.8 shows the fitness of the model. The MIMO model can fit the data around 99%.
Figure 10.8. Comparison between MIMO model and measurement data.

The MPC controller for two-tank loop was developed basing on the above MIMO model. Several MPC controllers with different parameters designs have been tested online on the two-tank loop, but none of them gave a good results for two-tank water level control.

The observations of the performance of former MPC controllers show that the water level in tank 1 can follow the target values quite well, but the water level in tank 2 could not track the set points. The probable reason is the control valve which controls the inlet flow rate for tank two is not functioning properly. In order to find a good MPC design for two-tank loop, we may need to change that control valve.

10.5. Summary
The MPC controller for the single tank loop has been successfully implemented in the real-time loop. It shows a better performance over the traditional PI controller. By setting proper
constraints, the MPC controller can reach an optimal control. The improvement of the MPC controller for the two-tank case will continue as part of the future work.
11. CONCLUDING REMARKS AND FUTURE WORK

11.1. Summary and Concluding Remarks

As stated in the Executive Summary, the three major tasks of developing the SP-100 space reactor nodal model and transient simulations, development and implementation of classical and model-predictive controller to SP-100, development of fault detection and isolation and integration for the design of fault-tolerant controller, and finally the on-line demonstration of the control strategies with application to a multivariate experimental flow control loop have been accomplished. The specific accomplishments are restated below for easy reference:

- Development and testing of the SP-100 reactor system dynamics and the power conversion module (thermo-electric conversion) nodal model as an integrated system.

- Development and application of the Model Predictive Control (MPC) algorithm to the SP-100 system. The SP-100 reactor simulation model has been used in this task. The MPC algorithm has been evaluated against traditional controllers, for various transient conditions.

- Development of a fault detection and isolation module for monitoring incipient faults in various field devices. This method uses the Principal Component Analysis (PCA) and data-driven models for fault monitoring.

- Design and completion of a laboratory multivariate water level control test loop, for evaluating the various control strategies and the fault-tolerant control method.

- Development of a multivariate flow control loop and on-line demonstration of the classical proportional-integral controller and the model-predictive controller (MPC) using the laboratory test loop.

- Development of an autonomous control framework, with control mode reconfiguration and hierarchical control strategy.

- Presentation of seven papers in national and international conferences, and publication of four manuscripts in refereed journals.

The features of an autonomous control strategy, such as controller auto-tuning and fault-tolerant control features have been demonstrated in this project. The development of methods and their implementation to both reactor simulation models and to a laboratory control test loop have applications to terrestrial reactors. These include both current and next generation commercial power reactors, and modular grid-appropriate reactors for remote deployment. It is
anticipated that the technology developed under this NEER project will add to the autonomy of reactor operation and maintenance, uninterrupted operation of long-duration fuel cycles, and life extension of currently operating reactors.

11.2. Recommendations for Future Research

Based on the research and development performed under this NEER grant, the following research areas are recommended for further work.

- A reliable and robust autonomous control requires proper integration of basic control and fault monitoring functions, a control supervisor, and a system executive. It is recommended that this integration effort be pursued as part of future work.

- Autonomous control often requires increasing the knowledge base by continuous learning. This learning scheme and knowledge use must be developed and integrated into the control framework.

- Development of alternative control mode selection (also referred to as control reconfiguration).

- Demonstration of the integrated system using a multivariate laboratory control experiment.

- Evaluation of the effect of actuator dynamics in on-line control strategies.

- Technology transfer from the aerospace and robotics industry to the incorporation of operational autonomy in nuclear power reactors.

- Evaluation of the effect of digital instrumentation and control in enhancing the operation and maintenance of both large and modular reactors for power generation and other applications.
REFERENCES


BIBLIOGRAPHY


[81]. English, R.E., Evolving the SP-100 reactor in order to boost large payloads to GEO and to low lunar orbit via nuclear-electric propulsion , Springfield, 1991.


[135]. Kuhn, D. R., Fault classes and error detection in specification based testing, Gaithersburg, MD, 1998.


[174]. Overschee P.V., and B.D., Moor, Closed-Loop Subspace Identification, Esat-Sista/Tr 1996.


[216]. Seo J.T., SNPSAM-Space Nuclear Power System Analysis Model, Theses, University of New Mexico, 1986.


[278]. Zhao, Ke, Development of a data driven multiple observer and causal graph approach for fault diagnosis of nuclear power plant sensors and field devices, M.S. Thesis, Department of Nuclear Engineering, The University of Tennessee, 2002

APPENDIX A

Fault-Tolerant Control of the SP-100 Space Reactor

M.G. Na and B.R. Upadhyaya

Submitted for Publication in IEEE Transactions on Control System Technology (May 2006)

Abstract

The control system of an SP-100 space reactor is a key element of space reactor design to meet the space mission requirements of safety, reliability, survivability, economics, and life expectancy. In this work, a fault-tolerant controller (FTC) consisting of a model predictive controller and a fault detection and diagnostics algorithm for the input and output measurements for the controller is developed to control the thermoelectric (TE) power in the SP-100 space reactor. A fault-tolerant controller makes the control system stable and retains acceptable performance even under system faults. The future TE power is predicted by using the fuzzy model identified by a subtractive clustering method of a fast and robust algorithm. The objectives of the proposed model predictive controller are to minimize both the difference between the predicted TE power and the desired power, and the variation of control drum angle that adjusts the control reactivity. Also, the objectives are subject to maximum and minimum control drum angle and maximum drum angle variation speed. The genetic algorithm that is effective in accomplishing multiple objectives is used to optimize the model predictive controller. The model predictive controller is equipped with a fault detection and diagnostics algorithm so that the controller can work properly even under input and output measurement faults. A lumped
parameter simulation model of the SP-100 nuclear space reactor is used to verify the proposed controller. Simulation results of the proposed controller show that the TE generator power level controlled by the proposed controller could track the target power level effectively even under measurement faults, satisfying all control constraints.

**KEY WORDS**: Fault detection and diagnostics, fault-tolerant control, fuzzy model, genetic algorithm, model predictive control, space reactor power control, sequential probability ratio test, subtractive clustering method

1. **Introduction**

The SP-100 space reactor is designed to provide a realistic and reliable source of long-term power for space exploration and exploitation activities. The SP-100 system is a fast spectrum lithium-cooled reactor system with an electric power rating of 100 kW [1]. The primary heat transport system consists of the working fluid, liquid lithium, and electromagnetic (EM) pumps. The energy conversion system uses a direct thermoelectric (TE) conversion mechanism.

The control system is a key element of space reactor design to meet the space mission requirements of safety, reliability, survivability, economics, and autonomous action. In order to accomplish a space mission with uncertain environment, rare events, and communication delays, all the control functions must be achieved through a sophisticated control system with a limited degree of human intervention from the earth. The control functions needed for SP-100 and other space reactors can be ensured only by an autonomous control system, which assumes the responsibilities for normal control, abnormal event response and fault tolerance, and provides an interface with operators on earth for high-level decision-making. Many studies have been conducted to control the SP-100 space reactor [2-5]. These studies focus on normal control of the reactor. This work deals with the normal and fault-tolerant control for the autonomous control system of the SP-100 space reactor.

In order to optimize the reactor power control performance, techniques for the optimal power control of nuclear reactors have been studied extensively in the past two decades [6-9]. But it is very difficult to design optimized controllers for nuclear systems because of variations in nuclear system parameters and modeling uncertainties, and in particular, for the long-term operation of the SP-100 space reactor.
This work employs the model predictive control method, which has received increased attention as a powerful tool for the control of industrial process systems [10-16]. It incorporates the fuzzy modeling for the output prediction of the model predictive control and combines the fault detection and diagnostics for fault tolerance of which the combined controller is a fault-tolerant controller. The basic concept of the model predictive control is to solve an optimization problem for a finite future at the current time. Once a future input trajectory is chosen, only the first element of that trajectory is applied as the input to the plant, and the calculation is repeated at each subsequent instant. This method has many advantages over the conventional infinite horizon control because it is possible to handle input and state (or output) constraints in a systematic manner during the design and implementation of the control. In particular, it is a suitable control strategy for nonlinear time varying systems because of the model predictive concept. Recently, the problem of controlling uncertain dynamical systems has been of considerable interest to control engineers. The model predictive control method has been applied to a nuclear engineering problem [17-18].

In this work, the fuzzy model is used to predict the future output that should be calculated at first to minimize the control cost function. That is, at the present time the behavior of the process over a prediction horizon is considered and the process output to changes in the manipulated variable is predicted by using a fuzzy model based on a subtractive clustering method of a fast and robust algorithm. Based on the predicted future output, a model predictive controller is designed to minimize the output error that means the difference between the predicted output and the setpoint. In addition, another fuzzy model estimates the input and output of the control system by using other process signals, and the errors between the estimated signals and the measured signals are used to determine the health of the measurement instruments by using the sequential probability ratio test (SPRT). A lumped parameter simulation model of the SP-100 space reactor is used to verify the proposed FTC for a space nuclear reactor.

2. Model Predictive Control Combined with a Fuzzy Model

The model predictive control (MPC) is combined with the fuzzy model based on the subtractive clustering approach. The model predictive controller combined with a fuzzy model is called a fuzzy model predictive controller. The MPC method is to solve an optimization problem
for a finite future at current time and to implement the first optimal control input as the current control input. The procedure is then repeated at each subsequent instant. Figure 1 shows the basic concept of the model predictive control [12]. That is, for any assumed set of present and future control moves, the future behavior of the process outputs can be predicted over a prediction horizon $L$, and the $M$ present and future control moves ($M \leq L$) are calculated to minimize a quadratic objective function. Although $M$ control moves are calculated, only the first control move is implemented. At the next time step, new values of the measured output are obtained, the control horizon is shifted forward by one step, and the same calculations are repeated.

![Fig. 1](image)

The purpose of taking new measurements at each time step is to compensate for unmeasured disturbances and model inaccuracies, both of which make the measured system output to be different from the one predicted by the fuzzy model. At every time instant, model predictive control requires the on-line solution of an optimization problem to compute optimal control inputs over a fixed number of future time instants, known as the time horizon. The basic idea of model predictive control is to calculate a sequence of future control signals in such a way that it minimizes a multistage cost function defined over a prediction horizon. A performance index for deriving an optimal control input is represented by the following quadratic function:

$$J = \frac{1}{2} \sum_{k=1}^{L} \left[ \hat{y}(t + k | t) - w(t + k) \right]^2 + \frac{1}{2} \sum_{k=1}^{M} R[\Delta u(t + k - 1)]^2,$$

subject to constraints

$$\left\{ \begin{array}{l} \Delta u(t + k - 1) = 0 \quad \text{for} \quad k > M, \\ u_{\min} \leq u(t) \leq u_{\max}, \\ -du_{\max} \leq \Delta u(t) \leq du_{\max}. \end{array} \right.$$
control horizon, respectively. The prediction horizon represents the limit of the instant in which it is desired for the output to follow the reference sequence. The constraint, \( \Delta u(t+k-1)=0 \) for \( k > M \), means that there is no variation in the control signals after a certain interval \( M < L \), which is the control horizon concept. \( u_{\text{min}} \) and \( u_{\text{max}} \) are the minimum and maximum values of input, respectively, and \( \Delta u_{\text{max}} \) is a maximum allowable control move per time step.

In order to obtain control inputs, the predicted outputs have to be first calculated by using a fuzzy model, in which the inputs consist of past values of control system inputs and outputs, and of assumed future control system input signals. The fuzzy model has been widely used for function approximation problems and has been proved to have good characteristics. Therefore, the fuzzy model is used to predict the future output based on past inputs and outputs; this approach is described in the following section.

### 2.1. Output Prediction Using a Fuzzy Model

In this work, a fuzzy model based on subtractive clustering (SC) is used to predict the future output of the model predictive controller. The \( i \)-th fuzzy rule for \( k \)-th time instant data is described as follows:

\[
\text{If } y(k-d-1) \text{ is } A_{i,j}(k) \quad \text{AND} \quad y(k-d-n_y) \text{ is } A_{i,n_y}(k) \quad \text{AND} \quad \Delta u(k-1) \text{ is } A_{i,n_y+1}(k) \quad \text{AND} \quad \Delta u(k-n_u) \text{ is } A_{i,n_y+n_u}(k),
\]

then \( \hat{y}_i(k) \) is \( f_i\left(y(k-d-1), \cdots, y(k-d-n_y), \Delta u(k-1), \cdots, \Delta u(k-n_u)\right) \),

where

\( A_{i,j}(k) \) = membership function value of the \( j \)-th input variable for the \( i \)-th fuzzy rule,

\( \hat{y}_i(k) \) = estimated output for the \( i \)-th fuzzy rule,

\( d \) = time delay which is equal to or greater than the prediction horizon \( L \),

\( n_y \) = number of time points used for output signals,

\( n_u \) = number of time points used for input signals.

The fuzzy model consists of a total of \( n \) fuzzy rules. The integer parameter \( d \) is introduced to estimate the present and future output signal by using only the past output signal.
The input vector to the fuzzy model consists of \( y \) and \( \Delta u \) which are past values of output and control input move, respectively, and can be indicated as a vector consisting of a total of \( m \) elements (\( m = n_y + n_u \), number of input variables to the fuzzy model):

\[
x(k) = \left[ y(k-d-1) \cdots y(k-d-n_y) \ \Delta u(k-1) \cdots \Delta u(k-n_u) \right].
\]

The fuzzy model identification can be accomplished through clustering of numerical data. A SC method is used as the basis of a fast and robust algorithm for identifying a fuzzy model and assumes the availability of \( N \) input/output training data \( \mathbf{z}(k) = (\mathbf{x}(k), y(k)) \), \( k = 1, 2, \ldots, N \). It is assumed that the data points have been normalized in each dimension. The method starts by generating a number of clusters in the \( m \times N \) dimensional input space. The SC method considers each data point as a potential cluster center and uses a measure of the potential of each data point, which is defined as a function of the Euclidean distances to all other input data points [19-20]:

\[
\Psi_1(k) = \sum_{j=1}^{N} e^{-\left|\mathbf{x}(k) - \mathbf{x}(j)\right|^2/\alpha^2}, \quad k = 1, 2, \ldots, N,
\]

where \( \alpha \) is a radius, defining a neighborhood of a cluster center, which has considerable influence on the potential. Obviously, the potential of a data point is high when it is surrounded by many adjacent data. After the potential of every data point has been computed, the data point with the highest potential is selected as the first cluster center, which means that \( \Psi^*(1) = \max_k \{\Psi_1(k)\} \). After finding out the first cluster center \( \mathbf{x}^*(1) \) and its corresponding potential value \( \Psi^*(1) \), to solve the second cluster center, the potential of each data point is revised by the following formula:

\[
\Psi_2(k) = \Psi_1(k) - \Psi^*(1)e^{-\left|\mathbf{x}(k) - \mathbf{x}^*(1)\right|^2/\beta^2}, \quad k = 1, 2, \ldots, N,
\]

where \( \beta \) is another radius, usually greater than \( \alpha \) in order to limit the number of generated clusters. The second series of potentials is subtracted from the first series of potentials as a function of its distance from the first cluster center. The data points near the first cluster center will have greatly reduced potential, and therefore are unlikely to be selected as the next cluster center. When the potentials of all data points have been revised according to Eq. (4), the data point with the highest remaining potential is selected as the second cluster center that
corresponds to \( \Psi^*(2) = \max_k \{ \Psi_k' \} \). After the \( i \)-th cluster center has been obtained, the potential of each data point is revised by the following equation:

\[
\Psi_{i+1}(k) = \Psi_i(k) - \Psi^*(i)e^{-\frac{\|h_{i}(k)-x^*(i)\|^2}{\hat{\sigma}^2}}, \quad k = 1, 2, \ldots, N, \tag{5}
\]

where \( x^*(i) \) is the location of the \( i \)-th cluster center and \( \Psi^*(i) \) is its potential value. If the inequality \( \Psi^*(i) < \varepsilon \Psi^*(1) \) is true, these calculations stop, else these calculations are repeated. The parameter \( \varepsilon \) is a design parameter, which controls the number of generated clusters which is the number of fuzzy rules, \( n \).

When the SC method is applied to a collection of input/output data, each cluster center is in essence a prototypical data point that exemplifies a characteristic behavior of the system and each cluster center can be used as the basis of a fuzzy rule that describes the system behavior. Therefore, a fuzzy system identification model can be developed based on the results of the SC technique. The number of \( n \) fuzzy rules can be generated, where the premise parts are fuzzy sets, defined by the cluster centers that are obtained by the SC algorithm. The membership function value \( A_i(x(k)) \) of an input data vector \( x(k) \) to a cluster center \( x^*(i) \) can be defined as follows:

\[
A_i(x(k)) = e^{-\frac{\|x(k)-x^*(i)\|^2}{\hat{\sigma}^2}}, \tag{6}
\]

The fuzzy model output \( \hat{y}(k) \) is calculated by the weighted average of the consequent parts of the fuzzy rules as follows:

\[
\hat{y}(k) = \frac{\sum_{i=1}^{n} A_i(x(k))f_i(x(k))}{\sum_{i=1}^{n} A_i(x(k))}, \tag{7}
\]

where the function \( f_i(x(k)) \) which is an output of a fuzzy rule is a polynomial in the input variables and represented by the first-order polynomial of inputs as given in Eq. (8):

\[
f_i(x(k)) = \sum_{j=1}^{m} q_{i,j}x_j(k) + r_i, \tag{8}
\]

where

\[q_{i,j} = \text{weighting value of the } j\text{-th input on the } i\text{-th fuzzy rule output},\]
\( r_i = \text{bias of the } i\text{-th fuzzy rule output.} \)

Therefore, the estimated output by Eq. (7) can be rewritten as

\[
\hat{y}(k) = \sum_{i=1}^{n} \tilde{w}_i(k) f_i(x(k)) = \mathbf{w}^T(k) \mathbf{q},
\]

where

\[
\tilde{w}_i(k) = \frac{A_i(x(k))}{\sum_{i=1}^{n} A_i(x(k))},
\]

\[
\mathbf{q} = \left[ q_{1,1} \cdots q_{1,n} \cdots q_{m,1} \cdots q_{m,n} \, r_1 \cdots r_n \right]^T,
\]

\[
\mathbf{w}(k) = \left[ \tilde{w}_1(k)x_1(k) \cdots \tilde{w}_n(k)x_1(k) \cdots \tilde{w}_1(k)x_m(k) \cdots \tilde{w}_n(k)x_m(k) \cdots \tilde{w}_1(k) \cdots \tilde{w}_n(k) \right]^T, \quad k = 1, 2, \ldots, N.
\]

The value \( \tilde{w}_i(k) \) is the normalized firing level of the \( i\)-th fuzzy rule. For a series of the \( N \) input/output data pairs, the following equation is derived from Eq. (9):

\[
\hat{y} = \mathbf{Wq},
\]

where

\[
\hat{y} = \left[ \hat{y}(1) \, \hat{y}(2) \, \cdots \, \hat{y}(N) \right]^T,
\]

\[
\mathbf{W} = \left[ \mathbf{w}(1) \, \mathbf{w}(2) \, \cdots \, \mathbf{w}(N) \right]^T.
\]

The vector \( \mathbf{q} \) is called the consequent parameter vector. By replacing the vector \( \hat{y} \) and the matrix \( \mathbf{W} \) of Eq. (10) with the \( N \) training input/output data pairs, the vector \( \mathbf{q} \) is calculated. The current estimated output may be expressed generally as follows:

\[
\hat{y}(t/t) = f \left( \mathbf{x}(t) \right) = f \left( y(t-d-1), \, \cdots, \, y(t-d-n_y), \, \Delta u(t-1), \, \cdots, \, \Delta u(t-n_u) \right).
\]

The fuzzy model-based output prediction can be calculated as

\[
\hat{y}(t+k/t) = d(t/t) + f \left( \mathbf{x}(t+k) \right) = d(t/t) + f \left( y(t+k-d-1), \, \cdots, \, y(t+k-d-n_y), \, \Delta u(t+k-1), \, \cdots, \, \Delta u(t+k-n_u) \right),
\]

where \( d(t/t) \) is a disturbance estimation. The disturbance estimation is defined as the difference between the actual output values and the predicted output and is calculated as follows:

\[
d(t/t) = y(t) - \hat{y}(t/t).
\]
The disturbance estimation is assumed to be constant over the prediction horizon.

**2.2. Control Input Optimization by a Genetic Algorithm**

Since the genetic algorithm has been known to be effective in solving multiple objective functions, in this work, a genetic algorithm is used to minimize the control objective function. Compared to the conventional optimization methods that move from one point to another, genetic algorithms start from many points simultaneously climbing many peaks in parallel. Accordingly, genetic algorithms are less susceptible to getting stuck at local minima compared to conventional search methods [21-22].

The term *chromosome* of the genetic algorithm is referred to as a candidate solution that minimizes a control objective function. A chromosome which is a candidate solution of the optimization problem is represented by $s_g$, whose elements consist of present and future control inputs and has the following structure [20,23]:

$$s_g = [u_g(t) \quad u_g(t+1) \quad \cdots \quad u_g(t+M-1)], \quad g = 1, \ldots, G,$$

where $t$ indicates the current time. Assuming we have chosen the number of chromosomes $G$ which will constitute the initial population, the crossover probability $P_c$ and the mutation probability $P_m$, the algorithm proceeds according to the following steps:

Step 1 (initial population generation): Generate an initial population consisting of a total of $G$ chromosomes. The values are allocated randomly, but they should satisfy both input and input move constraints of Eq. (1). For this purpose, a simple procedure is used as follows:

(a) Read the measured value $u(t-1)$ of the input variable at the previous time point $t-1$, which has already been implemented. Then select the current input value using the following equations:

$$u_g(t) = u(t-1) + r \cdot \Delta u_{\text{max}}.$$

where $r$ is a random number with uniform distribution between

$$\left[\max\left\{\frac{u_{\text{min}} - u(t-1)}{\Delta u_{\text{max}}}, -1\right\}, \min\left\{\frac{u_{\text{max}} - u(t-1)}{\Delta u_{\text{max}}}, 1\right\}\right]$$

so that it satisfies the maximum and minimum input constraints.

(b) Select the rest of the input moves using the following equations:
\[ u_g(t+k) = u_g(t+k-1) + r \cdot \Delta u_{\text{max}}, \quad 1 \leq k \leq M-1. \] (13)

where \( r \) is a random number with uniform distribution between
\[
\left[ \max\left\{ u_{\text{min}} - u_g(t+k-1), -1 \right\}, \min\left\{ u_{\text{max}} - u_g(t+k-1), 1 \right\} \right].
\]

A new random number \( r \) is generated each time Eq. (12) or Eq. (13) is used.

Step 2 (fitness function evaluation): Evaluate the objective function of Eq. (1) for all the chosen chromosomes. Then invert the objective function values to obtain the fitness value of the \( g \)-th chromosome as follows:

\[ F_g = \frac{1}{J_g(t)} , \quad g = 1, \ldots, G \] (14)

where \( J_g(t) \) is the objective function value for the \( g \)-th chromosome. Then, calculate the normalized fitness value of each chromosome, meaning that the selection probability \( p_g \) is calculated by

\[ p_g = \frac{F_g}{\sum_{i=1}^{G} F_i} , \quad g = 1, \ldots, G . \] (15)

Step 3 (selection operation): Calculate the cumulative probability \( q_g \) for each chromosome using the following equation:

\[ q_g = \sum_{j=1}^{g} p_j , \quad g = 1, \ldots, G . \] (16)

Generate a random number \( r \) between 0 and 1 and select the chromosome for which \( q_{g-1} \leq r \leq q_g \).

At this point of the algorithm a new population of chromosomes has been generated. The chromosomes with high fitness value have more chance to be selected.

Step 4 (crossover operation): For each chromosome \( s_g \), generate a random number \( r \) between 0 and 1. If \( r \) is lower than \( p_c \), this particular chromosome will undergo the process of crossover, otherwise it will remain unchanged. Mate the selected chromosomes and for each selected pair \((s_g, s_{g+1})\), generate a random integer number \( c \) between 1 and \( M-1 \) which is a crossing point. Two new chromosomes \( (s^{\text{new}}_g, s^{\text{new}}_{g+1}) \) are produced by interchanging all the
members of the parent chromosomes following the crossing point. The crossover operation might produce infeasible offsprings if the input values at the cross point do not satisfy the input move constraints. Therefore, the crossover operation is conducted if and only if the following two conditions are satisfied for the values of the input parameters before and after the cross position so that the input move constraints are satisfied:

\[
\left| u_g^{\text{new}}(t+c) - u_g^{\text{new}}(t+c-1) \right| \leq \Delta u_{\text{max}} , \tag{17}
\]

\[
\left| u_{g+1}^{\text{new}}(t+c) - u_{g+1}^{\text{new}}(t+c-1) \right| \leq \Delta u_{\text{max}} , \tag{18}
\]

where \( u_g^{\text{new}} \) and \( u_{g+1}^{\text{new}} \) are members of newly generated chromosomes.

Step 5 (mutation operation): For every member of each chromosome \( s_g \), generate a random number \( r \) between 0 and 1. If \( r \) is lower than \( p_m \), this particular member of the chromosome will undergo the process of mutation, otherwise it will remain unchanged. For the selected members define upper and lower bounds as follows:

\[
b_u = \min(\Delta u_{\text{max}} + u(t-1), \Delta u_{\text{max}} + u_g(t+k+1), u_{\text{max}}) \quad \text{if } k = 0,
\]

\[
b_l = \max(-\Delta u_{\text{max}} + u(t-1), -\Delta u_{\text{max}} + u_g(t+k+1), u_{\text{min}})
\]

\[
b_u = \min(\Delta u_{\text{max}} + u_g(t+k-1), \Delta u_{\text{max}} + u_g(t+k+1), u_{\text{max}}) \quad \text{if } 0 < k < M-1,
\]

\[
b_l = \max(-\Delta u_{\text{max}} + u_g(t+k-1), -\Delta u_{\text{max}} + u_g(t+k+1), u_{\text{min}})
\]

\[
b_u = \min(\Delta u_{\text{max}} + u_g(t+k-1), u_{\text{max}}) \quad \text{if } k = M-1.
\]

\[
b_l = \max(-\Delta u_{\text{max}} + u_g(t+k-1), u_{\text{min}})
\]

The above bounds define the region of values of \( u_g(t+k) \) which will produce a feasible solution. This definition is followed by the generation of a random binary number \( b \). Based on the value of \( b \), \( u_g(t+k) \) is modified by the following equations:

\[
u_g(t+k) = u_g(t+k) + (b_u - u_g(t+k)) \left(1 - r^{(\text{iter}/\text{iter}_{\text{max}})}\right) \quad \text{if } b = 0,
\]

\[
u_g(t+k) = u_g(t+k) - (u_g(t+k) - b_l) \left(1 - r^{(\text{iter}/\text{iter}_{\text{max}})}\right) \quad \text{if } b = 1.
\]

where \( r \) is a random number between 0 and 1, \( \text{iter} \) is the number of iterations performed so far and \( \text{iter}_{\text{max}} \) is the expected final number of iterations.
Step 6 (repeat or stop): If the maximum allowed time has not expired, return the algorithm to Step 2. Otherwise, stop the algorithm and select the chromosome that produced the lowest value of the objective function throughout the entire procedure.

3. Fault Detection and Diagnostics for Input and Output Measurements

In this work, a failure detection and diagnostics algorithm is developed to estimate the input and output measurements using a fuzzy model based on the subtractive clustering method [19] and to check the operability of existing hardware sensors using a sequential probability ratio test so that the FTC can handle the fault situations of the input and output measurements. There are two types of approaches in estimating the input and output measurements. One is a method that estimates required parameters on the basis of a deterministic model, and the other is the black-box modeling method that depends only on the measured values. Black-box modeling approaches, such as artificial intelligence, are favored because they can model complicated processes that are difficult to be described by analytical and mechanistic methods. Therefore, black-box model approaches for estimating the signals have been widely attempted. Also recently, artificial intelligence techniques, such as fuzzy inference systems and artificial neural networks have been given close attention from many researchers because artificial intelligence can model complex nonlinear systems easily. Therefore, in this work, a fuzzy model is used to estimate the input and output measurement signals.

3.1 Signal Estimation

The estimation of sensor signals uses the fuzzy model based on the subtractive clustering method mentioned in the subsection “2.1 Output Prediction Using a Fuzzy Model”. Therefore, the detailed description for estimating the signals is abbreviated here. The design method that optimizes the fuzzy model is accomplished by a genetic algorithm, combined with a least-squares method. The genetic algorithm is used to optimize the cluster radii, \( r_\alpha \) and \( r_\beta \), for the subtractive clustering of numerical data, and the least squares algorithm is used to calculate the consequent parameters, \( q_0 \) and \( r_i \).

The problem of learning a smooth mapping from data samples is ill-posed in the sense that the reconstructed mapping is not unique. Constraints can be imposed to the mapping to make the problem well-posed. Typical constraints are smoothness and piecewise smoothness. This
method to exchange an ill-posed problem into a well-posed one is called regularization [24-25]. Regularization is a well-known method for the treatment of mathematically ill posed problems. It has been successfully applied to numerous machine-learning problems for avoiding overfitting in neural network training. In this work, the regularization is accomplished by making the fuzzy model have smaller consequent parameters, which causes the fuzzy model to respond smoother and less likely to overfit. The following multiple objectives are suggested to minimize a root mean squared error along with the small consequent parameters:

\[ F = \exp(-\mu_1 E_1 - \mu_2 E_2), \]  

where \( \mu_1 \) and \( \mu_2 \) are the weighting coefficients, and \( E_1 \) and \( E_2 \) are defined as

\[ E_1 = \left( \frac{1}{N} \sum_{k=1}^{N} (y(k) - \hat{y}(k))^2 \right)^{\frac{1}{2}}, \]

\[ E_2 = \frac{\mathbf{q}^T \mathbf{q}}{D}. \]

The variables \( y(k) \) and \( \hat{y}(k) \) denote the actual measured signal and the estimated signal, respectively. The parameter \( D \) is the number of consequent parameters, \( q_{ij} \) and \( r_i \), that is, the number of the elements of the vector \( \mathbf{q} \) which is \( n(m+1) \). \( E_2 \) is introduced for the regularization of the input. The smaller the consequent parameters, better is generalization capability of the fuzzy model. The fitness function depends strongly on the relative value of \( \mu_2/\mu_1 \). The ratio \( \mu_2/\mu_1 \) is iteratively altered in the training stage by the values of \( E_1 \) and \( E_2 \) so that the best chromosome with the maximum fitness keeps the specified relative magnitude of two terms of \( \mu_1 E_1 \) and \( \mu_2 E_2 \) in Equation (24).

Since the genetic algorithm requires much computational time if there are many parameters involved, the genetic algorithm is combined with a least-squares algorithm. If some parameters of the fuzzy model are fixed by the genetic algorithm, the resulting fuzzy model output can be described as a series of expansions of some basis functions. This basis function expansion is linear in its adjustable parameters as shown in Equation (9) since the parameters needed to calculate \( \mathbf{w}^T(k) \) such as \( r_\alpha \) and \( r_\beta \) and have been already known by the genetic algorithm. Therefore, the least-squares method can be used to determine the remaining
parameters of the fuzzy model to estimate the input and output measurement signals. The least-squares method was already explained in the subsection 2.1.

### 3.2 Failure Detection and Diagnostics

In failure detection and diagnostics, at every new sample of a signal, a new mean and a new variance of the signals may be required to check the health of the sensor. However, this procedure requires too many samples to obtain its meaningful mean and variance. During the acquisition of the samples, a significant degradation of the monitored process may occur. Therefore, sequential probability ratio test (SPRT) is used to detect a sensor failure based on the degree of failure and the continuous behavior of the sensor, without having to calculate a new mean and a new variance at each sampling instant. The SPRT is a statistical model developed by Wald in 1945 [26, 27].

The objective of sensor failure detection and diagnostics is to diagnose sensor health as soon as possible with a very small probability of making a wrong decision. The SPRT uses the residual (difference between the measured value and the estimated value, \( y(k) - \hat{y}(k) \)). Normally the residual signals are randomly distributed, so they are nearly uncorrelated and have a Gaussian distribution function \( P_i(\varepsilon_k, m_i, \sigma_i) \), where \( \varepsilon_k \) is the residual signal at time instant \( k \), and \( m_i \) and \( \sigma_i \) are the mean and the standard deviation under hypothesis \( i \), respectively. The sensor degradation or failure can be stated in terms of a change in the mean \( m \) or a change in the variance \( \sigma^2 \). Therefore, the SPRT detects sensor health by sensing the alteration of the probability distribution. If a set of samples, \( x_i \), \( i = 1, 2, \cdots, n \), is collected with a density function describing each sample in the set, an overall likelihood ratio is given by

\[
\gamma_n = \frac{P_1(\varepsilon_1 | H_1) \cdot P_1(\varepsilon_2 | H_1) \cdot P_1(\varepsilon_3 | H_1) \cdots P_1(\varepsilon_n | H_1)}{P_0(\varepsilon_1 | H_0) \cdot P_0(\varepsilon_2 | H_0) \cdot P_0(\varepsilon_3 | H_0) \cdots P_0(\varepsilon_n | H_0)},
\]

(27)

where \( H_0 \) represents a hypothesis that the sensor is normal and \( H_1 \) represents a hypothesis that the sensor is degraded.

By taking the logarithm of the above equation and replacing the probability density functions in terms of residuals, means and variances, the log likelihood ratio can be written as the following recurrent form:
\[ \lambda_n = \lambda_{n-1} + \ln \left( \frac{\sigma_0}{\sigma_1} \right) + \frac{(\epsilon_n - m_0)^2}{2\sigma_0^2} - \frac{(\epsilon_n - m_1)^2}{2\sigma_1^2}. \]  

(28)

This is the form used for deriving the sensor drift detection algorithm. For a normal sensor, the log likelihood ratio would decrease and eventually reach a specified bound \( A \), a smaller value than zero. When the ratio reaches this bound, the decision is made that the sensor is normal, and then the ratio is reinitialized by setting it equal to zero. For a degraded sensor, the ratio would increase and eventually reach a specified bound \( B \), a larger value than zero. When the ratio is equal to \( B \), the decision is made that the sensor is degraded. The decision boundaries \( A \) and \( B \) are chosen by a false alarm probability \( \alpha \) and a missed alarm probability \( \beta \); \[ A = \ln \left( \frac{\beta}{1-\alpha} \right) \] and \[ B = \ln \left( \frac{1-\beta}{\alpha} \right) \] [26].

4. Application to the SP-100 Space Reactor

The SP-100 space reactor system is a fast spectrum lithium-cooled reactor system that can generate electric power of 100 kW for space exploration and exploitation activities. The reactor system is made up of a reactor core, a primary heat transport loop, a thermoelectric generator, and a secondary heat transport loop to reject waste heat into space through radiators. The reactor core is composed of small disks of highly enriched (93%) uranium nitride fuel contained in sealed tubes. Figure 2 shows a schematic of one loop of the reactor system. The heat generated in the reactor core is transported by liquid lithium and is circulated by electromagnetic (EM) pumps. The interface between the primary heat transport system and the energy conversion system is a set of primary heat exchangers. The energy conversion system uses the direct thermoelectric (TE) conversion mechanism. A temperature drop of about 500 K is maintained across the TE elements by the cooling effect of a second liquid lithium loop that transfers the waste heat from the converter to a heat-pipe radiator.

Table 1 shows a comparison between the simulation results [28] and the design parameters of SNPSAM design [29] for the rated operation condition. Most of the simulation results are in good agreement with the SNPSAM design parameters. The noticeable differences
are in the TE hot side and cold side temperatures. The reason for the differences is that a secondary lithium loop is added to the developed simulation model while TE cold shoes are directly connected to radiators in SNPSAM design. However, the calculated electric conversion efficiency is close to the value of SNPSAM. Therefore, the developed simulation model can be used to provide a credible control design for SP-100 space reactor system. The schematic block diagram of the proposed FTC is illustrated in Figure 3.

The FTC for the TE power control is subject to constraints as follows:

\[
\Delta u(t + j - 1) = 0 \quad \text{for } j > M,
\]

\[
0^\circ \leq u(t) \leq 180^\circ,
\]

\[
|\Delta u(t)| \leq 1.4^\circ T.
\]

The sampling interval \( T \) is 1 second. The external reactivity control uses the mechanism of the stepper motor control drum system [2]. The control drum angle of the stepper motor shaft can be rotated from 0 to 180 degrees. The maximum angular velocity of the drums is \( 1.4^\circ / \text{sec} \). The input-weighting factor \( R \) is expressed by the following equation:

\[
R = \omega \left( \frac{y_o}{u_o} \right)^2
\]

where \( y_o \) is a rated value of the TE generator power and \( u_o \) is a rated value of control drum angle. It is known that the response speed depends much on the input-weighting factor that is a tuning variable. According as the input-weighting factor increases, the response speed becomes slower.

Figure 4 describes the performance of a failure detection and diagnostics algorithm. Figure 4(a) shows the training and test data used to design and test the algorithm which are different. Figure 4(b) and (c) shows the estimation performances of the algorithm for the training and test data, respectively. In addition, Figure 4(d) shows the fault detection and diagnostics performance when the output measurement is assumed to begin to be gradually degraded.
artificially from 300 sec. The gradual degradation of the output measurement is detected at 321 sec since the beginning of the gradual degradation.

Figure 5 shows the performance of the proposed FTC for normal transients such as the setpoint change of TE power. The setpoint starts to change by ramp at 100 sec and 300 sec, respectively and also changes by step at 600 sec and 800 sec. The performance of the proposed controller was checked with a roughly adjusted factor of \( \omega = 1 \). It is shown that the TE generator power follows its desired setpoint change very well. It was known that the proposed controller meets several constraints very well and accomplishes the fast and stable responses.

Figure 6 shows the performance of the proposed FTC against output measurement fault. The TE setpoint starts to change by ramp at 100 sec and 300 sec, respectively and also changes by step at 600 sec and 800 sec. The output measurement is assumed to start to be gradually degraded on purpose from 300 sec and the failure detection and diagnostics algorithm detects the output measurement degradation at 241 sec since the beginning of the gradual degradation. After detecting the fault, the FTC uses the estimated output signals instead of the measured output signal. Therefore, the FTC try to recover its actual desired output at 541 sec. It is shown that the TE generator power follows its desired setpoint change very well.

Figure 7 shows the performance of a proposed FTC against the change of the control input constraint. The maximum control input move is \( 1.4^\circ / \text{sec} \) but the maximum value changes to \( 0.14^\circ / \text{sec} \) from 500 sec. In order to observe that the response is slower after the maximum control input move becomes low, differently from the setpoint changes of the above simulation cases, the TE setpoint starts to change by ramp at 100 sec and changes by step at 350 sec, respectively and also changes by step at 550 sec and starts to change by ramp at 750 sec. Although its response becomes slower because of low control input change speed, it is shown that the TE power eventually follows its desired power.
5. Concluding Remarks

In this work, the fault-tolerant controller combining a model predictive controller and the fault detection and diagnostics algorithm was developed to control the nuclear power in the SP-100 space reactor system. Based on a fuzzy model consisting of the control drum angle change and the TE power, the future TE power is predicted by using the fuzzy model identified by a subtractive clustering method of a fast and robust algorithm. Another fuzzy model combined with the sequential probability ratio test estimates the input and output measurement signals and diagnoses the health of input and output measurements. The genetic algorithm was used to optimize the model predictive controller and both the fuzzy models. It was determined from many numerical simulation results that the proposed FTC was able to actuate the control drum to regulate the control reactivity so that the TE generator electric power followed the setpoint changes. Also, the performance of the new proposed controller was proved to be efficient even under constraint changes and gradual sensor degradation (fault).

References


### List of Tables

Table 1. Comparison between simulation results and design parameters.

### List of Figures

Figure 1. Model predictive control concept.

Figure 2. Schematic of the thermoelectric (TE) SP-100 space reactor system.

Figure 3. Block diagram of the proposed FTC for the SP-100 space reactors.

Figure 4. Performance of a failure detection and diagnostics algorithm.

(a) training data and test data  
(b) estimation performance for the training data  
(c) estimation performance for the test data  
(d) fault detection and diagnostics

Figure 5. Performance of the proposed FTC for normal transients.

(a) TE power and thermal reactor power  
(b) control drum angle  
(c) reactivity  
(d) temperature

Figure 6. Performance of the proposed FTC against output measurement fault.

(a) TE power and thermal reactor power  
(b) control drum angle  
(c) reactivity  
(d) temperature  
(e) fail flag

Figure 7. Performance of a proposed FTC against the change of the control input constraint.
(a) TE power and thermal reactor power  
(b) control drum angle  
(c) reactivity  
(d) temperature
Table 1. Comparison between simulation results and design parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>SNPSAM Design [28]</th>
<th>Current Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal thermal power (kW)</td>
<td>2000.0</td>
<td>2000.3</td>
</tr>
<tr>
<td>Electric power output (kW)</td>
<td>112.0</td>
<td>112.6</td>
</tr>
<tr>
<td>Thermal efficiency (%)</td>
<td>5.59</td>
<td>5.63</td>
</tr>
<tr>
<td>Core inlet temperature (K)</td>
<td>1254.0</td>
<td>1254.0</td>
</tr>
<tr>
<td>Core outlet temperature (K)</td>
<td>1284.0</td>
<td>1284.0</td>
</tr>
<tr>
<td>Fuel temperature (K)</td>
<td>1376.0</td>
<td>1376.0</td>
</tr>
<tr>
<td>Cladding temperature (K)</td>
<td>1288.0</td>
<td>1288.0</td>
</tr>
<tr>
<td>TE hot side temperature (K)</td>
<td>1237.0</td>
<td>1228.5</td>
</tr>
<tr>
<td>TE cold side temperature (K)</td>
<td>857.0</td>
<td>841.8</td>
</tr>
</tbody>
</table>
Fig. 1. Model predictive control concept.

Fig. 2. Schematic of the TE SP-100 space reactor system.
Fault-tolerant controller (FTC)

- Fault detection & diagnostics
- Optimization by GA
- Output prediction by a fuzzy model

Fig. 3. Block diagram of the proposed FTC for an SP-100 space reactor.
(a) training data and test data

(b) estimation performance for the training data

(c) estimation performance for the test data

(d) fault detection and diagnostics

Fig. 4. Performance of a failure detection and diagnostics algorithm.
Fig. 5. Performance of the proposed FTC for normal transients.
(a) TE power and thermal reactor power

(b) control drum angle

(c) reactivity

(d) temperature
Fig. 6. Performance of the proposed FTC against output measurement fault.

(a) TE power and thermal reactor power
(b) control drum angle
(c) reactivity
(d) temperature

Fig. 7. Performance of a proposed FTC against the change of the control input constraint.
APPENDIX B

Orifice Meter Calibration Procedure and Results

In order to calibrate the orifice meters, rigorous mechanical timing experiments were performed to find the proportionality of flow to pressure drop across the orifice using the relationship, Flow rate = K*SQRT (Pressure difference). Using a timer and pre-calibrated liter marks on the level tanks, average flow rates were determined for different positions of the control valves. Simultaneously, output voltages from the pressure sensors were compared to values set with the calibration software, AMS Device Manager. Using linear interpolation of voltages between the maximum and minimum pressure differences, a $\Delta P$ was found for each output voltage recorded at each control valve position. Four sets of test positions were performed for each tank; from these an average proportionality constant was determined. Any anomalous data points were retaken to allow a more accurate calculation to be made.

Procedure:

1) Find Flow Rate

1. Isolate one tower.
2. Close all Valves.
3. Turn on Pump.
4. Open Bypass
5. Input test voltage into Control Valve in order to open it.
6. As water fills to zero line start timer.
7. Stop timer when a whole number a liters are filled.
8. Record Time to fill in seconds (T) and Liters filled (L)
9. From these Calculate Flow rate for each test point

2) Find Pressure Difference ($\Delta P$)

1. Find Max and Min Voltage and corresponding $\Delta P$ using calibration software
2. Record output voltage from pressure sensor at each test point
3. Perform linear interpolation between max and min to determine $\Delta P$ at each point.

3) Determine Proportionality constant $K$

1. Using flow rate and $\Delta P$ relationship find $K$ at each point
2. Find Average $K$

Calculations:

Flow Rate:
$F_{Rate}$ (L/s) = L/T

$\Delta P$:
$\Delta P$ (mm H$_2$O) = SLOPE($\Delta P_{min}$, $\Delta P_{max}$, $V_{min}$, $V_{max}$)*$\Delta P_{test}$ + INTERCEPT($\Delta P_{min}$, $\Delta P_{max}$, $V_{min}$, $V_{max}$)
$K = (F_{Rate} * 0.001) / \text{SQRT}(\Delta P * 9.806652)$
## Results

### Orifice Meter 1

<table>
<thead>
<tr>
<th>Voltage to Control Valve 1</th>
<th>Liters Filled</th>
<th>Time To Fill (sec)</th>
<th>Fill Flow Rate (Liter/sec)</th>
<th>Flow Rate (GPM)</th>
<th>Voltage Recorded</th>
<th>ΔP in mm H2O</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>168.33</td>
<td>0.0475257</td>
<td>0.7532917</td>
<td>0.843152</td>
<td>84.36427</td>
<td>1.6523E-06</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>64.54</td>
<td>0.1239493</td>
<td>1.9646218</td>
<td>1.112355</td>
<td>585.1418</td>
<td>1.63626E-06</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>41.01</td>
<td>0.1950982</td>
<td>3.0923448</td>
<td>1.582293</td>
<td>1459.329</td>
<td>1.63086E-06</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>29.65</td>
<td>0.2697918</td>
<td>4.2762533</td>
<td>2.252593</td>
<td>2706.235</td>
<td>1.6561E-06</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>24.38</td>
<td>0.3281715</td>
<td>5.2015834</td>
<td>3.026973</td>
<td>4146.753</td>
<td>1.62737E-06</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>18.72</td>
<td>0.3740315</td>
<td>5.9284745</td>
<td>3.712115</td>
<td>5421.27</td>
<td>1.62217E-06</td>
</tr>
<tr>
<td>9</td>
<td>7</td>
<td>17.36</td>
<td>0.4032839</td>
<td>6.3921302</td>
<td>4.149825</td>
<td>6235.507</td>
<td>1.63085E-06</td>
</tr>
</tbody>
</table>

Avg K: 1.63656E-06
### K Orifice 1

**Graph:**
- **K Orifice 1**
- Voltage to Control Valve on the x-axis, ranging from 2.5 to 9.5.
- K on the y-axis, ranging from 0.0000158 to 0.000018.
- Line graphs for Series 1 to Series 4, all with a trend of decreasing K as voltage increases.
- A blue cross indicates the average value.

### Orifice Meter 2

<table>
<thead>
<tr>
<th>Voltage to Control Valve</th>
<th>Liters Filled</th>
<th>Time To Fill (sec)</th>
<th>Fill Flow Rate (Liter/sec)</th>
<th>Flow Rate (GPM)</th>
<th>Voltage Recorded</th>
<th>ΔP in mm H2O</th>
<th>K</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>1</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>86.43</td>
<td>0.0925578</td>
<td>1.4670593</td>
<td>0.971</td>
<td>318.8873</td>
<td>1.65514E-06</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>46.30</td>
<td>0.1727768</td>
<td>2.7385476</td>
<td>1.403</td>
<td>1128.569</td>
<td>1.64233E-06</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>32.54</td>
<td>0.2458513</td>
<td>3.8967916</td>
<td>2.007</td>
<td>2260.625</td>
<td>1.65119E-06</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>25.52</td>
<td>0.3134489</td>
<td>4.9682280</td>
<td>2.735</td>
<td>3629.775</td>
<td>1.66137E-06</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>22.09</td>
<td>0.3621548</td>
<td>5.7402263</td>
<td>3.48</td>
<td>5021.416</td>
<td>1.632E-06</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>20.40</td>
<td>0.3921569</td>
<td>6.2157647</td>
<td>4.0175</td>
<td>6028.833</td>
<td>1.61281E-06</td>
</tr>
<tr>
<td>8.5</td>
<td>8</td>
<td>19.56</td>
<td>0.4089980</td>
<td>6.4826994</td>
<td>4.16382</td>
<td>6303.075</td>
<td>1.64507E-06</td>
</tr>
</tbody>
</table>

**Average K:** 1.64284E-06
Concluding Remarks

After analysis of the data, for orifice meter 1, the proportionality constant K is found to be 1.63666E-6, making the calibration curve for Orifice Meter 1, including all conversions:

\[
\sqrt{(1860.2209 \times \text{Voltage} - 1484.084253) \times 9.806652} \times 1.63666E-6 \times 15850.2 = \text{Flow Rate (GPM)}.
\]

For Orifice Meter 2 K is found to be 1.64296E-6. The corresponding calibration curve is:

\[
\sqrt{(1874.264032 \times x - 1501.023093) \times 9.806652} \times 1.64296E-6 \times 15850.2 = \text{Flow Rate (GPM)}.
\]
% the initial values of each state variable
clear all
close all
clc
lambda=[0.0127, 0.0317, 0.115, 0.311, 1.4, 3.78];
beta=[0.000252,0.00147,0.001344,0.002941,0.001024,0.000237];
beta_tot=sum(beta);

%% average generation time (s)
kapa=1.55e-7;

%% Coolant flow rate
Wc=14.1;
Ws=14.1;

%% fuel doppler coefficient
alpha = 0.24e-6;

%% core expansion
alpha_fuel = -3.72e-6 - 8.50e-6;
alpha_clad = 0.12e-6;

%% coolant expansion
alpha_co = -3.25e-6;

%% reactor core thermal model
mCf = 1.465e2 * 109.0;
hAf = 2.0e6/(1376-1288);
mClad = 0.276 * 25.5;
hAclad = 2.0e6/(1288-1267.1);
CpLi = 4200.0;

%% initial state values
theta=fzero(inline('9.41E-11*x^5-3.57E-8*x^4+2.21E-6*x^3+3.72E-4*x^2-2.5E-3*x-4.0'),125);
\[ d\theta = 0; \]
\[ \text{Power} = 2.0000 \times 10^6; \]
\[ \text{Cfission} = [2.56044980203996, 5.98380618825848, 1.50806782543339, 1.22026479381736, 0.09438247614957, 0.00809051173591] \times 10^{11}; \]
\[ T_{\text{fuel}0} = 1376.0; \]
\[ T_{\text{clad}0} = 1288.0; \]
\[ T_{\text{core}0} = 1267.1; \]
\[ T_{\text{thxin0}} = 1284.; \]
\[ T_{\text{thxout0}} = 1250.2; \]
\[ T_{\text{thxsout0}} = 848.0; \]
\[ T_{\text{thxsin0}} = 828.0; \]
\[ N_{\text{hx}} = 12; \]
\[ N_{\text{channel}} = 30; \]
\[ \epsilon_{\text{n}} = 0.85; \]

\[ \text{Mdot\_tot} = \frac{W_c}{N_{\text{hx}}}; \]
\[ \text{Sdot\_tot} = \frac{W_s}{N_{\text{hx}}}; \]

%%% parameter definition for thermal calculation

%%% channel inner and outer size (m);
\[ w_{\text{channel}} = 0.0254; \]
\[ t_{\text{channel}} = 0.0035; \]
\[ l_{\text{channel}} = 0.254; \]
\[ t_{\text{wall}} = 0.0005; \]

%%% # of segment per channel
\[ N_{\text{segp}} = 1; \]

%%% # of TE couples per channel
\[ N_{\text{te}} = 480; \]
%%header size (m)

Lhead = 0.8;
Dhead_in = 0.0413;
Dhead_end = 0.019;
lhead = 0.2825;

mdot=Mdot_tot/Nchannel;
sdot = Sdot_tot/Nchannel;

vchannel =wchannel * tchannel *lchannel;
cp = 4200.0;
wp = mdot;
ws=sdot;
Tav_ref = 1267.1 ;
Tav_ref2 = 865.0 ;

%%% the parameters to compute the thermal electric resistance
%%% Nb-Zr

t1 = 0.05e-2;
k1 = 41.9;

%%% Al203

t2=0.0254e-2;
k2=25.0;

%%% Tungstern

t3 = 0.127e-2;
k3 = 163.3;

%%% Forsterite

t4 = 0.0025e-2;
k4 = 8.0;
% Nickel

t5 = 0.03e-2;
k5 = 60.7;
Ap = 0.14e-4;
An = 0.14e-4;
Apn = Ap+An;

Tref3 = 0.5 * (828.0+841.2);
rho = interp1([0 600 1600 1.e10],[500 500 400 400],Tref3,'linear','extrap');
Cprad = cp*rho*vchannel;
sbc = 5.67e-8;
Arad = 17.8051*0.14e-4*2.0*480*1.175;
Tatm = 250.0;

%%% Cross sections: An:Ap = 0.14cm^2, 0.14 cm^2
Ap = 0.14;
An = 0.14;

%%% Length: 0.38 cm
Length = 0.38;

%%% Compute the initial state

Y0 = [Thxout0, Thxsout0, Tfuel0, Tclad0, Tcore0, Thxin0, Power, Cfission, Thxin0];
Tp = Thxout0;
Ts = Thxsout0;
Tpin = Thxin0;
Tsin = Thxin0;
Rl = 0.016;
APPENDIX D

Fault Detection and Isolation (FDI) Algorithm for SP-100 System

close all
clear all

load data_pca_0802
load dat_test_0801
n=1200;

Tpout=state(1:n,7);
Tsout=state(1:n,2);
Tpin=state(1:n,1);
Tsin=state(1:n,8);
Thot=state(1:n,10);
Tcol=state(1:n,11);
Pt=state(1:n,3);
Pe=state(1:n,9);
a=0.002;
X1=Tpout+a*randn(n,1)*mean(Tpout);
X2=Tsout+a*randn(n,1)*mean(Tsout);
X3=Tpin+a*randn(n,1)*mean(Tpin);
X4=Tsin+a*randn(n,1)*mean(Tsin);
X5=Thot+a*randn(n,1)*mean(Thot);
X6=Tcol+a*randn(n,1)*mean(Tcol);
X7=(Pt+a*randn(n,1)*mean(Pt))/1000;
X8=Pe+a*randn(n,1)*mean(Pe);

dat=[X1,X2,X3,X4,X5,X6,X7,X8];
Tpout_test=dat_test(:,1);
Tsout_test=dat_test(:,2);
Tpin_test=dat_test(:,3);
Tsin_test=dat_test(:,4);
Thot_test=dat_test(:,5);
Tcol_test=dat_test(:,6);
Pt_test=dat_test(:,7);
Pe_test=dat_test(:,8);

m=length(Tpout_test);
X1_test=Tpout_test+a*randn(m,1)*mean(Tpout);
X2_test=Tsout_test+a*randn(m,1)*mean(Tsout) ;
X3_test=Tpin_test+a*randn(m,1)*mean(Tpin);
X4_test=Tsin_test+a*randn(m,1)*mean(Tsin) ;
X5_test=Thot_test+a*randn(m,1)*mean(Thot) ;
X6_test=Tcol_test+a*randn(m,1)*mean(Tcol) ;
X7_test=(Pt_test+a*randn(m,1)*mean(Pt))/1000;
X8_test=Pe_test+a*randn(m,1)*mean(Pe);

test=[X1_test,X2_test,X3_test,X4_test,X5_test,X6_test,X7_test,X8_test];

[datnew,mv,std]=zscore(dat);
[pc,score,latent,tsquare] = princomp(datnew);
PP=pc(:,1);

test_new=zscore(test,mv,std);
for i=1:m
    test_pc(i,:)=mv+std.*(test_new(i,:)*PP*PP');
    error(i,:)=datnew(i,:)*(eye(8,8)-PP*PP');
end
% Q statistics
[q]=qstat(datnew,pc,1);
[qa]=qlim(datnew,1);

% b==drift magnitude
b=0.005 ;

% fault position: coolant Temperature outlet Primary loop
fault1=X1_test(end)+b*mv(1);
dat_f1=[fault1,test(end,2:8)];
[dat_f1_new,mean1,std1]=zscore(dat_f1,mv,std);
[q1]=qstat(dat_f1_new,pc,1);

err1=dat_f1_new*(eye(8,8)-PP*PP');
[FF1,E1]=eig(err1'*err1);
F1=FF1(:,8);

% fault position: coolant Temperature inlet Primary loop
fault2=X2_test(end)+b*mv(2);
dat_f2=[test(end,1),fault2,test(end,3:8)];
dat_f2_new=zscore(dat_f2,mv,std);

err2=dat_f2_new*(eye(8,8)-PP*PP');
[FF2,E2]=eig(err2'*err2);
F2=FF2(:,8);

% fault position: coolant temperature outlet secondary loop
fault3 = X3_test(end) + b*mv(3);
dat_f3 = [test(end,1:2), fault3, test(end,4:8)];
dat_f3_new = zscore(dat_f3,mv, std);

err3 = dat_f3_new*(eye(8,8)-PP*PP');
[FF3,E3] = eig(err3'*err3);
F3 = FF3(:,8);

% fault position: coolant temperature inlet secondary loop
fault4 = X4_test(end) + b*mv(4);
dat_f4 = [test(end,1:3), fault4, test(end,5:8)];
dat_f4_new = zscore(dat_f4,mv, std);

err4 = dat_f4_new*(eye(8,8)-PP*PP');
[FF4,E4] = eig(err4'*err4);
F4 = FF4(:,8);

% fault position: hot leg temperature
fault5 = X5_test(end) + b*mv(5);
dat_f5 = [test(end,1:4), fault5, test(end,6:8)];
dat_f5_new = zscore(dat_f5,mv, std);

err5 = dat_f5_new*(eye(8,8)-PP*PP');
[FF5,E5] = eig(err5'*err5);
F5 = FF5(:,8);

% fault position: cold leg temperature
fault6 = X6_test(end) + b*mv(6);
dat_f6 = [test(end,1:5), fault6, test(end,7:8)];
dat_f6_new = zscore(dat_f6,mv, std);
err6 = dat_f6_new * (eye(8, 8) - PP * PP');
[FF6, E6] = eig(err6' * err6);
F6 = FF6(:, 8);

% fault position: reactor thermal power
fault7 = X7_test(end) + b * mv(7);
dat_f7 = [test(end, 1:6), fault7, test(end, 8)];
dat_f7_new = zscore(dat_f7, mv, std);

err7 = dat_f7_new * (eye(8, 8) - PP * PP');
[FF7, E7] = eig(err7' * err7);
F7 = FF7(:, 8);

% fault position: electric power
fault8 = X8_test(end) + b * mv(8);
dat_f8 = [test(end, 1:7), fault8];
dat_f8_new = zscore(dat_f8, mv, std);

err8 = dat_f8_new * (eye(8, 8) - PP * PP');
[FF8, E8] = eig(err8' * err8);
F8 = FF8(:, 8);
F = [F1, F2, F3, F4, F5, F6, F7, F8];

for i = 1:8

    Q1(i) = err1 * (eye(8, 8) - F(:, i) * F(:, i))' * 2 * err1';
    QQ1 = err1 * err1';
    Q2(i) = err2 * (eye(8, 8) - F(:, i) * F(:, i))' * 2 * err2';
    QQ2 = err2 * err2';
    Q3(i) = err3 * (eye(8, 8) - F(:, i) * F(:, i))' * 2 * err3';

end
QQ3 = err3 * err3';
Q4(i) = err4 * (eye(8,8) - F(:,i) * F(:,i))' * err4';
QQ4 = err4 * err4';
Q5(i) = err5 * (eye(8,8) - F(:,i) * F(:,i))' * err5';
QQ5 = err5 * err5';
Q6(i) = err6 * (eye(8,8) - F(:,i) * F(:,i))' * err6';
QQ6 = err6 * err6';
Q7(i) = err7 * (eye(8,8) - F(:,i) * F(:,i))' * err7';
QQ7 = err7 * err7';
Q8(i) = err8 * (eye(8,8) - F(:,i) * F(:,i))' * err8';
QQ8 = err8 * err8';

end
FI1 = ones(1,8) - Q1 / QQ1;
FI2 = ones(1,8) - Q2 / QQ2;
FI3 = ones(1,8) - Q3 / QQ3;
FI4 = ones(1,8) - Q4 / QQ4;
FI5 = ones(1,8) - Q5 / QQ5;
FI6 = ones(1,8) - Q6 / QQ6;
FI7 = ones(1,8) - Q7 / QQ7;
FI8 = ones(1,8) - Q8 / QQ8;
FI = [FI1', FI2', FI3', FI4', FI5', FI6', FI7', FI8'];
bar(FI)
title('Fault Isolation Index');
xlabel('Fault Direction #');
ylabel('FI')
legend('1','2','3','4','5','6','7','8')