Universita' degli studi di Bari

Dottorato di Ricerca in Fisica

Tesi di Dottorato

Design, construction and commissioning of the Thermal Screen Control System for the CMS Tracker detector at CERN

Dottorando:
Enzo Carrone

Gennaio 2005

Table of contents

Introduction. ix

CHAPTER 1. CERN and CMS 1

1.1. LHC 1
1.2. Physics Motivation 3
1.3. The Compact Muon Solenoid 4
1.3.1. The CMS detectors 6
1.3.2. The Tracker Architecture 8
1.3.3. Silicon Sensors 11
1.3.4. Radiation Damage 15
1.3.5. The Tracker Control System 16

CHAPTER 2. Control Systems 17

2.1. SCADA 20
2.2. Hierarchical Decomposition of Control Systems 20
2.3. The Regulatory Layer and its Model 22
2.4. Decentralized Control of the Regulatory Layer 24
2.5. Control Systems Design 24
2.6. Definition of the Problem 26
2.7. Physics and Controls 27
2.8. Model Based Design 27
2.9. System Modeling 29
2.9.1. Mechanistic Models 29
2.9.2. Black Box Models 30
2.10. Model Representation 30
2.11. Feedback Systems 32
2.12. A generic Control System with Error Feedback 33
2.13. System Identification 34
2.14. Control Performances 37
2.15. Stability of Control Systems 39
2.16. Controller Design 40
2.17. PID Controllers 41
2.17.1. Proportional Controller 43
2.17.2. Proportional Control of First-order Systems 44
2.17.3. Proportional Control of Second Order Systems 45
2.17.4. Integral Controller 48
2.17.5. Proportional plus Integral (PI) Controller 49
2.17.6. Derivative Controller 51
2.17.7. Proportional plus Derivative (PD) Control 52

2.17.8. Proportional plus Integral plus Derivative (PID)Control 53
2.17.9. PID Controllers and Measures of Control Performances 54
2.18. PID Tuning 56
2.18.1. Internal Model Control (IMC) 59
2.19. Control of Multiple-Input, Multiple-Output Processes 60
2.19.1. Multiloop Control Strategy 62
2.20. The Relative Gain Array (RGA) 63

CHAPTER 3. Thermal screen constraints and specifications 67

3.1. General Description 69
3.1.1. Modularity 71
3.1.2. Accessibility and Maintenance 71
3.1.2.1. The Heating Foils 72
3.1.3. The Insulating Layer 72
3.1.4. The Cold Plates 74
3.1.5. The Assembled Panel 74
3.2. Cooling 75
3.3. Heating 76
3.4. Thermal Model 77
3.5. Thermal Resistances and Heat Fluxes Evaluation 78
3.6. Heat Fluxes Evaluation 79

CHAPTER 4. The thermal screen control system 81

4.1. A Systems Engineering Approach 81
4.2. Risk Management 85
4.3. The Thermal Screen: a Safety System 86
4.4. A common Design Methodology 89
4.5. Performance Specifications 90
4.6. Overview 91
4.7. System Modeling 92
4.8. Controller Design 96
4.9. The MIMO Problem 98
4.9.1. PLC Programming 100
4.10. Operation and Programming 102
4.10.1. Programming Standards: IEC 1131 103
4.11. Industrial Communication 103
4.11.1. OPC 103
4.12. The Thermal Screen PLC 105
4.13. Power Supply 106
4.14. PID Programming 108
4.14.1. Human Machine Interface 112
4.15. Commissioning 116
4.16. Interactions with CMS Detector Control System 119
4.17. Risk Assessment 121

4.17.1. Initial Risk Assessment 122
4.17.2. HAZOP 123

CHAPTER 5. Commissioning, System Identification and Tests 125

5.1. Test Bench and Climatic Chamber 126
5.2. Identification Process 129
5.3. System Identification of a Panel 132
5.4. Controller Tuning 138
5.5. Interlock Tests 144

Conclusions. 149

Appendix. 151

A.1. Polynomial Representation of Transfer Functions 151
A.2. CMS Cooling Plant 154
A.3. System Identification with MatLab Toolbox 157
A.4. Initial Risk Evaluation 159
A.5. HAZOP Analysis 160
A.6. Event Tree 161
A.7. Test plan 163
A.8. Acceptance Test 166

Bibliography. 169

List of figures

Figure 1-1 CERN Accelerator Complex 2
Figure 1-2 LHC Experiments 3
Figure 1-4 CMS Detectors 5
Figure 1-3 CMS 5
Figure 1-5 Particle tracks 6
Figure 1-6 CMS Tracker 9
Figure 1-7 One quarter of the CMS Tracker layout. Pixel detector layers are shown in purple, while strip detectors are in red (single-sided) and blue (double-sided). 9
Figure 1-8 Tracker's Functional Groups 10
Figure 1-9 A charged particle traversing the detector generates free electron-hole pairs along its track, which are moved by electric field. 11
Figure 1-10 Schematic of a Silicon Detector 12
Figure 1-11 Charge density, electric field and potential in a one-dimensional model of a silicon detector at full depletion. 13
Figure 1-12 The expected radiation fluences of photons, neutrons and charged hadrons in the CMS experiment over 10 years of operation as a function of the distance z from the collision point along the beam axis and the radius r. 15
Figure 1-13 Tracker Control System Architecture 16
Figure 2-1 Hierarchical decomposition of a control architecture 21
Figure 2-2 SCADA implementation at CERN 22
Figure 2-3 A General Feedback System 23
Figure 2-4 Process and its associated inputs/outputs 30
Figure 2-5 Generic control system with error feedback 34
Figure 2-6 Control Systems Design Procedure 35
Figure 2-7 System Identification 36
Figure 2-8 Response to a step input 37
Figure 2-9 Gain and Phase Margin 38
Figure 2-10 Control System with its Controller 41
Figure 2-11 A PID Controller Implementation 43
Figure 2-12 Proportional Controller 44
Figure 2-13 Effect of Proportional Gain Variation 47
Figure 2-14 Proportional effect 48
Figure 2-15 Proportional Controller Implementation 48
Figure 2-17 Integral effect 50
Figure 2-16 Integral Controller Implementation 50
Figure 2-19 Derivative effect 51
Figure 2-18 Effect of variations of the integral time 51

Figure 2-20 Derivative Controller Implementation 52
Figure 2-21 Effect of variations of derivation time 53
Figure 2-22 PID “Parallel” Structure 54
Figure 2-24 PID parameters - 2 55
Figure 2-25 PID parameters - 3 55
Figure 2-23 PID parameters - 1 55
Figure 2-26 Errors minimization 56
Figure 2-27 Schematic of Internal Model Control Strategy 60
Figure 2-28 Process Interactions 61
Figure 2-29 2×2 system 62
Figure 2-30 Block diagram for conventional 2×2 multiloop schemes 64
Figure 3-1 Tracker - External rails for the thermsl screen 67
Figure 3-2 Thermal Screen’s Three Layers 69
Figure 3-1 Tracker Support Tube (Front View) 70
Figure 3-2 Tracker Support Tube 70
Figure 3-3 A Thermal Screen Panel Installed 71
Figure 3-4 Heating foil - RTD detail 73
Figure 3-5 A thermal screen panel laying on the floor 73
Figure 3-6 Cold plate pipes schematic 76
Figure 3-7 Section view of one panel 77
Figure 3-8 Thermal screen cooling plant 78
Figure 3-9 Thermal model 79
Figure 3-10 Thermal model II 80
Figure 4-1 A systems Engineering process 84
Figure 4-2 Safety Instrument Functions 89
Figure 4-3 Tracker DCS interactions 92
Figure 4-4 Thermal screen process representation 94
Figure 4-5 ISA drawing of a single panel circuit 94
Figure 4-6 Thermal network describing heat conduction from junction temperature through the case to ambient temperature. Current source equals the power dissipation. 95
Figure 4-7 Electro-thermal model 96
Figure 4-9 Bode diagram 97
Figure 4-8 Simulink diagram 97
Figure 4-10 Step Response 98
Figure 4-11 MatLab Optimization toolbox output 99
Figure 4-12 Panels mutual interactions 100
Figure 4-13 MIMO electro-thermal model 100
Figure 4-14 Block diagram of a typical PLC 103
Figure 4-16 OPC communications 106
Figure 4-15 Industrial communications layers 106

Figure 4-17 Thermal screen control system schematics 107
Figure 4-18 Xantrex Power Supply 108
Figure 4-19 Sequence of functions of the standard PID controller. 109
Figure 4-20 PID parameters selection points 111
Figure 4-21 Simatic PID HMI 113
Figure 4-22 Thermal Screen HMI 114
Figure 4-24 PID Dynamic Trends 115
Figure 4-23 Detailed PID HMI 115
Figure 4-25 First PLC Rail 116
Figure 4-26 Second PLC Rail 116
Figure 4-27 Third PLC Rack 117
Figure 4-28 Patch Panel 118
Figure 4-29 LHC Rack 118
Figure 4-30 Siemens PLC 119
Figure 4-31 Rack Backplane 119
Figure 5-1 Climatic chamber artistic view 126
Figure 5-2 Climatic chamber - Side view 127
Figure 5-3 Climatic chamber - Front view 127
Figure 5-5 Climatic Chamber - Internal View 128
Figure 5-4 Climatic Chamber - External View 128
Figure 5-6 Matlab Identification Toolbox GUI 129
Figure 5-7 Models comparison 130
Figure 5-8 Time and frequency response of the BJ model 132
Figure 5-9 IDENT screenshot for the PRBN test 134
Figure 5-10 Temperature and voltage vs. time diagram 134
Figure 5-11 Spectrum Distribution 135
Figure 5-13 Step response of the P1 model 136
Figure 5-12 Bode diagram of the P1 Model 136
Figure 5-14 Skogestad tuning 138
Figure 5-15 IMC Tuning 139
Figure 5-16 IAE Minimization Tuning 140
Figure 5-17 MatLab Optimization Toolbox tuning 141
Figure 5-18 Heatig Foils behaviour 142
Figure 5-19 CST Cooling Down 143
Figure 5-20 Beam Test Experimental Setup 145
Figure 5-21 CAN power supply with the interlock line 146
Figure 5-22 Beam Test HMI Main 147
Figure 5-23 TIB HMI 147
Figure 5-24 TEC HMI 148
Figure A-1. Thermal Screen Pumping Station 154

Figure A-2. CMS Cooling Plant 155
Figure A-3. CMS Tracker Cooling Requirements 156
Figure A-4. Ident GUI 157
Figure A-5. Test Scenario Specification Form 165

Il maestro e' nell'anima

e dentro l'anima per sempre restera'.

(Paolo Conte).

Introduction

The CERN (European Organization for Nuclear Research) laboratory is currently building the Large Hadron Collider (LHC). Four international collaborations have designed (and are now constructing) detectors able to exploit the physics potential of this collider. Among them is the Compact Muon Solenoid (CMS), a general purpose detector optimized for the search of Higgs boson and for physics beyond the Standard Model of fundamental interactions between elementary particles.

This thesis presents, in particular, the design, construction, commissioning and test of the control system for a screen that provides a thermal separation between the Tracker and ECAL (Electromagnetic CALorimeter) detector of CMS (Compact Muon Solenoid experiment).

Chapter 1 introduces the new challenges posed by these installations and deals, more in detail, with the Tracker detector of CMS.

The size of current experiments for high energy physics is comparable to that of a small industrial plant: therefore, the techniques used for controls and regulations, although highly customized, must adopt Commercial Off The Shelf (COTS) hardare and software.

The "slow control" systems for the experiments at CERN make extensive use of PLCs (Programmable Logic Controllers) and SCADA (Supervisory Control and Data Acquisition) to provide safety levels (namely interlocks), regulations, remote control of high and low voltages distributions, as well as archiving and trending facilities. The system described in this thesis must follow the same philosophy and, at the same time, comply with international engineering standards.

While the interlocks applications belong straightforwardly to the category of DES (Discrete Event System), and are therefore treated with a Finite State Machine approach, other controls are more strictly related to the regulation problem. Chapter 2 will focus on various aspects of modern process control and on the tools used to design the control system for the thermal screen: the principles upon which the controller is designed and tuned, and the model validated, including the Multiple Input-Multiple Output (MIMO) problematics are explained.

The thermal screen itself, the constraints and the basis of its functioning are described in Chapter 3, where the thermodynamical design is discussed as well.

For the LHC experiments, the aim of a control system is also to provide a well defined SIL (Safety Interlock Level) to keep the system in a safe condition; yet, in this case, it is necessary to regulate the temperature of the system within certain values and respect the constraints arising from the specific needs of the above mentioned subsystems.

The most natural choice for a PLC-based controller is a PID (Proportional Integral Derivative) controller. This kind of controller is widely used in many industrial process, from batch production in the pharmaceutics or automotive field to chemical plants, distillation columns and, in general, wherever a reliable and robust control is needed.

In order to design and tune PID controllers, many techniques are in use; the approach followed in this thesis is that of black-box modeling: the system is modeled in the time domain, a transfer function is inferred and a controller is designed. Then, a system identification procedure allows for a more thorough study and validation of the model, and for the controller tuning.

Project of the thermal screen control including system modeling, controller design and MIMO implementation issues are entirely covered in Chapter 4. A systems engineering methodology has been followed all along to adequately manage and document every phase of the project, complying with time and budget constraints. A risk analysis has been performed, using Layer of Protection Analysis (LOPA) and Hazard and Operability Studies (HAZOP), to understand the level of protection assured by the thermal screen and its control components.

Tests planned and then performed to validate the model and for quality assurance purposes are described in Chapter 5. A climatic chamber has been designed and built at CERN, where the real operating conditions of the thermal screen are simulated. Detailed test procedures have been defined, following IEEE standards, in order to completely check every single thermal screen panel. This installation allows for a comparison of different controller tuning approaches, including IAE minimization, Skogestad tuning rules, Internal Model Control (IMC), and a technique based upon the MatLab Optimization toolbox.

This installation is also used for system identification purposes and for the acceptance tests of every thermal screen panel (allowing for both electrical and hydraulic checks).

Also, tests have been performed on the West Hall CERN experimental area , where a full control system has been set up, dedicated to interlock high- and low- voltage lines. The interlock system operating procedures and behaviour have been validated during real operating conditions of the detector esposed to a particle beam.

The satisfactory results of tests take the project to full completion, allowing the plan to reach the "exit" stage, when the thermal screen is ready to be installed in the Tracker and ready to be operated.

Acknowledgments

Whenever somebody refers to "St. Paul enlightened on the way to Damascus"... well, somehow I can relate (clearly without the faintest expectation to become a saint myself!); because this is exactly what has happened to me during these years spent at CERN: a light has always guided me through major and minor complications, giving comfort when I was down, providing stunningly bright insights on the world of physics and controls when I was erratically looking for a solution, and helping me become a better person.

This light has the name of Andromachi Tsirou, to whom goes my admiration for being a mentor, my deep gratitude for being my supervisor and my love to a friend of rarest qualities.

She has entrusted me, making me feel confident of myself and showing what it means being a charismatic leader with a human touch; her words are ... just worth as much as gold to me. Had I not met her, I would certainly not be what I am now.

Along came Piero Giorgio Verdini. If you think you are able to think fast, provide solutions which are not only effective but also have the sign of the genius, the perfection of the scholar and the unmistakable touch of elegance, well, you had better meet him. Piero Giorgio's heart is as large as much as his mind is bright: every time I was confronted with a new problem, not only he could provide a solution on the spot, but he went as far as offering his help straightforwardly. Isn't it amazing? Thank you, PGV, you are simply the best!

But life is not all the times easy: the first time I met Nikolay Luzhetskiy I almost cried: his handshake was so strong! Then, after a few months, I am ready to cry of joy for having met him: his generosity and dedication to his job made of the climatic chamber that he built a wonder: we use to refer to it as the "russian submarine" because it is so sturdy and so well built that we could probably be able to navigate the oceans with it!

I was desperately trying to model my process when I met Guy Baribaud; while I was attending his classes on "Analog and Digital Techniques in Closed Loop Regulation Applications" I thought: this man really knows what he is talking about! Now, after three years, I can say without hesitation that without Guy's contribution probably no thermal screen

model would exist, and I would have missed one of the most elegant approaches to controls. On top of that, the attention to details has given our testing procedures the rigorous structure that they currently have.

At the same time, should someone ask me what is a true leader is, I could provide an asnwer without hesitation: Gigi Rolandi. He has supervised my work and tought to me how to put leadership in practice, leading our CERN group with charisma. Moreover, he made me understand the meaning of the word "trust": Gigi trusted me and empowered me as much as he could to make me feel responsible of my job, treating me as a true professional.

It is also true that there is no professional growth whatsoever without facing new challenges: Ariella Cattai provided me with new opportunities to think about original issues related to my work, giving to me also the chance to present my ideas and discuss them with experts during several CMS collaboration meetings. Thank you, Ariella, for always believing in me!

This project actually involves a number of persons who have been dealing with it from the beginning: to Paolo Petagna goes my gratitude for having introduced me to the thermal screen design and problematics. Working with him is just so pleasant. When struggling with the thermal screen tests I was lucky to be not alone: Blanca Perea Solano provided a big smile along with a great devotion to our actvities.

The thermal screen test bench also needed a cooling plant, which has been provided by Paola Tropea; she has been so patient in handling my weird requests to keep the plant working overnight, in the early morning, in the weekend, during vacations... always. Thank you!

I would also like to thank Patrice Siegrist for his numerous and purposeful suggestions on how to test the thermal screen panels and for the invaluable help in all the logistics.

The barrack 598 at CERN is located underneath a water reservoir which, in french, is called *chateau d'eau* (water castle). During the years, the barrack's inhabitants have started to refer to

the place itself as *le chateau*; and now, this is how it is called among *connoisseurs*. I tend to believe that it works because the people in the barrack are of such noble souls and minds: Jean Martin, Duccio Abbaneo, Roberto Chierici, Matteo Risoldi, Michael Eppard, Alexander Dierlamm, Martin Weber, Bruno Wittmer, Georges Roiron, Flic Nicholson, Eric Albert, Jeannine Muffat-Joly. Others don't live in the 598, but are part of the Tracker community, and have played a role in this thesis work: Horst Breuker, Hans Postema, they all contributed to making my life there easier, more pleasant and more interesting.

If one has to work with Siemens PLC at CERN, meeting Jacky Brahy, Claude-Henry Sicard, Claude Dehavay, Raymond Lafay is a must; and it's thanks to them and to all the help they constantly provided that now we "play" with confidence with Siemens modules in the laboratory!

These three years spent at CERN have been just amazing; and my life would have certainly been less rich without colleagues with whom I shared the ups and downs of this thesis: they are many, and I shall mention one for all, Peter Cwetanski: without his precious help, this thesis would have been an unformatted text document; and, at the same time, I would not have been able to appreciate the value of true friendship and of vintage, classy Mercedes cars!

Then there is somebody who started as a boss or a colleague, and then became a friend: it's the ALICE HMPID group: Eugenio Nappi, Giacinto De Cataldo, Antonio Franco, Vincenzo Rizzi; starting from the *mythopoietic* discussions we used to have, they always supported me in these years, providing the same encouraging words and advices they gave to me while we were working together. Sometimes it still feels we are working together: they are my old family!

And it actually is the same family that Domenico Di Bari belongs to, and it does not come as a surprise: not only was he a great supervisor, but he made me feel admitted to the "club" of experimental physicists, trusting me all the times. Thank you, Nico, for helping me become a more mature person, and proud of my profession!

But there is another person, very dear to me, who endorsed my admission to the "club" of professionals working on High Energy Physics; his name is Luca Trentadue, and it is thank to him that I feel now part of a community, even if I am coming from a different background!

For their patience in proof-reading this thesis, I would also like to express my gratitude to Raffaele De Leo from University of Bari, and Franceso Corsi, from Polytechnic of Bari; not only they expressed their opinions in the most gracious way, but they also provided very useful suggestions on how to improve this thesis.

Being a student at the University of Bari also meant to me starting new and strong friendships; the constant support of Cosimo Pastore, Irene Sgura, Luigi Lagamba and Emanula Cavallo has been unvaluable when I felt like I wouldn't make it. It is thanks to them I could still feel at home when I was back in Italy! I would also like to mention the *XVII Ciclo di Dottorato* as one of the most incredible group of PhD students. I hope we will meet again, all together!

So far, so good for those who shared my scientifc work; but this work could not have been carried on without the discrete and efficient work of the secretariat of Geneva and Bari: Madeleine Azeglio, Dorothee Denise, Marie-Claude Pelloux, Nadejda Bogoulioubova, Dawn Hudson, at CERN, and Anna Massarelli, Claudia Ceglie, Maria Valentino in Bari.

They had a smile for me every time I popped up with some annoying request, and every time I left their office feeling lighter and more optimistic. They have been my guardian angels, that's it!

I like to think of a thesis as a Formula 1 Grand Prix: the pilot can be as good as you want, but without the Team, he would accomplish nothing. Those who wait at the pit stop to check that everything is in order, to give support and comfort, a smile and a hug, those who are there just waiting for him, when he will stop from time to time... Francesca Romana, for all her love and patience, and for always believing that good things would happen, from the very beginning... and Anna, Mario, Miriam, Franco, Elisabetta. My family. They are beautiful. They are just everything to me, and the love they give to me and that I feel for them is neverending.

1 CERN and CMS

1.1. LHC

LHC is so far the accelerator with the highest discovery potential built for high energy physics.

The very high energy density provided by this collider will create the conditions of the young universe shortly after the big bang, and will give the opportunity of new tests on the Standard Model (SM). Still numerous questions arising from the Standard Model could not be answered by previous accelerators due to energy limitations, motivating the development of an even bigger, new machine. The LHC opens new perspectives in high energy physics; it is the successor of the LEP (Large Electron Positron) collider (and also of Tevatron) and has no comparable counterpart worldwide.

In 2007, the construction will be finished and the experiments will be taking data for at least 10 years. The LHC will reuse the LEP ring tunnel with 27 km circumference, providing much higher energies for proton on proton (and alternatively heavy ion) collisions. While LEP was designed for a centre of mass energy of 200 GeV, LHC will reach 14 TeV with protons and 1312 TeV with Pb ions. Furthermore, the LHC bunch crossing (Bx) frequency of 40 MHz (corresponding to 25 ns every crossing) will be approximately three orders of magnitude higher compared to the bx frequency of LEP. On average, every bunch crossing will produce about 18 proton-proton intercations, generating 500 charged particle tracks. Compared to the LEP electron-positron collider, where collisions occurred rarely due to the low cross-sections of electrons and positrons, the collision rate will be almost 109 times higher in LHC.

The performance of a collider can essentially be characterized by two parameters: the energy in the centre of mass of the particle collision, and the luminosity L.

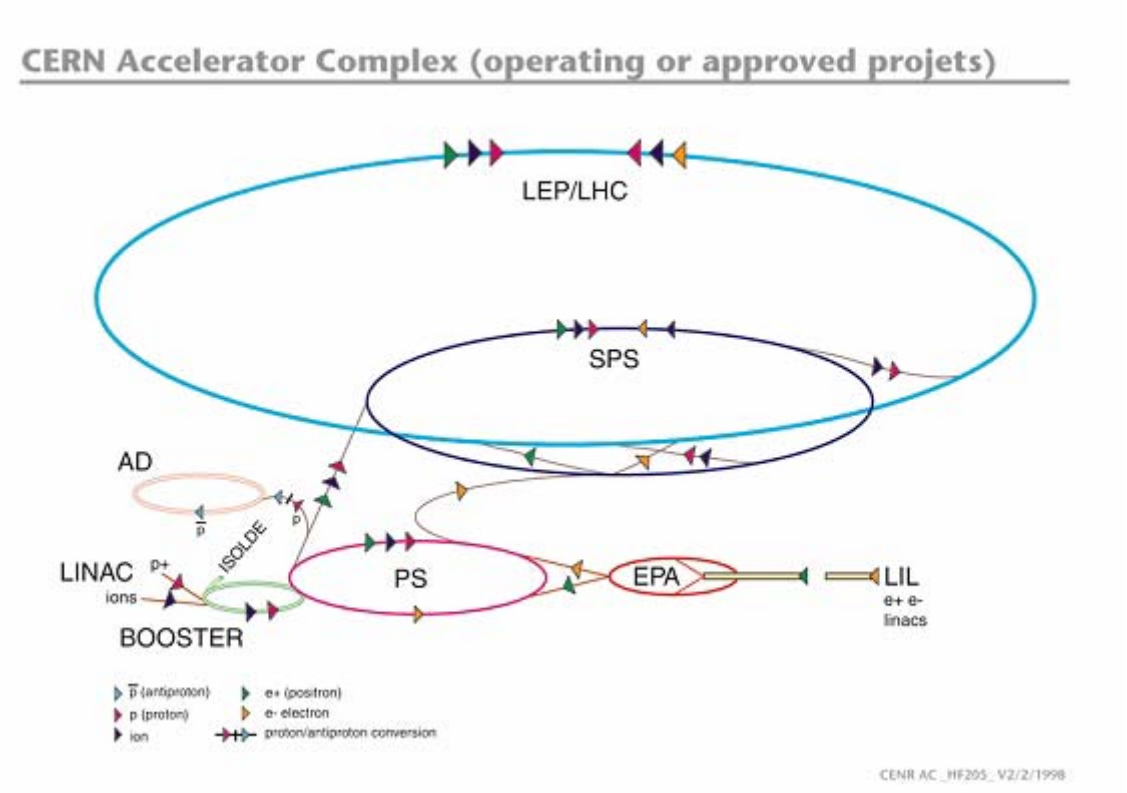


Figure 1-1 CERN Accelerator Complex

The luminosity depends on the number of particles per bunch in the two colliding beams, N_1 and N_2, on the bunch crossing frequency f, on the cross-sectional area of each particle bunch. If the particle distribution in the bunches is Gaussian, the luminosity can be expressed as:

Eq. (1.1)

$$L = \frac{1}{4\pi} \frac{N_1 N_2 f}{\sigma_x \sigma_y}$$

where σ_x and σ_y are the R.M.S. sizes of the distribution in the directions transverse to the beam, x and y. The LHC design luminosity is equal to $10^{-34}\ cm^{-2}\ s^{-1}$ and will be obtained with a bunch crossing frequency of 40 MHz, a particle density of about 1000 protons per bunch. An integrated luminosity of $5 \cdot 10^5\ pb^{-1}$ is expected over ten years of operation, corresponding to $5 \cdot 10^7$ seconds of running at high luminosity $10^{-33}\ cm^{-2}\ s^{-1}$. An integrated luminosity of $3 \cdot 10^4\ pb^{-1}$ will be accumulated during the low luminosity $1.7 \cdot (10)^{-34}\ \ cm^{-2}\ s^{-1}$ of start-up phase. The LHC beams, being of the same polarity, will circulate in two separate vacuum pipes.

A magnetic field of 8.4 Tesla is required in order to bend the trajectory of 7 TeV protons along the LEP tunnel. It will be created by superconducting coils, cooled at 1.9 K by superfluid

helium. Two antiparallel fields are needed to bend the two counter-rotating proton beams along the same circular path. The solution minimizing the cost of the magnets is a single iron yoke and cryostat shared by the two coils.

The beams cross at four points, where the experiments are located. The transverse dimensions of the beams at the intersections are reduced by additional focusing magnets, in order to achieve the luminosity values required by each experiment. ATLAS (A Toroidal Large Apparatus) and CMS, the two general purpose detectors of the LHC, are designed for physics studies at the highest LHC luminosity. ALICE (A Large Ion Collider Experiment) is a detector conceived for the study of heavy ion collisions, and LHC-B is a dedicated experiment for B physics (figure 1-2).

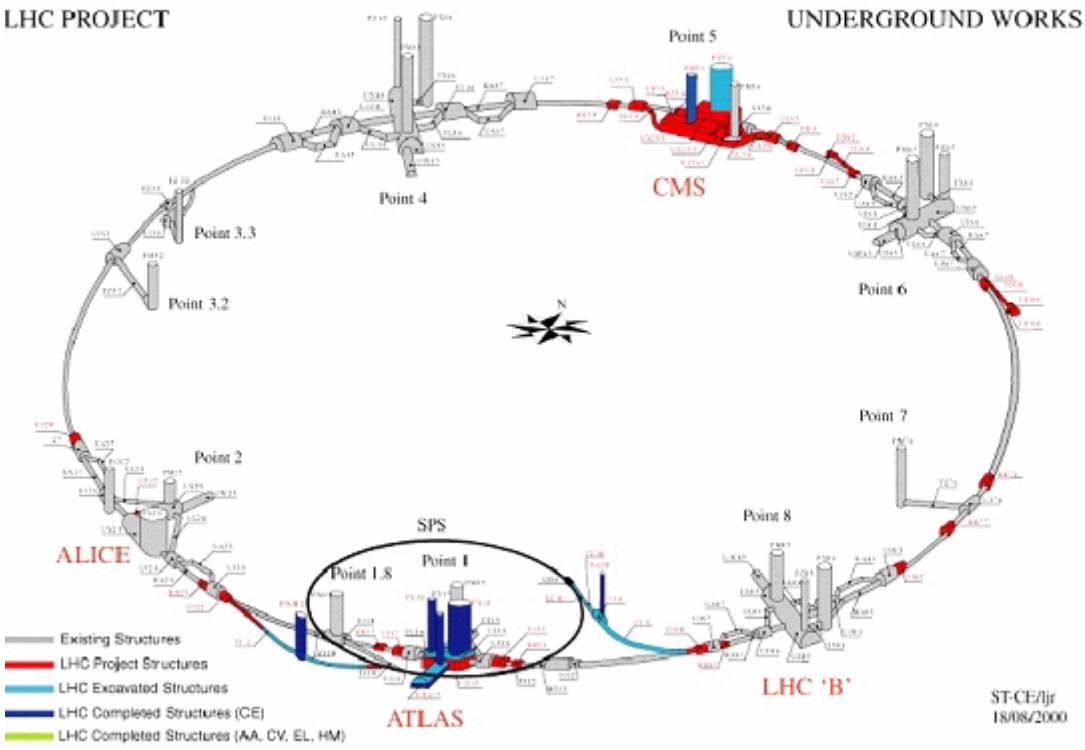


Figure 1-2 LHC Experiments

1.2. Physics Motivation

The main physics goal of LHC is the discovery of the Higgs boson which is anticipated by the Standard Model. Theory only provides an upper limit for its mass of about 1TeV, while LHC will reach much higher energies. In the last period of LEP, when energies were pushed

to the limit of the machine, a few possible Higgs candidates were observed suggesting a mass of about 114 GeV [59].

Due to the extended energy range of LHC, these particles will be undoubtedly confirmed and characterized if they exist as predicted by the Standard Model. Although 40 million bunch crossings occur per second in each of the four LHC experiments, the Higgs boson is expected to appear only about once every day, yet it is enough to accumulate good statistics.

Another goal is the Charge-Parity (CP) Violation. At an early stage, the universe was dominated by energy. While expanding and cooling down, gradually matter and anti-matter formed and became dominant. However, today's world appears to be entirely made of matter. The CP Violation implies an asymmetry between matter and anti-matter behaviour. This could explain today's dominance of matter.

First reported in the 1960s, several experiments have measured the CP violation first in the Kaon decays and then in the B-meson decays, as in the recent BaBar experiment at SLAC (Stanford Linear Accelerator) [11].

LHC will enter a new energy range, allowing to study the CP violation on B-mesons; in particular, the LHCb experiment will be dedicated to this study.

Another domain of interest will be represented by the search for supersymmetic particles (sparticles), which are predicted by supersymmetry theories (SUSY). According to SUSY, sparticles are the supersimmetric partners of the elementary particles we know. Since they have not been observed so far, SUSY must be a broken symmetry, which means that sparticles have masses different than their counterparts. The SUSY masses are expected in the TeV range, which makes them visible to LHC. Theory predicts at least five SUSY Higgs bosons, which can provide an explanation for the dark matter of the universe.

During a few weeks every year, LHC will work with lead ion beams instead of protons; the energy density, in this case, is much higher. The ALICE experiment is expected to reconstruct a very early stage of the universe called quark-gluon plasma, which may reveal different physical properties.

1.3. The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a detector designed for discoveries at the highest luminosity available in proton collisions at the LHC. Like most of modern collider experiments it comprises several subdetectors of complementary functions: muon chambers, calorimeters, and tracking system.

Figure 1-3 shows a three-dimensional view of the CMS detector.

The detector has a cylindrical shape, with an approximate symmetry in the azimuthal angle Φ. Its overall dimensions are 15 m in diameter and 21 m in length. It will be built around a superconducting solenoid generating a uniform magnetic field of 4 Tesla inside the coil. The magnetic flux is returned through a thick saturated iron yoke instrumented with muon chambers.

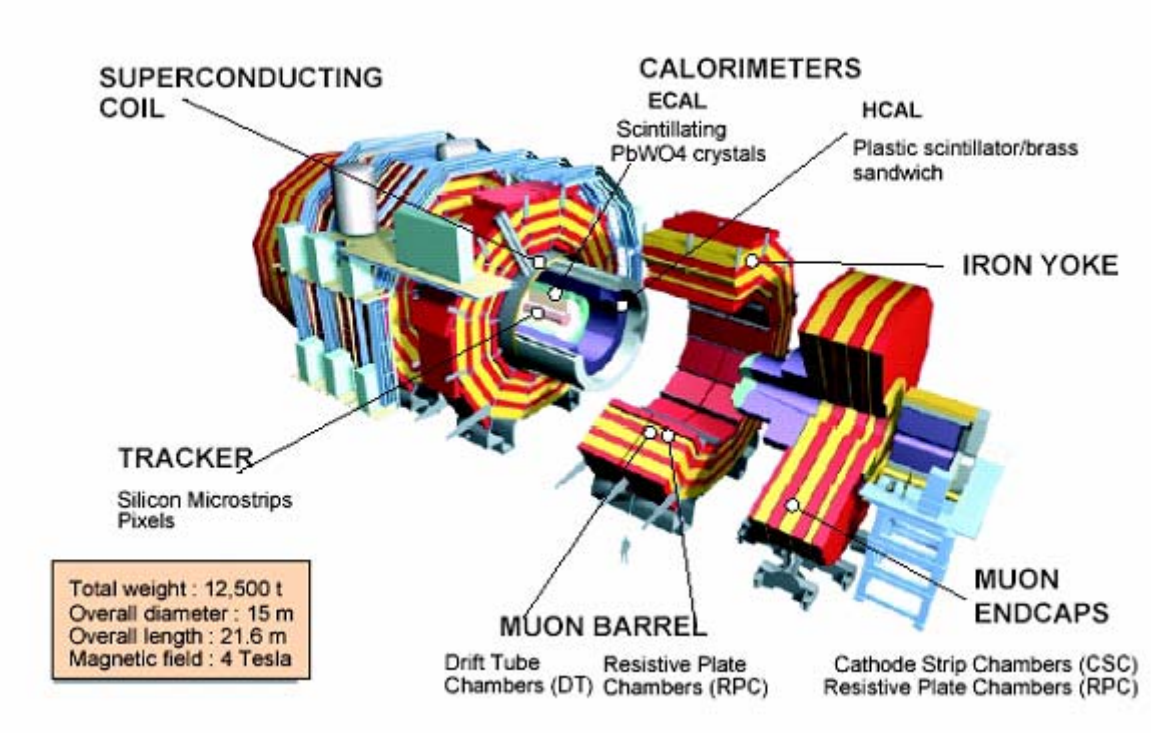


Figure 1-3 CMS

Thanks to the large coil dimensions, 13 m in length and 3 m in radius, both tracking system and calorimetry can be accommodated inside the magnet coil. This prevents the performance of the calorimeters from being affected by the coil material.

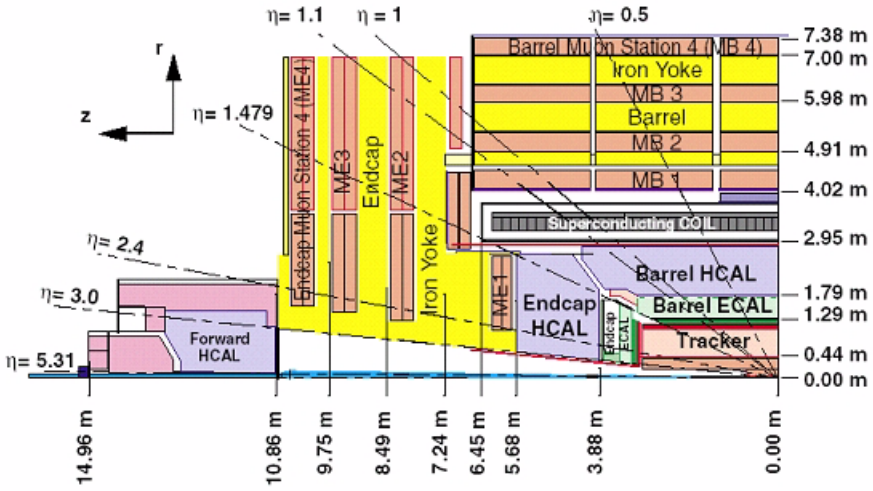


Figure 1-4 CMS Detectors

A schematic of the particles path in CMS is given in figure 1-5.

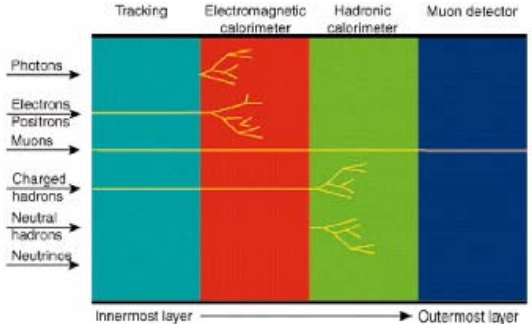


Figure 1-5 Particle tracks

As for the Trigger and Data Acquisition System, they consist of four blocks.

The first two stages, frontend detector electronics and first level trigger processors, are synchronous and pipelined. The first level trigger must reduce the 40MHz bunch crossing rate to an event rate of 100 kHz by filtering only interesting events. To avoid dead time, the data collected within this period must be stored in the front-end in order to pass it on after a trigger request.

The two later stages are a large switching network ("event builder") with a throughoutput of 500Gbit/s and an on-line event filtering system implemented in a computer farm. These stages are made of commercial components and thus can be upgraded as technology develops. The combined information of all detector subsystems is used for the total event reconstruction and quantification. While the silicon tracker is intended for momentum and polarity identification, the energy is measured by the calorimeters, and penetrating muons are detected in the outermost layer. The triggering information is derived from calorimeters and the muon system.

1.3.1. The CMS detectors

- **The central tracking system**

The detector closest to the point of collision is the central tracking system, whose purpose is to reconstruct the trajectory of charged particles [3].

The central tracker consists of three pixel layers and ten microstrip layers, with a sensitive are of more than 206 m^2.

The tracker constitutes a substantial amount of material in front of the electromagnetic calorimeter, which makes the measurement of photons and electrons difficult. The reason for this thick tracker is the high particle rate expected at the LHC, which can only be handled with

a large number of small detection cells.Its task is to measure the tracks of charged particles with a minimum of interaction.

In particular, the tracker allows measuring the momentum and direction of charged particles at their production vertex. In vacuum, in a uniform magnetic field, charged particles describe a helicoidal trajectory around the field axis. The curvature of the track is directly related to the component of the momentum in the plane transverse to the field: $P_T = qBR_c$, where q is the particle charge, B is the magnetic field intensity and Rc is the radius of curvature.

In a practical unit system this expression reads P_t[GeV/c]=0.3 B[T] Rc[m], for a particle charge e.

Thus in principle the transverse momentum can be determined by measuring the (x,y) or (R, Φ) coordinates of a set of points along the track in the plane transverse to the magnetic field, and by fitting a circle through the measured points. In practice the situation is more complicated, since the particles lose energy and are scattered in the tracker material.

The particle momentum is computed as $p = \frac{p_T}{\sin\theta}$, where θ is the angle of the particle direction with respect to the field axis. The value of θ is calculated from the z-coordinates of the measured points.

- **The hadron and very forward calorimeters**

The main goal of the hadron calorimeter (HCAL) is to measure the energy and direction of hadron jets.

An absorber thickness of 7 nuclear interaction lengths is required in order to contain 95% of the energy of a hadronic shower. The absorber chosen for the CMS hadron calorimeter is copper, as it is non-magnetic and has a short interaction length (15 cm). Copper layers will be interleaved with plastic scintillator tiles read out by wavelength shifting fibers. The main part of the hadron calorimeter (HCAL) is located inside the magnet, which is surrounded by an additional small part in the central region ("tail catcher").

The central HCAL consists of a brass/scintillator sampling calorimeter. Its scintillation light is captured, wavelength shifted and guided to hybrid photodiodes. The forward part of the HCAL consists of a steel absorber with quartz fibers. Traversing charged particles produce Cherenkov light in the fibers which is guided to photomultipliers. The return yoke will be instrumented with scintillator tiles to form an outer hadron calorimeter.

- **The electromagnetic calorimeter**

An electromagnetic calorimeter of excellent energy resolution is indeed required in order to permit the discovery of the Higgs in the two photon channel, which letic calorimeter (ECAL) consists of approximately 76000 scintillating $PbWO_4$ crystals. Electrons and photons are converted to light pulses, which are read out by silicon avalanche photodiodes.

The principle of the electromagnetic calorimeter (ECAL) is to absorb the energy of photons and electrons and to deliver a signal proportional to the deposited energy.

- **The muon detection system**

The muon system [58] is composed of four muon stations interleaved with the flux return yoke plates. It consists of four stations in both barrel (MB1…MB4) and endcap (ME1…ME4) parts, which are integrated in the iron return yoke of the magnet. In the barrel part, each station consists of twelve layers of Drift Tube Chambers (DT). Resistive Plate Chambers (RPC) are used for bunch crossing identification and provide a cut on the muon transverse momentum at the first level trigger. In the endcap region, each station consists of six stations of Cathode Strip Chambers (CSC).

The detector is divided into a central part ($|\eta| < 1.2$) and a forward part $0.9 \le |\eta| \le 2.4$. The muon detector should fulfil three basic tasks: muon identification, trigger, and momentum measurement. Muon identification relies on the fact that, in contrast with most charged particles, muons do not interact much with matter. The muon detector can be placed outside the magnet coil, after the calorimeters, which reduces the hadron background.

Furthermore the total thickness of absorber before the last muon station amounts to 16 nuclear interaction lengths, so that only muons can reach it.

As part of the new physics shows up by the presence of muons of high transverse momentum in the final state, the muon detector has to take part in the trigger decision. It must therefore be composed of fast detectors, delivering a signal that can be associated to a single bunch crossing, and capable of providing a fast estimation of the muon transverse momentum in order to allow P_t cuts at the trigger level.

The CMS muon system is requested to measure the transverse momentum with an accuracy between 8% and 40% for muons of P_t between 10 and 1000 GeV/c. This can be done in two ways: as the muon track gets curved by the magnetic field in the return yoke, a sagitta measurement allows to determine the muon transverse momentum; the second way is to measure the muon direction in the transverse plane just after the magnet coil. For muons originating from the beam crossing point the angle of the track with respect to the radial direction is directly related to the transverse momentum.

1.3.2. The tracker architecture

A covered area of 206 m^2 makes the CMS tracker the largest silicon detector ever built. The sensors are arranged in a total of about 16000 modules, which consist of one or two strip detectors in series together with the associated readout electronics. The barrel modules will be placed on the surface of cylindrical support structures. To allow better area coverage, the modules will overlap like roof tiles. A carbon fiber frame holds one or two silicon sensors

which are connected to the readout hybrid via a pitch adapter. On both ends of the frame, cooling pipes are sinking the heat produced by sensors and electronics

The main structure consists of a central (barrel) part with three pixel and ten strip layers and the disk and endcap sections with two pixel and twelve strip layers (Fig. 1-6).

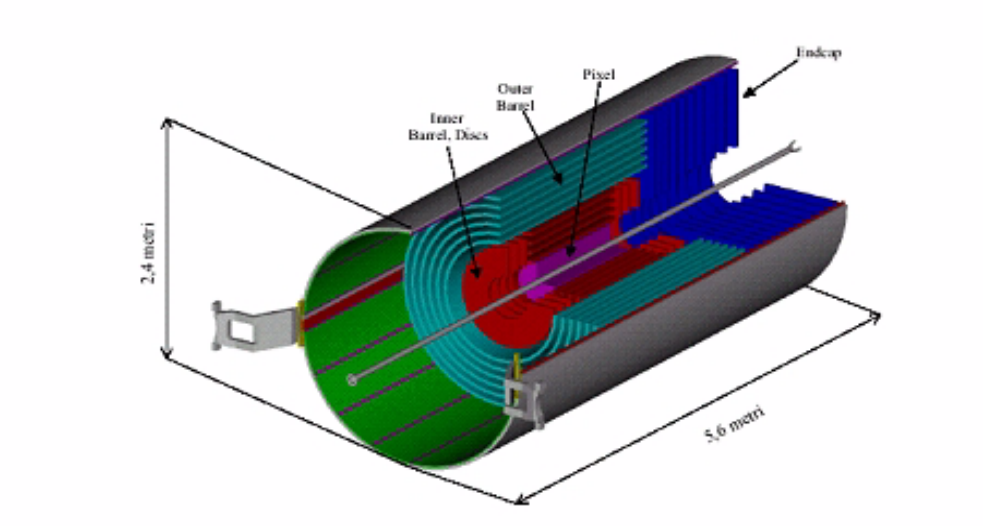


Figure 1-6 CMS Tracker

In figure 1-7 is shown a cross section of one quadrant.

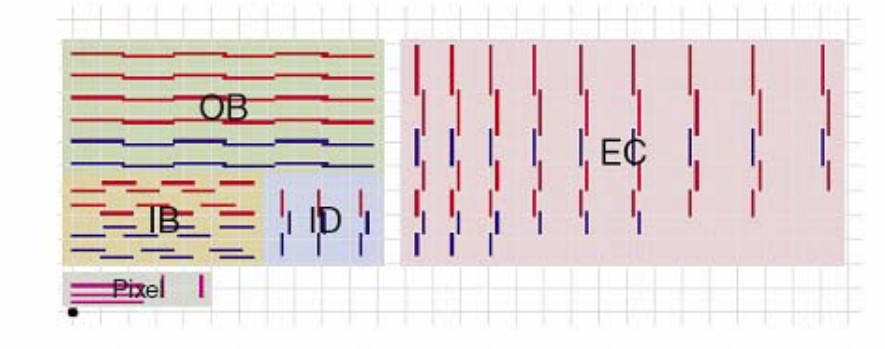


Figure 1-7 One quarter of the CMS Tracker layout. Pixel detector layers are shown in purple, while strip detectors are in red (single-sided) and blue (double-sided).

The pixel layers in barrel and endcap parts are shown in purple, while the strip layers are drawn in red (single-sided detector modules) and blue (double-sided detector module). The double-sided modules are made of two single-sided detectors mounted back to back, with a 100 mrad angle between their strips. Thus, these 'stereo" modules deliver two-dimensional hit positions.

The number of detector layers is a trade-off between tracking efficiency, material budget and cost. On one hand, the number of hits increases with the number of layers penetrated, which makes the track reconstruction easier. On the other hand, the amount of material within the tracker should be kept as low as possible because multiple scattering, which spoils the tracks, is proportional to the amount of material traversed by the particles. An even tougher constraint is the cost of the tracker, which reduces the number of layers to an affordable design. In an average event, about 750 charged particles are produceed from each bunch crossing, which gives a few thousand hits in the tracker.

The simplest approach to track reconstruction from a set of hit points is to start with a pixel hit in the innermost layer and project a cone onto the next layer in radial direction. If no hit can be found there, the starting point was either noise or a particle of very low energy which get stuck or was detected by multiple scattering, so the original hit can be discarded. Otherwise, the procedure can be repeated until finally the full track through all planes is found. Of course, the procedure is much more complicated in reality: dead or inefficient regions have to be taken into account (e.g., by skipping a layer) and the magnetic field bends the tracks depending on the particle momentum. Since there is a lot of low-momentum background in the innermost part of the tracker, a more advanced concept starts its track search from the outside. With this approach, a preselection of interesting tracks is provided by the first-level trigger, which is obtained from calorimeter and muon detector data.

Fig. 1-8 shows the functional groups of the CMS silicon tracker.

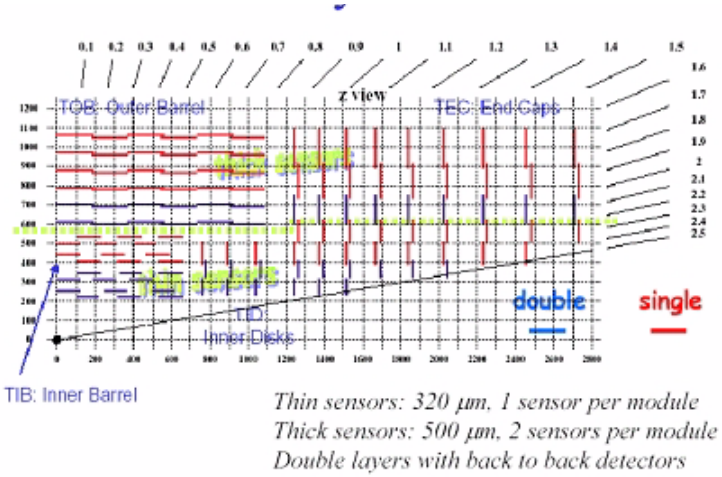


Figure 1-8 Tracker's Functional Groups

The ten strip layers in the barrel are divided into the inner barrel (IB) and the outer barrel (OB), which are numbered in ascending order with the radius. The seven rings of the disk modules are divided into the three inner disk layers (ID) and the seven endcap layers (EC), again numbered with ascending radius.

The tracker sensor operating temperature will be -10 °C. This is required by the silicon sensors, which suffer from radiation damage. Defects are "frozen" so they can not gradually decrease the detector quality, as will be discussed later.

1.3.3. Silicon Sensors

The principle of solid state detectors is based on the energy loss of traversing particles. Free electron-hole pairs are generated, which move towards opposite electrodes under the influence of an electric field. The energy loss of heavy particles in matter was described by H.A.Bethe and F.Bloch [7].

The ionization of particles traversing the detector leads to the creation of free electron-hole pairs. Fig. 1-9 shows the principal layout of a solid state detector with opposite electrodes.

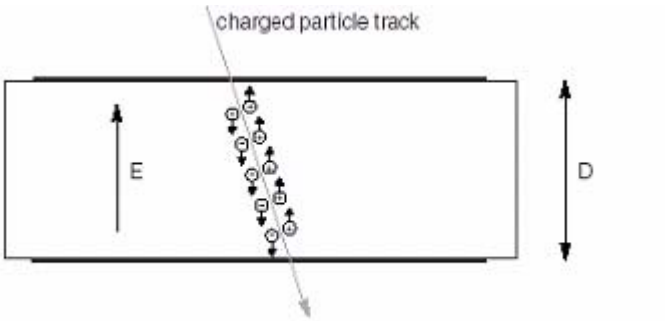


Figure 1-9 A charged particle traversing the detector generates free electron-hole pairs along its track, which are moved by electric field.

An electric field between the electrodes is required to move the carriers according to the relation $v=\mu E$. While the charges move inside the detector bulk, a current is induced in the electrodes, no matter whether the carriers finally reach the electrodes or not. This current i is proportional to the sum of all carrier velocities, with the elementary charge e and the detector thickness D.

Eq. (1.2)

$$Q_c = \frac{e}{D}\int\left(\sum v_e + \sum v_h\right)dt$$

The integrated current gives the total collected charge Qc, which is usually measured with integrating and, thus, charge-sensitive amplifiers. After the generation of a free electron-hole pair, the electron moves to the positive electrode while the hole moves to the negative. If no charges are trapped, the sum of the distances they travel equal the detector thickness,

regardless of their initial position. Thus, the integral term in Eq. 1.2 equals the detector thickness multiplied by the number of pairs n, and the total collected charge is Q_c =ne.

The collected charge is stated in terms of electrons, which might mislead to the false conclusion that only the electrons contribute to the charge measured at the electrode. In fact, both electrons and holes are responsible for the charge collection in equal parts, since on average both carriers travel through half of the detector if no charge trapping occurs.

In usual silicon detectors, virtually all charges finally reach the electrodes. After heavy irradiation however, charge traps are created and the mean travel distance shrinks. In other solid state detector materials, such as diamond or gallium arsenide, charge traps are always present, resulting in a mean pair travel distance less than the detector thickness.

Depending on the sensor material, the electric field can be homogeneous or not. Diamond detectors merely consist of a thin film with ohmic contacts on both sides. An applied voltage results in a homogeneous field between the electrodes. With silicon however, it is impossible to operate a detector this way, because the intrinsic number of free carriers due to thermal excitation would be orders of magnitude higher than the expected detector signal. Thus, a reverse-biased pn-junction is introduced, by doping with acceptors and donors resulting in zones of p and n types. Starting with a homogeneously doped material, a thin layer with a high doping density of the other type is applied onto the surface, resulting in a pn-junction. On the opposite surface, which is known as backplane, the bulk type doping is enhanced to get a good ohmic contact [7].

Fig. 1-10 shows the schematic layout of a silicon detector based on n-type bulk material, which is mostly used. Once the junction is under reverse bias, all free carriers in the bulk are drained by the electric field. The thickness of the implants with high doping concentration (p+ and n+) is in the order of a micrometer, so that the difference between bulk and total detector thickness is negligible.

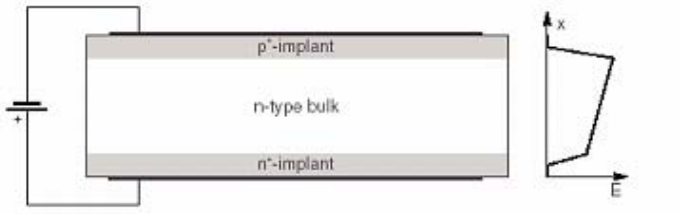


Figure 1-10 Schematic of a Silicon Detector

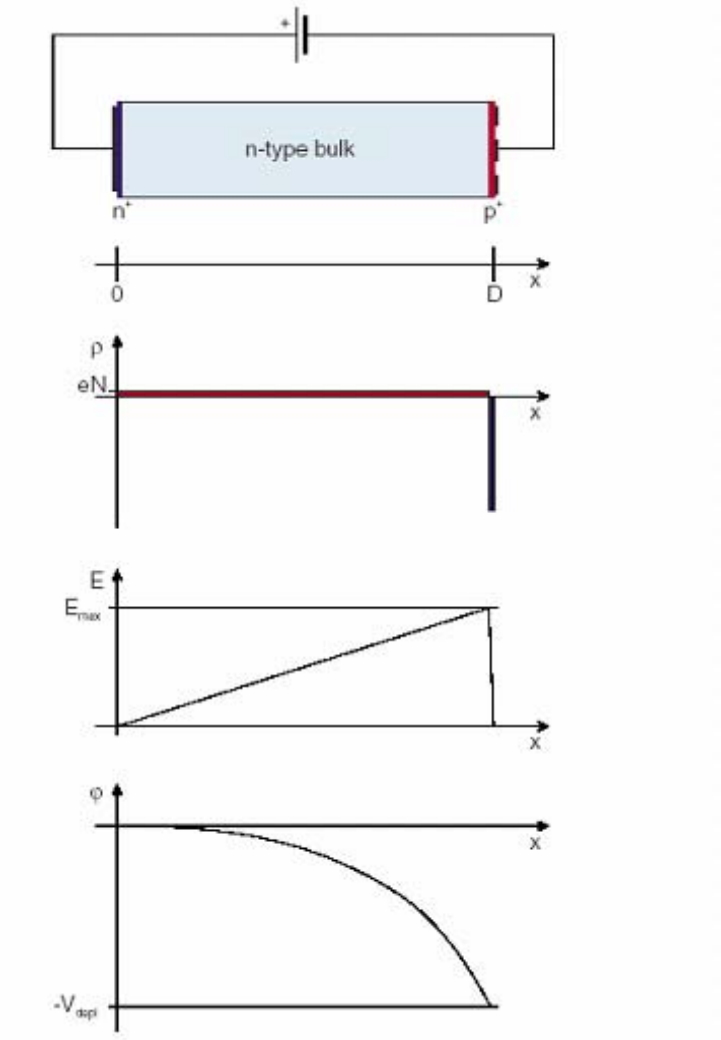


Figure 1-11 Charge density, electric field and potential in a one-dimensional model of a silicon detector at full depletion.

The bulk charge is constant over the full width of the bulk. Since the global charge must be balanced in a steady state, positive and negative charges have to match. This condition leads to the width of constant non-zero implant charge density which is equal to the bulk width but scaled down by the ratio of doping concentrations,

Eq. (1.3)

$$d_p = D \cdot \frac{N_{Bulk}}{N_p}$$

The charge density ρ can be obtained by the Poisson equation

Eq. (1.4)

$$\frac{dE}{dx} = \frac{\rho}{\varepsilon}$$

where ε is the dielectric constant. In the case of full depletion, the electric field has a triangular shape, rising from zero at the backplane (x=0) to its maximum at the junction (x=D). In the implant, the field quickly drops to zero again. The maximum electric field E_{max} is given by

Eq. (1.5)

$$E_{max} = \frac{q \cdot D \cdot N_{bulk}}{\varepsilon}$$

The relation

Eq. (1.6)

$$\frac{d\varphi}{dx} = -E$$

defines the electrical potential φ. The potential difference between backplane and implant electrodes is the voltage V applied to the detector. At full depletion, V_{depl} is given by

Eq. (1.7)

$$V_{depl} = \frac{eN_{bulk}D^2}{2\varepsilon} = \frac{D^2}{2r\mu_e\varepsilon}$$

If the applied voltage is higher than the depletion voltage, a constant offset adds to the electric field. With $V \gg V_{depl}$, the electric field can be approximately considered constant. When the applied voltage is below full depletion, the field does not extend over the whole bulk. With its maximum still at the junction, only a fraction of the sensor is depleted. Obviously, the charge collection is inefficient in that case, since the carriers do not mov outside the electric field. The width of the depletion zone, or collection distance d_c, is given by

Eq. (1.8)

$$c_c = D\sqrt{\frac{V}{V_{depl}}}$$

Thus, the efficiency of a silicon detector is given by

Eq. (1.9)

$$\eta_e = \sqrt{\frac{V}{V_{depl}}} \qquad for \qquad 0 \le V \le V_{depl}$$

$$\eta_e = 1 \qquad for \qquad V \ge V_{depl}$$

In a homogeneus electric field the carrier velocities are constant. Initially all carriers move towards the electrodes and gradually aredrained at the electrodes until all charges are gone. As the mobilities of electrons and holes differ, the current waveform is a superposition of two triangles of the same area, but different slope.

1.3.4. Radiation Damage

The total fluences of photons, neutrons and charged hadrons expected in the CMS experiment over the scheduled 10 years of LHC operation is shown in fig. 1-12, where z is the distance from the vertex along the beam axis, while the other parameter is the radius r.

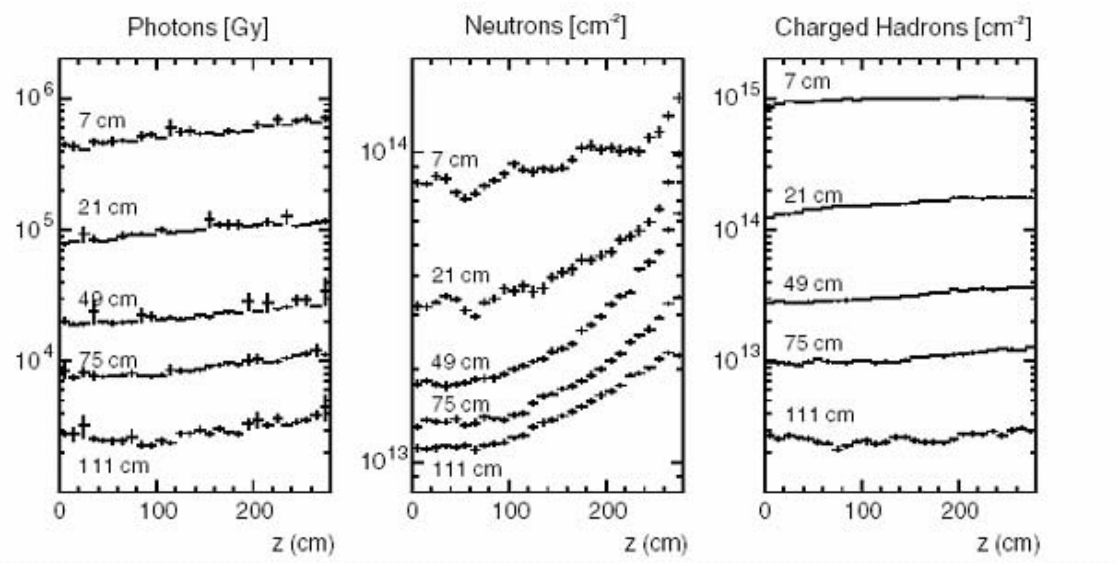


Figure 1-12 The expected radiation fluences of photons, neutrons and charged hadrons in the CMS experiment over 10 years of operation as a function of the distance z from the collision point along the beam axis and the radius r.

After irradiation, the increased current is still changing. There is a beneficial short-term effect called "annealing" with a time constant of a few days at room temperature. Unfortunately, it is followed by a deterioration effect called "reverse annealing" in the long run (about one year at room temperature). Both effects are strongly temperature dependent. At room temperature, annealing initially causes the leakage current to decrease, while later it rises due to reverse annealing process, until it finally saturates at a value which is significantly above the initial level. Due to the intrinsic positive feedback in the sensors, this can in turn lead to the situation in which *thermal runaway* happens: the increased dark current heats the silicon, which in turn causes an increase an increase in the dark current (approximately a factor 2 every 6 °C of temperature change), quickly leading to catastrophic failure and permanent damage.

At -10 °C however, both effects are virtually frozen, so the detector current remains constant. Thus, irradiated detectors in general should be operated and stored at low temperature, while it is favourable to shortly expose them to room temperature (for handling, service, transportation etc.) to take advantage of the beneficial annealing [60], [61].

Additionally, operation at -10 °C also reduces the dark currents by a factor of thirty with respect to the room temperature, thus further reducing the potential for thermal runaway.

1.3.5. The tracker control system

The Distributed Interlock System is going to be treated as a safety instrument function operating in the continuous mode (see 4.3).

The goal of the Distributed Interlock System is to handle the interlock lines of the power supplies of the Tracker subdetector, as shown in figure 1-13. The Tracker subdetector must be kept at a temperature around -10 °C throughout its lifetime, expected to be longer than 10 calendar years. A sophisticated and reasonably fail-safe cooling system, adapted to the Tracker needs has been designed and will be supplying the Tracker support structure with coolant, its role accentuated by the thermal screen. However, in case of failure of the cooling system, blockage of part of the cooling structure, risk of condensationor the occasional maintenance, any possible heat source inside the Tracker must to be shut off. It has to be noted that the Tracker has no direct (hardware) connection to the controls of the cooling system, which operates autonomously.

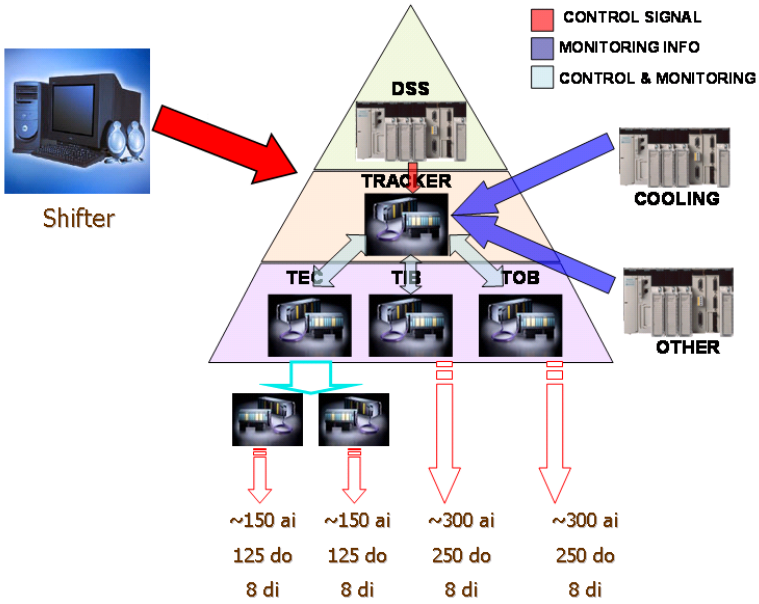


Figure 1-13 Tracker Control System Architecture

The main risks to which the tracker will be subject are:

- Temperature increase due to a cooling plant failure, leading to shortened detector lifetime (Reverse Annealing) and possibly to catastrophic failure (Thermal Runaway);

- Condensation or sublimation of water due to a dry-gas plant failure, leading to local temperature increases (formation of thermally insulating ice layers), possible contamination of the silicon sensors, possible mechanical damage, and possibly spectacular failures (electrical short-circuits).

- Sudden drop of the Low Voltage and High Voltage power supplies due to a main power cut.

The Distributed Interlock System consists of three separate hardware layers: the sensors which monitor the variable (temperature) describing the thermal state of the Tracker, the PLC system which here can be considered as an intelligent relay system controlling the Tracker state and setting the interlock line status, and the interface of this system to the CMS experiment.

As an additional complication, even in case of a main power cut, the Tracker power supplies must be ramped down gently rather than set abruptly to "OFF".

The reason for this is to avoid any sudden change of the bias and supply voltages which would result in fast voltage spikes across the AC-coupling capacitors integrated in the sensors. On one hand, this could generate "pin-holes" by breakdown of the thin dielectric layer, which would subsequently create DC connections between the detector strips and the readout electronics. On the other hand, these spikes would also be transmitted to the front-end electronics stages and simply destroy them. Both effects would obviously have catastrophic consequences for the Tracker.

The generic scheme for facing such a problem with a control system would be to have the control system sense the mains cut and send a signal to the power supplies which drives them to ramp down. This requires the power supplies to be equipped with permanent backup power supplies, with enough capacity to maintain power throughout the time it takes to ramp down. This interlock, referred to as Power Cut Interlock has been directly implemented on the Tracker Power supply without requiring the intervention of the external interlock system.

From the above description of the Tracker interlock system it becomes obvious that the main Tracker element controlled is the Tracker Power supplies.

It is therefore obvious that the interlocking system will need to posses sufficient intelligence to be able to address the several power supply groups individually, adjust to variable thresholds for the temperature interlocks and be capable of timing its actions correctly. It should also be capable of dynamically reconfiguring the power supply groups and the alarm thresholds without being interrupted (and thus leaving the Tracker unprotected) and generally it should be reliably operational over the whole lifetime of the Tracker (approximately 10 years).

2 Control Systems

A control system is a set of hardware and software components used to monitor and control one or more devices.

When multiple devices are being controlled in a co-ordinated fashion this collection of devices is often referred to as the "process".

A control system may be passive, that is to say it provides information on the status of the process to an operator who then decides whether control actions are required and if so submits the necessary commands to the system. The control system then passes these commands to the appropriate devices in the form understood by that device. In the case of simple devices this may be an electrical signal with well defined characteristics, whereas in the case of a more complex device this may be a complex formatted message. Control systems are often, as in this case, not purely passive and in many cases may be programmed to perform actions in an automated fashion.

When considering large scale systems, such as plants in the process industry or, as in this case, large high energy physics experiments, control systems are usually highly structured and are typically hierarchically decomposed into different layers. Based on cost and performance functions, the optimal operating point for the plant is determined by the top layers. The lower layers implement the setpoints corresponding to the operating point and focus on the avoidance of input saturations. The lowest layer, implementing the basic feedback control, is usually also highly structured, consisting mainly of single-input-single-output control loops.

The structural problem implies selecting the variables to control, the manipulated variables (inputs), measurements, and outputs to use for control and the structure of interconnection between inputs and outputs. For large scale systems, in particular, the decision about the control systems structure is fundamental for the achievable performance. A poor choice of structure will impose control performance limitations that can not be overcome by any advanced parametric design.

2.1. SCADA

Supervisory Control and Data Acquisition (SCADA) is a system that allows an operator to monitor and control processes that are distributed among various remote sites. There are many processes that use SCADA systems: hydroelectric, water distribution and treatment utilities, natural gas, etc. SCADA systems allow remote sites to communicate with a control facility and provide the necessary data to control processes. For many of its uses, SCADA provides an economic advantage. As distance to remote sites increase and difficulty to access increases, SCADA becomes a better alternative to an operator or repairman's visiting the site for adjustments and inspections. There are four major elements to a SCADA system: the operator, the master terminal unit (MTU, often a PLC), communications, and the remote terminal unit (RTU, often a PC Workstation). The operator exercises control through information that is depicted on a video display unit. Input to the system normally initiates from the operator via the master terminal unit's keyboard. The MTU monitors information from remote sites and displays information for the operator. The relationship between MTU and RTU is analogous to master and slave. Depending on the complexity or sophistication the MTU may employ heuristics embedded into its programming which allow it to make modifications to the system to maintain optimality. In the same fashion, the sophistication in the RTU may allow local optimization of functions. From a software point of view, a SCADA system is essentially a toolkit used to build supervisory systems [1] [2]. As such, it includes a number of standard facilities:

- Development environment
- Set of basic functionality (e.g. HMI, Trending, Alarm Handling, Access Control, Logging/Archiving, Scripting, etc.)
- Networking/redundancy management facilities for distributed applications
- Flexible & Open Architecture- Support for major PLCs and OPC
- Powerful Application Programming Interface (API)
- Many other software interfaces: Open DataBase Connectivity (ODBC), Dynamic Data Exchange (DDE), Object Linking and Embedding (OLE), ActiveX and Web connectivity.

2.2. Hierarchical decomposition of control systems

Plants in the process industry (and particle physics detectors) usually have hundreds or thousands of measurements; thus, control systems are usually hierarchically structured in several layers in order to increase the flexibility of the system.

The lowest layer in a plantwide control system is the regulatory control layer. This is the layer where the manipulated variables and the outputs are connected with controllers. This layer is typically highly decentralized, meaning that most of the controllers are single-input/single-output controllers.

The second layer, the supervisory control layer, does not act on the manipulated variables of the regulatory layer, but on the setpoints of the control loops of the lower layer.

The main task of the third layer, the supervisory, is to handle constraints on the manipulated variables as well as on the controlled outputs of the system.

Thus, the objective is to avoid saturations in the inputs and to keep certain variables within specified bounds, imposed by security and environmental considerations (e.g., temperatures, pressures, etc.).

The Supervisory level also provides the interface to the operator of the system as well as providing functions such as alarm handling, logging and interfacing to other systems. At this level some form of high level processing which combines information from many of the lower level elements is performed. An example of this might be the Finite State Machine (FSM) modelling of the system for automated control.

The regulatory and the supervisory control layer together form the control layer. Above this layer there are usually one or two optimization layers. The optimizations determine the setpoints for the control layer in order to maximize performances and economic profit. They are usually performed on a daily or weekly basis using steady-state models of the overall plant and/or models of parts of the plant.

Figure 2-1 shows a typical hierarchical decomposition of a plantwide control system.

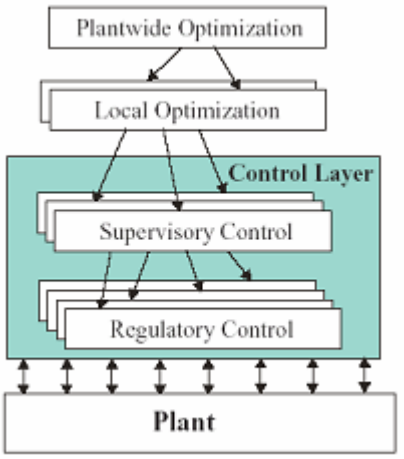


Figure 2-1 Hierarchical decomposition of a control architecture

The functionality of the regulatory control layer may be broken down further into two logical levels.

At the lowest level, the Field Management level is responsible for the physical interaction with the devices connected to the control system. As such the field level acquires input signals and distributes output signals. The Input/Output (I/O) signals are processed at this level to convert them to a form understood by the higher levels of the control system. A typical example would be an analogue to digital converter (ADC). Typically the signals are acquired

through individual connections. Another role of the Field Management level is to reduce the large number of I/O connections to a smaller number of connections to the next level in the control system (for example via a field bus). In this way multiple I/O parameter values can be transported over a single cable to the higher levels in the control system. At the Process Management Level any closed-loop control processing (often called regulation) that is required is performed.

Also at this level some data reduction might be performed.

One example would be that I/O parameters required for the closed-loop processing but not required by the operator of the system may be not be passed to the next level in the control system. Another form of data reduction might be that only parameter values that have changed by more than a certain defined amount are passed upwards. Again parameters received from multiple devices at the field level may be communicated to the Supervision level by a reduced number of connections.

In Figure 2-2 the SCADA implementation on the CERN experiments is shown.

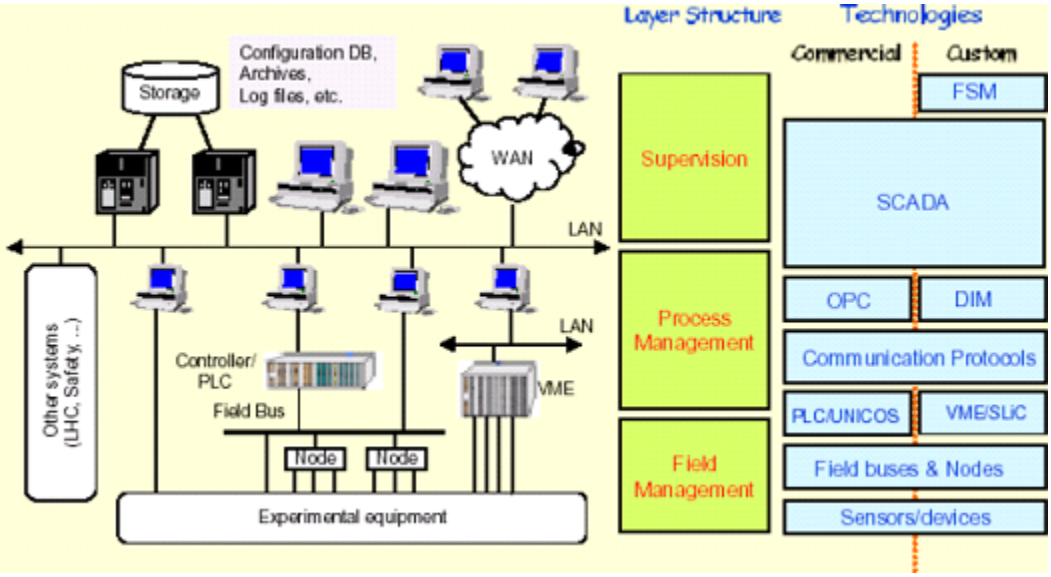


Figure 2-2 SCADA implementation at CERN

2.3. The regulatory layer and its model

Process control deals with systems that change in time.

In a typical control system, the Process Variable is the parameter to be controlled, such as temperature, pressure, or flow rate. A sensor is used to measure the process variable and provide feedback to the control system. The Set Point is the desired or command value for the process variable. If the Process Variable is less than the Set Point, then the Actuator Output will increase. This is called a Closed Loop Control System because the process of reading sensors and calculating the desired actuator output is repeated continuously and at a fixed rate.

The main task for the regulatory control layer is to reject immeasurable disturbances and to track the setpoints provided by the upper layer as tight as possible [15]. For a general system it can be described by the block diagram in figure 2-3.

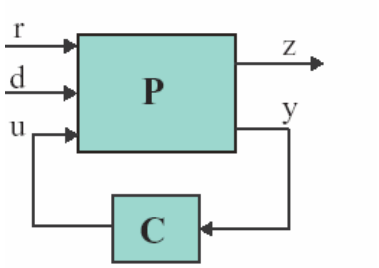


Figure 2-3 A General Feedback System

The displayed variables are

- • z - controlled outputs
- • y - outputs (to be used for feedback control)
- • u - manipulated variables
- • r - setpoints for z
- • d - disturbances

The plant P is modelled by the transfer functions G, Gd, Gyu and Gyd defined by
$z = G(s)u+Gd(s)d$
$y = Gyu(s)u+Gyd(s)d$

Control structure design involves the following steps:

1. Selection of the controlled variables z (controlled outputs with setpoints)
2. Selection of inputs u (actuators)
3. Selection of measurements m (for construction of the output vector y)
4. Selection of a structure interconnecting outputs y and inputs u (structure of the controller-matrix C).

This structure is also called control configuration or pairing.

The relation between controlled outputs, outputs and measurements can be described as follows. If every controlled output z is measured in a control system, then there is no difference between controlled outputs z, outputs y and measurements m. On the other hand,

if a controlled output z can not be measured, the output y might be chosen as y = f (m), with appropriately chosen measurements m, such that yi can be used instead of z. In this case the setpoint for z and y might be different.

2.4. Decentralized control of the regulatory layer

The regulatory control layer is usually highly decentralized, meaning that it mostly contains single-input/single-output elements and only a few small multivariable controllers. In comparison to full multivariable control of the whole plant, highly decentralized control has the following advantages [16] [17].

1. *Modelling.* For full multivariable control, very detailed and exact models of the plant are required. These models are usually very complex, not easy to derive and thus very expensive. In the same way, the costs connected to controller design, maintenance and implementation of the control system are much lower for the case of the decentralized control system.

2. *Uncertainty and failures.* Multivariable controllers are much more faultprone to uncertainty or failures in measurements and actuators, than decentralized controllers. Because of this inherent robustness, they can also deal better with changes in operating conditions of a plant.

3. *Startup and shutdown.* A decentrally controlled regulatory layer allows startup of the whole plant one loop at a time. This startup phase would be much more complicated for full multivariable control. The same holds, of course, for the shutdown of a plant.

4. *Ease of retuning.* In a decentralized control structure the individual control blocks can usually be retuned more or less independently, e.g., in response to changes in operating conditions.

5. Understanding. The decentralized regulatory control layer is easier to understand and to handle by operators. It is important to understand that, in the case of decentralized control, the control configuration is crucial for the achievable performance. This means that the decision about which output to control using which input is by far more important than the selection of a certain controller type for the resulting loop and a good tuning of this controller.

2.5. Control systems design

Designing a control system is a creative process involving a number of choices and decisions. These choices depend on the properties of the system that is to be controlled and on the requirements that are to be satisfied by the controlled system. The decisions imply compromises between conflicting requirements. The design of a control system involves the following steps:

1. Characterize the system boundary, that is, specify the scope of the control problem and of the system to be controlled.

2. Establish the type and the placement of actuators in the system, and thus specify the inputs that control the system.

3. Formulate a model for the dynamic behaviour of the system, possibly including a description of its uncertainty.

4. Decide on the type and the placement of sensors in the system, and thus specify the variables that are available for feedforward or feedback.

5. Formulate a model for the disturbances and noise signals that affect the system.

6. Specify or choose the class of command signals that are to be followed by certain outputs.

7. Decide upon the functional structure and the character of the controller, also in dependence on its technical implementation.

8. Specify the desirable or required properties and qualities of the control system.

In several of these steps it is crucial to derive useful mathematical models of systems, signals and performance requirements. The models of systems we consider are in general linear and time-invariant. Sometimes they are the result of physical modelling obtained by application of first principles and basic laws; on other occasions they follow from experimental or empirical modelling involving experimentation on a real plant or process, data gathering, and fitting models using methods for system identification [13].

The functional specifications for control systems depend on the application. We distinguish different types of control systems:

- Regulator systems

  The primary function of a regulator system is to keep a designated output within tolerances at a predetermined value despite the effects of load changes and other disturbances.

- Servo or positioning systems

  In a servo system or positioning control system the system is designed to change the value of an output as commanded by a reference input signal, and in addition is required to act as a regulator system.

- Tracking systems

  In this case the reference signal is not predetermined but presents itself as a measured or observed signal to be tracked by an output.

For single-input single-output feedback systems realizing the most important design targets may be viewed as a process of loop shaping of a one-degree-of-freedom feedback loop. The targets include:

- closed-loop stability
- disturbance attenuation
- stability robustness within the limitations set by o plant capacity or corruption by measurement noise.

Further design targets, which may require a two-degree-of-freedom configuration, are:

- satisfactory closed-loop response
- robustness of the closed-loop response.

2.6. Definition of the problem

The central problem in control is to find a technically feasible way to act on a given process so that the process adheres as closely as possible to some desired behaviour. Furthermore, this approximate behaviour should be achieved in the face of uncertainty of the process and in the presence of uncontrollable external disturbances acting on the process.

The above definition introduces several ideas:

*** Desired behaviour.** This needs to be specified as part of the design problem.

*** Feasibility**. This means that the solution must satisfy various constraints, which can be of technical, environmental, economic, or other nature.

*** Uncertainty.** The available knowledge about a system will usually be limited and of limited accuracy.

*** Action.** The solution requires that action be somehow applied to the process, typically via one or more manipulated variables which command the actuators.

*** Disturbances.** The process to be controlled will typically have inputs other than those that are manipulated by the controller. These other inputs are called disturbances.

*** Approximate behaviour.** A feasible solution will rarely be perfect. There will invariably be a degree of approximation in achieving the specified goal.

*** Measurements.** These are crucial to let the controller know what the system is actually doing and how the unavoidable disturbances are affecting it.

2.7. Physics and controls

Examples of dynamic systems with automatic controllers abound. dvanced process controllers are operating in virtually every industrial domain; micro-controllers pervade an immense array of household and entertainment electronics; thermostats regulate temperatures in domestic- to industrial-sized ovens.

To particularize the principal goal of control engineering within this team effort, it is helpful to distinguish between a system's tangible realization and its behaviour. The aircraft's physical realization, for example, includes fuselage, wings, and ailerons. Its behaviour, on the other hand, refers to the aircraft's dynamic response to a change in throttle, aileron, or flap position.

To control such a system automatically, one needs to interface the system to a controller, which will also have a physical realization and behaviour. Depending on the application, the controller could be realized in a chip, analogue electronics, a PLC, or a computer. There also needs to be a channel by which the controller and system can interact via sensors and actuators: sensors to report the state of the system, actuators as a means for the controller to act on the system.

With this process and control infrastructure in place, the key remaining question pertains to the controller behaviour. In the aircraft application, for example, if the controller (here called an autopilot) detects a deviation in speed, height or heading via the sensors, just how should it command throttle and ailerons to get back on target?

The fundamental goal of controls is to find technically, environmentally, and economically feasible ways of acting on systems to control their outputs to desired values, thus ensuring a desired level of performance. Finding a good solution to this question frequently requires an involvement in process design, actuator and sensor selection, mathematical analysis, and modelling.

Such a cyclically dependent interaction between system behaviours is called feedback. This behaviour-altering effect of feedback is a key mechanism that is exploited deliberately to achieve the objective of acting on a system to ensure that the desired performance specifications are achieved.

2.8. Model based design

Model-based design emerged as a means of addressing the difficulties and complexities inherent in control systems designs. Developers recognized that software design needs to start before physical prototypes and systems are available. Traditional design processes resulted in the discovery of design and requirements errors late in the design cycle resulting in expensive delays and missed windows of opportunities. In traditional design processes, design information is communicated and managed as text based documentation. Frequently this documentation is difficult to comprehend. Code is created manually from specification and

requirements documents that are time consuming and error prone. There is little tracking to ensure that changes are correctly implemented.

Model-based design, used to its fullest, provides a single design environment that enables developers to use a single model of their entire system for data analysis, model visualization, testing and validation, and ultimately product deployment, with or without automatic code generation. At a minimum model-based design can be used as a specification that contains greater detail than text-based specifications. In real-time applications, it enables developers to evaluate multiple options, predict systems performance, test systems functionality by imposing I/O conditions that might be operationally expected (before product deployment), and test designs.

Once the model is built and completely tested, accurate real-time software for the production embedded design is automatically generated, thereby saving time and reducing costs compared to traditional manual coding. Model-based design with automatic code generation can also be used in rapid prototyping, enabling subsystem designs to be tested and optimized. Although it is very important in highly complex design applications (e.g., guidance systems, engine controls, autopilots, anti-lock braking systems) that otherwise might be difficult to realize without it, model-based design can be used effectively and economically for less complex designs.

Furthermore, model-based design creates a structure for software reuse that permits established designs to be effectively and reliably upgraded in a more simplistic and cost effective manner.

Model-based design works in the following manner [35]:

• The entire system model is visualized via block diagrams and state charts to describe knowledge and implementation details

• Design options can be evaluated and systems performance predicted via simulation of the system model

• Algorithm and behavioral models are optimized and refined yielding a fully tested specification

• Production quality software is automatically created for real-time testing and deployment from the fully tested specification.

Model-based design saves money by cutting design time and providing final designs that more closely approximate pre-design expectations for performance, systems functionality, and features and schedule.

It provides:

• Faster design iterations that produce desired performance, functionality and capabilities.

• Design cycles that are more predictable and result in faster product shipments

• Reduction in design, development and implementation costs.

2.9. System modeling

Any description of a system could be considered to be a model of that system. Although the ability to encapsulate dynamic information is important, some analysis and design techniques require only steady-state information. Models allow the effects of time and space to be scaled, extraction of properties and hence simplification, to retain only those details relevant to the problem.

In terms of control requirements, the model must contain information that enable prediction of the consequences of changing process operating conditions. Within this context, a model could either be a mathematical or statistical description of specific aspects of the process. It can also be in the form of qualitative descriptions of process behaviour. Depending on the task, different model types will be employed [40].

2.9.1. Mechanistic models

If much is known about the process and its characteristics are well defined, then a set of differential equations can be used to describe its dynamic behaviour. This is known as 'mechanistic' model development. The mechanistic model is usually derived from the physics and chemistry governing the process. Depending on the system, the structure of the final model may either be a lumped parameter or a distributed parameter representation. Lumped parameter models are described by ordinary differential equations (ODEs) while distributed parameter systems representations require the use of partial differential equations (PDEs). ODEs are used to describe behaviour in one dimension, normally time, e.g. the level of liquid in a tank. PDE models arise due to dependence also on spatial locations, e.g. the temperature profile of liquid in a tank that is not well mixed.

Obviously, a distributed parameter model is more complex and hence harder to develop. More importantly, the solution of PDEs is also less straightforward. Nevertheless, a distributed model can be approximated by a series of ODEs given simplifying assumptions. Both lumped and distributed parameter models can be further classified into linear or nonlinear descriptions. Usually nonlinear, the differential equations are often linearized to enable tractable analysis.

In many cases, typically due to financial and time constraints, mechanistic model development may not be practically feasible. This is particularly true when knowledge about the process is initially vague or if the process is so complex that the resulting equations cannot be solved. Under such circumstances, empirical or 'black-box' models may be built using data collected from the plant.

2.9.2. Black Box models

Black box models simply describe the functional relationships between system inputs and system outputs. They are, by implication, lumped parameter models. The parameters of these functions do not have any physical significance in terms of equivalence to process parameters such as heat or mass transfer coefficients, reaction kinetics, etc. This is the disadvantage of black box models compared to mechanistic models. However, if the aim is to merely represent faithfully some trends in process behaviour, then the black box modelling approach is just as effective. Moreover, the cost of modelling is orders of magnitude smaller than that associated with the development of mechanistic models.

Black box models can be further classified into linear and nonlinear forms. In the linear category, transfer function and time series models predominate. With sampled data systems, this delineation is, in a sense, arbitrary. The only distinguishing factor is that in time-series models, variables are treated as random variables. In the absence of random effects, the transfer function and time-series models are equivalent. Given the relevant data, a variety of techniques may be used to identify the parameters of linear black box models. The most common techniques used, though, are least-squares based algorithms.

2.10. Model representation

The processes and their associated variables can be described pictorially as shown in Figure 2-4. The main block represents the process while the arrows indicate the inputs, the outputs and the disturbances of the process, respectively.

A mathematical model is a convenient and useful tool for a control system designer and should conform to the figure above (i.e. be such that given the values of the inputs it provides directly the values of the outputs). In particular, the model should have the following general form for every output:

Output = f(input variables).

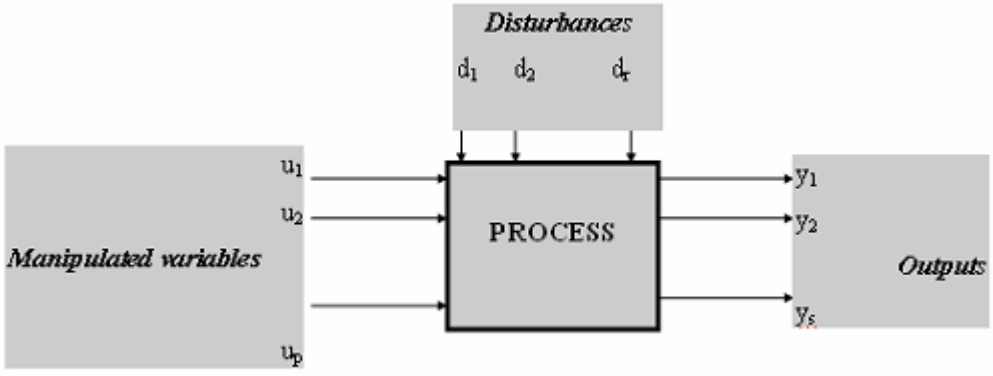


Figure 2-4 Process and its associated inputs/outputs

According to Figure 2-16, the relationship from above can be written as:

Eq. (2.1)

$$y_i = f(u_1, u_2, ...u_p, d_1, d_2, ..., d_r) i = 1,2,3,..,s$$

Such a model, describing directly the relationship between the input and the output variables of a process, is called an input-output model. It is a very convenient form since it represents directly the cause-effect relationship in systems.

The input-output relation of a linear time-invariant system with continuous-data input is often described by a differential equation or a system of simultaneous differential equations. Let us consider that the input-output relation of a linear time-invariant system is described by the following n_{th} order differential equation with constant coefficients:

Eq. (2.2)

$$\frac{d^n y(t)}{dt^n} + a_1 \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_n y(t) = b_m u(t) + \cdots + b_1 \frac{d^{m-1} u(t)}{dt^{m-1}} + b_0 \frac{d^m u(t)}{dt^m}$$

where the coefficients a_1, a_2, ..., a_n, and respectively b_1,..., b_m are real constants, and $n \geq m$. This last condition originates from the feasibility of the system. Once the input u(t) for every $t \geq t_0$ (t=0 usually) is known, respectively the initial conditions of the problem (i.e. the value of the unknown function), respectively the values of its derivatives:

Eq. (2.3)

$$y(t_0) = y_0, y^{(1)}(t_0) = y_1, y^{(2)}(t_0) = y_2, \cdots, y^{(n-1)}(t_0) = y_{n-1}$$

are given than the output response y(t) for t>=0 is determined solving this equation. Control systems are often modelled as linear, time-invariant systems. A convenient analysis tool is the Laplace transform. It is used in favor of the Fourier transform, since it is capable of handling one-sided initial value problems (transients) and signals with exponential growth.

The (one-sided) Laplace transform of a signal x is given by

Eq. (2.4)

$$X(S) = \int_0^{\infty} (x(t) \cdot e^{-st}) dt$$

where s is the complex Laplace-variable, s = s+ iw. The Laplace representation of a system (the transfer function H(s)) is obtained by Laplace transforming its impulse response h.

Hence, for a system H with input x and output y the transfer function is

Eq. (2.5)

$$H(S) = \frac{Y(S)}{X(S)}$$

with X(s) the Laplace transform of the input signal, and Y(s) the Laplace transform of the output signal.

2.11. Feedback systems

A simple form of feedback consists of two dynamical systems connected in a closed loop which creates an interaction between the systems. Simple causal reasoning about such a system is difficult because the first system influences the second and the second system influences the first, leading to a circular argument. This makes reasoning based on cause and effect difficult and it is necessary to analyze the system as a whole. A consequence of this is that the behaviour of a feedback system is often counterintuitive.

Feedback has many advantages: it is possible to create linear behaviour out of nonlinear components, and a system can be made very resilient towards external influences or disturbances. The major disadvantage is that instability may arise.

Feedback techniques have always had a central role in scientific instruments. An early example is the development of the mass spectrometer. In a paper from 1935 by Nier it is observed that the deflection of the ions depend on both the magnetic and the electric fields. Instead of keeping both fields constant Nier let the magnetic field fluctuate and the electric field was controlled to keep the ratio of the fields constant. The feedback was implemented using vacuum tube amplifiers. The scheme was crucial for the development of mass spectroscopy.

Another example is the work by the Dutch Engineer Van Der Meer. He invented a clever way to use feedback to maintain a high density and good quality of the beam of a particle accelerator. The scheme, called stochastic cooling, was awarded the Nobel prize in Physics in 1984, along with Carlo Rubbia. The method was essential for the successful experiments in CERN when the existence of the particles W and Z was first demonstrated. Another use of feedback called repetitive control was developed by Nakano for particle accelerators. The key idea was to obtain very precise control by exploiting the fact the particles move in circular orbits.

The atomic force microscope is a more recent example. The key idea is to move a narrow tip on a cantilever beam across the surface and to register the forces on the tip. Such systems rely on feedback systems for precise motion and precise measurement of the forces. Feedback is a powerful idea, which is used extensively in natural and technical systems. The principle of feedback is very simple: base correcting actions on the difference between desired and actual performance. In engineering feedback has been rediscovered and patented many times in many different contexts.

The benefits of feedback can often be obtained using simple forms of feedback such as on/off control and PID (Proportional Integral Derivative) control. Feedback can reduce the effects of disturbances and process variations, it can create well defined relations between variables, it makes it possible to modify the properties of a system, e.g. stabilize an unstable system. The discussion is based on block diagrams and simple static mathematical models. The major drawback is that feedback can create instability.

The function of a feedback control system is, therefore, to ensure that the closed loop system has desirable dynamic and steady-state response characteristics. Ideally, it should satisfy the following performance criteria:

1. The closed loop must be stable.

2. The effect of disturbances are minimized, providing good disturbance rejection.

3. Rapid, smooth responses to set-point changes are obtained, that is, good set-point tracking.

4. Steady-state error (offset) is eliminated.

5. Excessive control action is avoided.

6. The control system is robust, that is, is insensitive to changes in process conditions and to inaccuracies in the process model.

2.12. A generic Control System with Error Feedback

Feedback control means measuring the controlled variable (an output), comparing that measurement to the set point (desired value), and acting in response the error (difference between set point and controlled variable) by adjusting the manipulated variable (an input). The control algorithm is the set of calculations and decisions that lies between the error and the directions given to the final control element.

A generic representation of a system is shown in Figure 2-5 [41]. The system has two blocks. One block P represents the process and the other C represents the controller. Notice negative sign of the feedback. The signal r is the reference signal which represents the desired behaviour of the process variable x. There are two types of disturbances, labelled d and n: the disturbance labelled d is called a load disturbance and the disturbance labelled n is called measurement noise. Load disturbances drive the system away from its desired behaviour.

In Figure 2-5 it is assumed that there is only one disturbance that enters at the system input. This is called an input disturbance. In practice there may be many different disturbances that enter the system in many different ways. Measurement noise corrupts the information about the process variable obtained from the measurements.

It is also assumed that the measured signal y is the sum of the process variable x and measurement noise. In practice the measurement noise may appear in many other ways. The system is said to have error feedback, because the control actions are based on the error which is the difference between the reference r and the output y. In some cases there is no explicit information about the reference signal because the only information available is the error signal.

Attenuation of load disturbances is often a primary goal for control. This is particularly the case when controlling processes that run in steady state. Load disturbances are typically dominated by low frequencies. Consider for example the cruise control system for a car, where the disturbances are the gravity forces caused by changes of the slope of the road. These disturbances vary slowly because the slope changes slowly when one drives along a road.

Step signals or ramp signals are commonly used as prototypes for load disturbances. Measurement noise corrupts the information about the process variable that the sensors deliver.

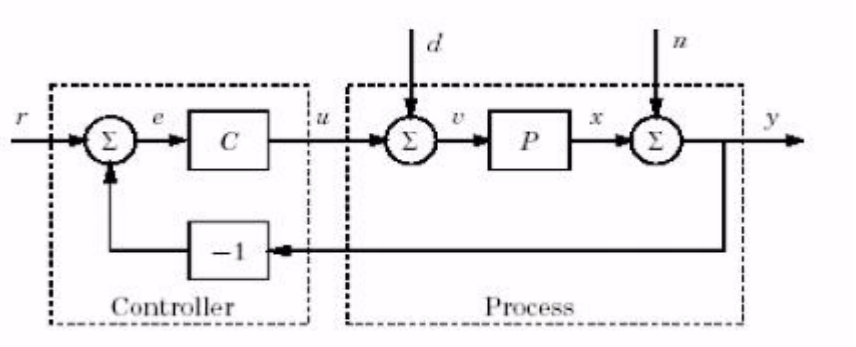


Figure 2-5 Generic control system with error feedback

Measurement noise typically has high frequencies. The average value of the noise is typically zero. If this was not the case the sensor would give very misleading information about the process and it would not be possible to control it well. There may also be dynamics in the sensor. Several sensors are often used. A common situation is that very accurate values may be obtained with sensors with slow dynamics and that rapid but less accurate information can be obtained from other sensors.

2.13. System identification

System identification is the experimental approach to process modelling.

It includes the following steps [44]:

• Experiment design: its purpose is to obtain good experimental data, and it includes the choice of the measured variables and of the character of the input signals.

• Selection of model structure: A suitable model structure is chosen using prior knowledge and trial and error.

• Choice of the criterion to fit: A suitable cost function is chosen, which reflects how well the model fits the experimental data.

• Parameter estimation: An optimization problem is solved to obtain the numerical values of the model parameters.

• Model validation: The model is tested in order to reveal any inadequacies.

Figure 2-6 shows an algorithm for modeling and system identification. Notice that the order of the blocks in the algorithm does not only describe the chronological order the tasks are performed, but also how they influence each other. For example, a certain model structure can be proposed by the physical model, the amount of data limits the model complexity etc.

It is sometimes possible to derive a model directly from physical laws. This model will most often, however, contain unknown parameters to be estimated.

If some parameters are known and some are unknown, it is sometimes (but not always) possible to perform an estimation using the values of the known parameters [50].

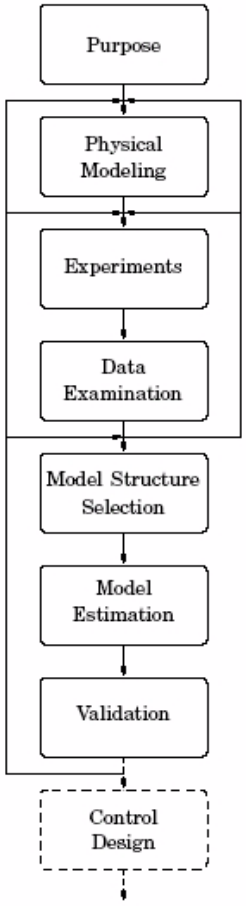


Figure 2-6 Control Systems Design Procedure

We commonly use only the structure of the model derived from the physical model. This structure can be in terms of model order, known pole locations (an integrator or dominating resonance), static nonlinearities etc. If there are no knowledge about the considered system, we use the term black-box identification. It is called grey-box identification, if some part of the system is known.

The goal of system identification is to utilize input/output experimental data to determine a system's model. For example, one may apply to the system a specific discrete-time input u(k), measure the observed discrete-time output y(k) an try to determine the system's transfer function (Figure 2-7).

The field of system identification starts from experimental test data, and develops a mathematical model of the observed physical system behavior. Typically one excites the system with some rich input signal such as a random white noise, measures the response at sample times, and from this data one creates a differential equation or a difference equation that can predict the system response to arbitrary inputs.

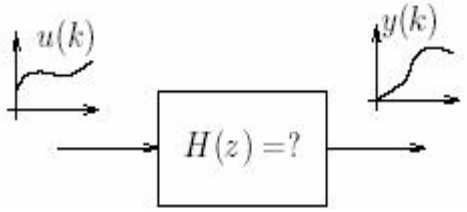


Figure 2-7 System Identification

An important advantage of this direct experimental approach to getting a model is that it does not require one to use sophisticated and often time-consuming modeling based on physical laws, and does not suffer from simplifications that one usually has to make in physical law based modeling [63].

Algorithms are developed for generating several different types of mathematical input-output models from data. These include: a single-input, single-output (SISO) autoregressive exogenous variable model (ARX), and a multiple-input, multiple-output (MIMO) version using a multivariable matrix version of an ARX model; both a SISO and a MIMO state variable model in observable canonical form; and a MIMO state variable model in block diagonal form, appropriate for identifying vibration modes directly. These models allow one to model any general linear time-invariant system, although one may need to use a dimension higher than that of a minimal order realization.

In parametric identification one attempts to determine a finite number of unknown parameters that characterize the model of the system. The following is a typical problem in this class: Parametric transfer function identification. Determine the coefficients α_i and β_i of the system's rational transfer function:

Eq. (2.6)

$$H(z) = \frac{\alpha_m z^m + \alpha_{m-1} z^{m-1} + \ldots + \alpha_1 m z + \alpha_0}{z^n + \beta_{n-1} z^{n-1} + \ldots + \beta_1 z + \beta_0}$$

The number of poles n and the number of zeros m are assumed known.

This can be done using the methods of least-squares, which we introduce next, starting from a simple line fitting problem and eventually build up to the Problem.

Given any of these model types, we seek to fit the input-output data so that the parameters of the model minimize the sum of the squares of the differences between the observed output in the data and the output produced by the model. This objective function is a nonlinear function of the model parameters, and the number of parameters can be high.

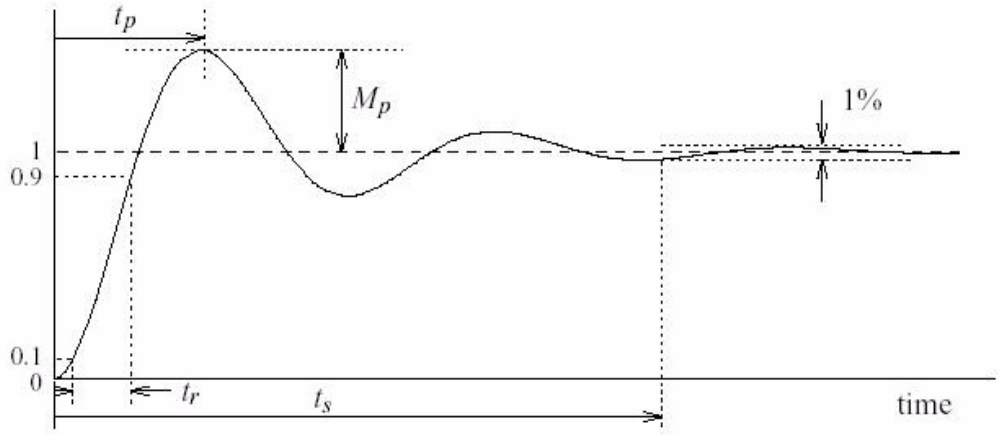
Although this objective function seems quite natural, and in fact is arguably the most logical function to minimize in fitting data, it is not the objective function used in most identification approaches. Perhaps the most common approach is to minimize the Euclidean norm of the equation error. This is attractive because it produces a linear set of equations to be solved, and avoids the minimization of a nonlinear function in many variables that must be addressed here.

We will see that the distinction between the usual minimization of the equation error norm versus minimizing the output error norm as done here, has practical significance.

The problem posed above is a nonlinear least squares problem with potentially a large number of parameters. Using routine methods that do not require derivative information, i.e., an explicit expression for the derivative of the objective function with respect to the parameters, was seen to be totally impractical in terms of run times. The result is the same when only first derivative information is used. Hence, we make use of sequential quadratic programming to construct the algorithms. An automatic step-size adjustment scheme is used that ensures that the updates not only decrease the objective function, but also maintain stability of the model. A major part of the effort in creating the algorithms is the generation of the formulas used to efficiently evaluate the first and second partial derivatives.

2.14. Control performances

The system response is often specified by certain parameters. We distinguish transient response specification parameters, which describe the response to a transient input (e.g., a step), and the steady state response parameters, which describe the system output after the transient has died out [49].



t_r	rise time	time to reach the vicinity of a new set point
t_s	settling time	time after which transient effects have damped out
M_p	overshoot	maximum deviation of system output from its final value
t_p	peak time	time to reach the maximum overshoot point

Figure 2-8 Response to a step input

For Single Input-Single Output (SISO) systems we have the following partial list of typical classical performance specifications. Consider the feedback loop of Fig. 21. These are the basic requirements for a well designed control system:

1. The transient response is sufficiently fast.

2. The transient response shows satisfactory damping.

3. The system is sufficiently insensitive to external disturbances and variations of internal parameters.

These basic requirements may be further specified in terms of both a number of frequency-domain specifications and certain time-domain specifications. Figures 2-20 and 2-21 illustrate several important frequency-domain quantities:

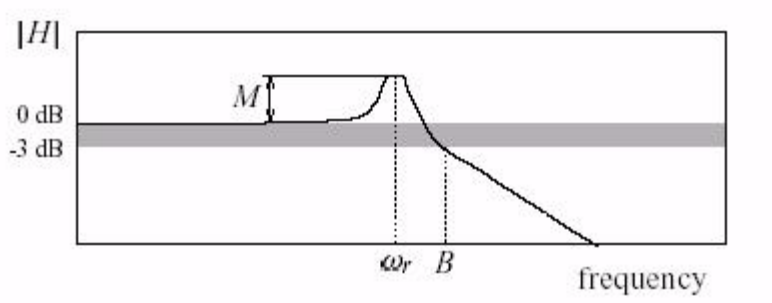


Figure 2-9 Gain and Phase Margin

Gain margin. The gain margin measures relative stability. It is defined as the reciprocal of the magnitude of the loop frequency response L, evaluated at the frequency ω_π at which the phase angle is -180 degrees. The frequency ω_π is called the phase crossover frequency.

Phase margin. The phase margin also measures relative stability. It is defined as 180° plus the phase angle Φ_1 of the loop frequency response L at the frequency ω_1 where the loop gain is unity. The frequency ω_1 is called the gain crossover frequency.

Bandwidth. The bandwidth B measures the speed of response in frequency-domain terms. It is defined as the range of frequencies over which the closed-loop frequency response H has a magnitude that is at least within a factor =0.707 (3 dB) of its value at zero frequency.

Resonance peak. Relative stability may also be measured in terms of the peak value M of the magnitude of the closed-loop frequency response H (in dB), occurring at the resonance frequency ω_r.

Rise time Tr. The rise time expresses the "sharpness" of the leading edge of the response. Various definitions exist. One defines TR as the time needed to rise from 10% to 90% of the final value.

Percentage overshoot MP. This quantity expresses the maximum difference (in % of the steady-state value) between the transient and the steady-state response to a step input.

Settling time Ts. The settling time is often defined as time required for the response to a step input to reach and remain within a specified percentage (typically 2 or 5%) of its final value.

2.15. Stability of control systems

Some control systems are stable and some not. There are many definitions of stability [51]:

- a system is stable if we move it from its original state it returns into that state (asymptotically stable systems), or to the close neighbourhood of it.
- a system is stable if it responds with an amplitude limited output to a limited input signal.

The possibility of unstable behaviour is a consequence of the feedback control structure. Real systems contain time lags and dead time, so delayed information is used to make decisions in control systems. The most important thing during the design of a control system is that the system must be stable.

As a general statement we can say that all systems are stable if the solution of the characteristic equation of the differential equation that describes the system has roots with negative real part, or in case of discrete systems the roots are inside of the unit circle. A continuous control system is stable if in the Laplace operator domain its closed loop transfer function has only left-hand poles (this means that the real part of these poles is negative), and a system is unstable if it has right hand side poles (the real part of these poles is positive). If a system has poles that have only imaginary component the system is on the verge of stability, which means that there are oscillations with constant amplitude in the output of the system.

We can also define stability as “bounded output for a bounded input" (also known as “BIBO”). It means that:

- a ramp disturbance is not fair – even stable systems can get into trouble if the input keeps rising;
- an impulse disturbance is fair – although it is briefly infinite, it soon passes. • a stable system should also handle a step change in input, ultimately coming to some new steady state. e must be realistic, however. a system is so sensitive that a small input step leads to an unacceptably high, though steady, output, we might declare it unstable for practical purposes)
- it should also handle a sine input; here the result is in general not steady state, because the output may oscillate. Thus we distinguish between 'steady state' and 'long-term stability'.)

Stability depends on:

- The system – certainly; we will discuss the characteristics of the system that determine stability.
- The type of disturbance – yes, as discussed above. It is possible that a system is stable to some bounded inputs, but not others. For example, the first-order integrator is stable to sine disturbances but not to steps. Thus, we declare such a system unstable
- The magnitude of disturbance – maybe.We represent real systems by linear models. While the magnitude of the disturbance does not influence stability of the linear system, a sufficiently large disturbance to the real system may move it to a regime of operation that the linear approximation does not describe.

The key is the poles of the linear characteristic equation. If any pole has a zero or positive real part, the system will be unstable to bounded disturbances. This happens because of the structure of the linear, constant-coefficient, ordinary differential equations that we are using – the solutions are exponential terms.

2.16. Controller design

In the classical control engineering era the design of feedback compensation to a great extent relied on trial-and-error procedures. Experience and engineering sense were as important as a thorough theoretical understanding of the tools that were employed.

Here we consider the basic goals that may be pursued from a classical point of view. In the classical view the following series of steps leads to a successful control system design [42]:

- Determine the plant transfer function P based on a (linearized) model of the plant.
- Investigate the shape of the frequency response, to understand the properties of the system fully.
- Consider the desired steady-state error properties of the system. Choose a compensator structure — for instance by introducing integrating action or lag compensation — that provides the required steady-state error characteristics of the compensated system.
- Plot the Bode, Nyquist or Nichols diagram of the loop frequency response of the compensated system. Adjust the gain to obtain a desired degree of stability of the system. M- and N-circles are useful tools. The gain and phase margins are measures for the success of the design.
- If the specifications are not met then determine the adjustment of the loop gain frequency response function that is required. Use lag, lead, lag-lead or other compensation to realize the necessary modification of the loop frequency response function. The Bode gain-phase relation sets the limits.

A general system, including the controller, can be expressed as follows:

where:

R(s) is the desired output value, or setpoint

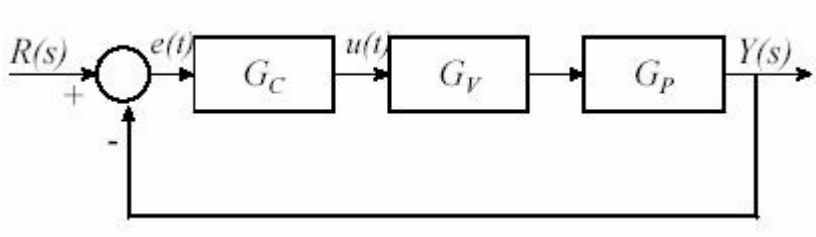


Figure 2-10 Control System with its Controller

Y(s) is the actual output value

Gp represents a process Gc represents a controller

Gv represents a control element, such as a valve or the ailerons on an aircraft.

The Closed Loop Transfer Function (CLTF) of this system is:

Eq. (2.7)

$$G(s) = \frac{Y(s)}{R(s)} = \frac{G_c G_v G_p}{1 + G_c G_v G_p}$$

The controller, Gc, must be designed so that given a setpoint it is capable of adjusting the operating conditions of the process to ensure that the output deviates within an allowable range of this value.

Whatever the nature of the process, there will be some desired point of operation, and some difference from that, or error.

We desire to minimize that error through feedback control, and we conceive of controller schemes to do that. First, it seems reasonable that the controller response be proportional to the magnitude of the error. Then we would like the controller to take stronger action should error persist in time. Finally, we would like the controller to act quickly at the onset of error in an attempt to get ahead of it.

Although the optimal form that a controller can take is very much dependent upon the characteristics of the system, one particular form of controller, known as a 3-term (or PID) controller, has been devised to be suitable for many general types of system. The performance of such controllers, although not optimal, is often adequate.

2.17. PID controllers

In industrial environments, the PID-controller is still a favourite feedback controller: practical experience confirms that the PID-controller has a structure which is sufficiently versatile to cope in an excellent way with a multitude of process control problems. Other advantages of the PID-law are the low number of design parameters and the fact that the controller parameters can easily be related to performance measures [27].

The PID-controller transforms the error signal e(t) (i.e. the difference between the actual system output and the desired system output) into the control signal u(t). Since it is a linear controller, it can be studied in the time domain as well as in the frequency domain.

We can realize these notions in the PID (proportional-integral-derivative) control algorithm. The PID algorithm is simply a plausible way to do a controller response. The time domain representation of the controller output CO(t) is:

Eq. (2.8)

$$CO(t) = \left(E(t) + \frac{1}{T}\int_0^t E(t))dt + T_d \cdot \frac{dE}{dt} \right) \cdot K_c$$

The definition is written in terms of deviation variables. The error E is the difference between set point and controlled variable, both written as deviation variables.

$$E(t) = SP(t) - CV(t) \qquad \text{Eq. (2.9)}$$

$$(E(t) - Es) = (SP(t) - SPs) - (CV(t) - CVs) \qquad \text{Eq. (2.10)}$$

where SP is the set point (that is, a value imposed by the user) and CV is the controlled varibale (the varibale that is actually afected by the effect of the control).

Basically, the controller consists of three terms, each with a typical function. The first term is proportional to the error signal, the second term is proportional to its integral and the third term is proportional to its derivative. However, to be practical, this algorithm needs to be modified, especially for the implementation of the derivative action. Pure derivation is a practical nonsense, as it has unbounded amplification.

The transfer function of the PID-controller is then given as:

Eq. (2.11)

$$C(s) = K_p(1 + \frac{1}{sT_i} + \frac{sT_d}{sT_c + 1})$$

Kp is the proportional gain, or simply gain. Conventionally the gain is applied across the other two modes, as well, as indicated in the definition.

Ti is the integral time. Its reciprocal is often called the reset rate. This reset rate has dimensions of "repeats/time". This is because at constant error input ΔE, the controller output will increase by $K_C \cdot \Delta E$ in each time increment Ti. Deactivate integral mode by setting Ti to infinity. Short Ti, or high reset rate, represents more aggressive control response.

Td is the derivative time. It is also called preact. Deactivate derivative mode by setting Td to zero.

Large Td represents more aggressive control response. Marlin[46] suggests that the derivative mode be applied only to the controlled variable and not the error term as shown in

the equation. By this, set point changes would not lead to sudden manipulated variable changes, driven by the derivative mode.

A possible PID-scheme is given in Fig. 2-23.

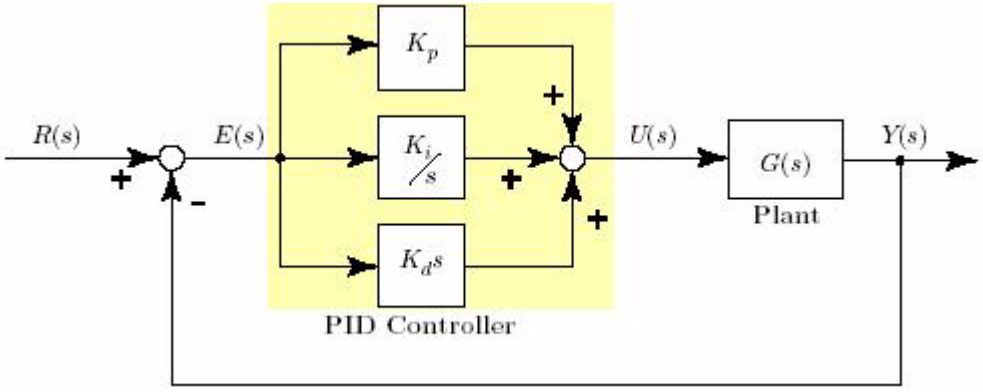


Figure 2-11 A PID Controller Implementation

The open loop time response of a PID controller to a step error input, is shown in Fig. 2-11. The P-action is fast since it reacts immediately to a change in process value. A constant error results in a constant control action due to the P-term. The I-action causes the control action to keep on growing as long as the error exists. Therefore the error will always decrease using an I-controller. Compared to the P-controller, the I-controller is rather slow. The D-action reacts very fast when an error appears. It is only active at the moment when the error changes.

Obviously, PID-controllers will always operate in a closed loop, this is: connected to the system. Basically 3 parameters (excluding tc) have to be chosen:

- The gain Kp
- The integration time or reset time Ti
- The derivation time or lead time Td.

We will now analyze in more detail the effect of the different options.

2.17.1. Proportional Controller

This is the simplest form of feedback control. The control signal, u(t), is made to be linearly proportional to the error in the measured output: u(t) = KP * e(t). The transfer function for the controller is given by: GC = KP.

A large error results in a large output from the controller, and a small error result in a small output from the controller.

The graph below shows how the output of the controller varies with error:

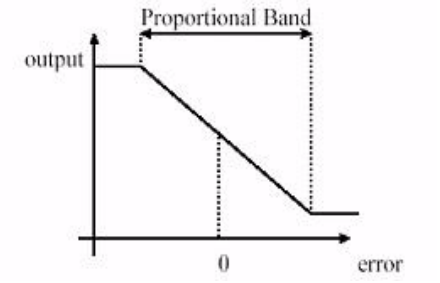


Figure 2-12 Proportional Controller

Note that the graph does not continue as a straight line forever; this is because in real life, the output of the controller tends to be bounded. For example the signal to a valve cannot be outside the range 0% open to 100% open. Similarly, if the speed of a motor is being controlled, a constraint on the motor may be that the maximum supply voltage is 240V.

The range over which the controller responds with a constant gain is known as the proportional band, and $K_P = \frac{100}{ProportionalBand}$

2.17.2. Proportional control of First-order systems

Let us now consider the response of a first-order system under proportional control:

Eq. (2.12)

$$G_p = \frac{K}{Ts + 1}$$

And G_V is assumed to be 1. With proportional control the Closed Loop Transfer Function (CLTF) is equal to:

Eq. (2.13)

$$\frac{Y(s)}{R(s)} = \frac{K \cdot K_P}{Ts + 1 + KK_P}$$

Which, rearranging, gives:

Eq. (2.14)

$$\frac{Y(s)}{K(s)} = \frac{\frac{KK_P}{1 + KK_P}}{\left(\frac{Ts}{1 + KK_P}\right) + 1}$$

The CLTF shows that as the gain of the controller is increased, the overall gain of the system approaches unity and the time constant gets smaller and smaller.

The addition of proportional control also affects the characteristic equation, which implies that the response of the system will also change. The time constant is reduced as the gain of the controller increases and therefore the system response will be faster.

The purpose of the controller is to maintain the system output at the setpoint value. To determine whether or not a proportional controller achieves this it is necessary to determine the steady state value of the output (i.e. the output as time) after a change in setpoint. This calculation is simplified by using the Final Value Theorem of Laplace transforms, which states:

Eq. (2.15)

$$\lim_{t \to \infty} e(t) = \lim_{s \to 0} sE(s)$$

Therefore,

$$E(s) = R(s) - Y(s) = R(s) - E(s)G_C G_P$$ Eq. (2.16)

Eq. (2.17)

$$E(s) = \frac{R(s)}{1 + G_C G_P}$$

Eq. (2.18)

$$\lim_{t \to \infty} e(t) = \lim_{s \to 0} \frac{1}{1 + K_P \frac{K}{1 + T_S}} \frac{1}{(s)} = \frac{1}{1 + KK_P}$$

This equation states that there will always be an error, unless the value of Kp is equal to ∞.

2.17.3. Proportional Control of Second Order Systems

Consider the system defined by:

Eq. (2.19)

$$G_p = \frac{K}{s^2 + as + b}$$

Assuming unity feedback with proportional control, the CLTF is given by:

Eq. (2.20)

$$\frac{Y(s)}{R(s)} = \frac{KK_p}{s^2 + as + (b + KK_p)}$$

This shows that when a proportional controller is used with a second order system, the resulting control system is still second order. The dynamics of the system are however altered by the controller. The natural frequency and damping ratio of the control system are as follows:

Eq. (2.21)

$$\omega_n = \sqrt{b + KK_p}, \zeta = \frac{a}{2\sqrt{b + KK_p}}$$

This implies that as the value of the proportional gain is increased the natural frequency will increase and the damping ratio will reduce. The result will be greater oscillation and the frequency of this oscillation will be increased.

The steady state error of the system is determined as follows:

Eq. (2.22)

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s), \text{ where } E(s) = R(s) - Y(s)$$

$$E(s) = R(s) - Y(s) = R(s)\left[1 - \frac{KK_p}{KK_p + s^2 + as + b}\right]$$

$$\lim_{s\to 0} sE(s) = s\frac{1}{s}\left[\frac{s^2 + as + b}{KK_p + s^2 + as + b}\right] = \frac{b}{KK_p + b}$$

Consequently, just like the case of a first order system, there will always be a steady state error unless Kp is equal to ∞. The graph below shows the response of the second order system:

Eq. (2.23)

$$G(s) = \frac{1}{s^2 + 2s + 1}$$

to a unit setpoint change for varying values of the proportional gain.

Notice that this steady state error decreases with increasing DC-gain K or Kp. Only with an infinite DC-gain K or Kp, the error will be zero. To conclude, a DC- offset between the reference value and process output will always exist with a pure P-controller, but can be made small by choosing Kp large.

However, when a step input is applied to a closed loop system with a very large Kp, the controller will respond very fast to any error with an almost full control action.

A too big Kp may then result in overshoot and oscillations, and the system can even become unstable. A trade-off has to be made between overshoot and DC-offset.

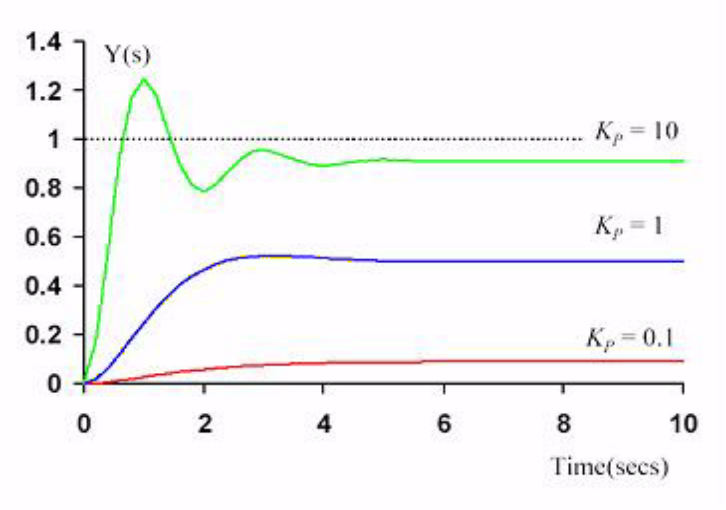


Figure 2-13 Effect of Proportional Gain Variation

For example, if we want to calculate the steady state error for the following system (with proportional control) when there is a unit step change in setpoint:

Eq. (2.24)

$$G_p = \frac{K}{s(Ts+1)}$$

The solution is as follows:

Eq. (2.25)

$$E(s) = R(s) - Y(s)$$

$$\frac{Y(s)}{R(s)} = \frac{\dfrac{KK_p}{s(Ts+1)}}{1+\dfrac{KK_p}{s(Ts+1)}} = \frac{KK_p}{KK_p + s(Ts+1)}$$

$$E(s) = R(s)\left[1 - \frac{KK_p}{KK_p + s(Ts+1)}\right]$$

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s) = s\frac{1}{s}\left[\frac{KK_p + s(Ts-1) - KK_p}{KK_p + s(Ts+1)}\right] = 0$$

This is a general result, which states that provided there is an integrator in the system, i.e. there is a factor of s in the denominator of the system (a pole at 0), then there will be no steady state error, for a change in setpoint, when the system is controlled using a proportional controller. This leads on to the second term in the 3-term controller, the integral control.

2.17.4. Integral Controller

If the requirement of a control system is no steady state error and the system contains no integrator, then one integrator can be introduced.

Despite the fact that proportional control does in general lead to a steady state error it is still a useful controller and has widespread application. For example when controlling the level of liquid in a tank, the requirement is typically that the tank must not overfill or run dry, which can be achieved using proportional control.

Integral control adjusts the output of the controller so that it is proportional to the integral of the error signal.

The output of an integral controller is given by:

Eq. (2.26)

$$Output = u(t) = K_i \int_0^t e(t)dt$$

$$G_c = \frac{U(s)}{E(s)} = \frac{K_i}{s}$$

The output of the controller with respect to the size of the error is demonstrated in the figure below:

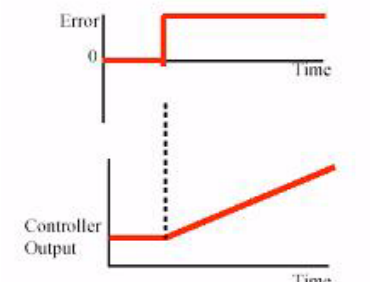


Figure 2-14 Proportional effect

For example, let us calculate the steady state error in the following system when there is a unit step change in setpoint:

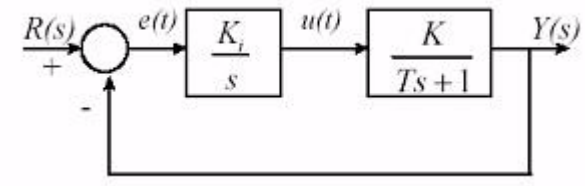


Figure 2-15 Proportional Controller Implementation

The solution is determined in the following way:

$$\frac{Y(s)}{R(s)} = \frac{\dfrac{KK_i}{s(Ts+1)}}{1+\dfrac{KK_i}{s(Ts+1)}} = \frac{KK_i}{Ts^2+s+KK_i}$$

$$E(s) = R(s) - Y(s)$$

$$E(s) = R(s)\left[1 - \frac{KK_i}{Ts^2+s+KK_i}\right]$$

$$E(s) = R(s)\left[\frac{Ts^2+s+KK_i-KK_i}{Ts^2+s+KK_i}\right]$$

$$\lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s) = s\frac{1}{s}\left[\frac{Ts^2+s}{Ts^2+s+KK_i}\right] = 0$$

This example demonstrates that integral action can be used to eliminate steady state offsets. Provided there is an error signal, the controller output will continue to change, thus driving the system until the error becomes zero. Unfortunately, using integral action alone can lead to poor damping, which causes the response of the system to become more oscillatory. To avoid this a combination of proportional and integral control is often used (PI control) - Two Term Control.

2.17.5. Proportional plus Integral (PI) Controller

The output of a PI controller is given by:

$$Output = u(t) = K_p e + K_i \int_0^t e(t)dt$$

$$G_c = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} = K_p\left(1+\frac{1}{T_i s}\right)$$

Where Ti is called Integral or reset time (= $\frac{K_P}{K_I}$), and is known as Reset Rate.

Schematically a PI controller can be represented by the following block diagram:

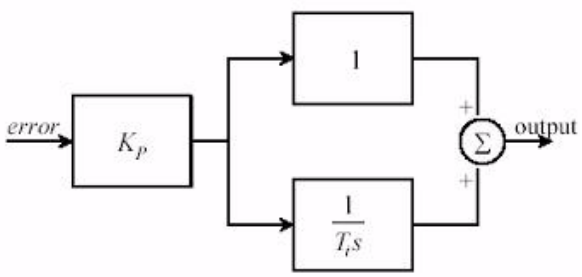


Figure 2-16 Integral Controller Implementation

The figure below shows the output of a PI controller against the error input.

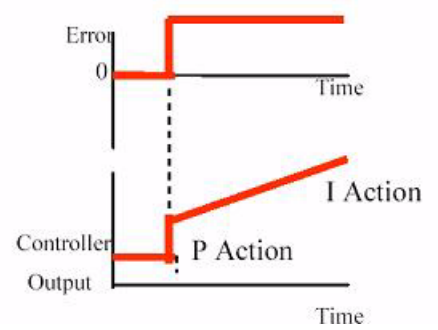


Figure 2-17 Integral effect

The graph below shows the response of the system:

Eq. (2.27)

$$G(p) = \frac{5}{10s + 1}$$

to step changes in set-point for varying values of the integral action. For each curve the value of the gain is fixed.

To summarize the observations from this graph, increasing the value of TI, increases the amount of oscillation in the system.

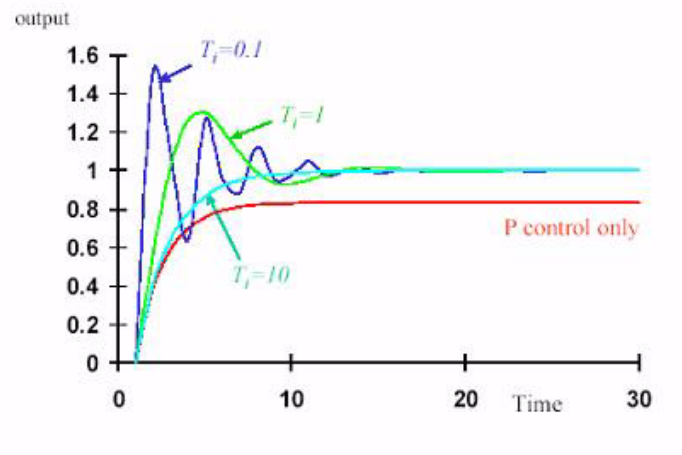


Figure 2-18 Effect of variations of the integral time

2.17.6. Derivative Controller

The final element of a 3-term controller is derivative action. Derivative control attempts to anticipate future error values from the rate of change of error in the control algorithm. The output of a derivative controller is set to be proportional to the rate of change of error:

Eq. (2.28)

$$Output = u(t) = K_d \frac{de(t)}{dt}$$

$$G_c = \frac{U(s)}{E(s)} = K_d s$$

Where Kd is the Derivative Gain. The relationship between error and controller output is shown below:

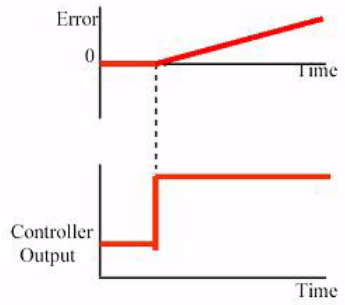


Figure 2-19 Derivative effect

The characteristics of a derivative controller can be summarized as follow:

• Derivative control increases the damping in the system, but tends to amplify noise.

• Derivative control will be insensitive to constant or slowly changing errors and is therefore typically used with other forms of controller.

2.17.7. Proportional plus Derivative (PD) Control

As explained above, proportional control can produce oscillatory behaviour. Since derivative control increases stability and consequently reduces oscillations, for those systems where it is important to minimize the amount of oscillation, PD control is applied. It should be noted that PD control will produce a steady state error unless the system contains an integrator. For many servo-mechanical systems, such as positional control by motors, there is an integrator (position is the integral of velocity – velocity = rate of change of position) and therefore PD control is suitable. The output of a PD controller is given by:

$$Output = u(t) = K_p e(t) + K_d \frac{de(t)}{dt}$$

$$U(s) = K_p E(s) + K_d s E(s)$$

$$\frac{U(s)}{E(s)} = K_p + K_d s = K_p \left(1 + \frac{K_d}{K_p} s \right)$$

$$G_c = K_p \left(1 + T_d s \right)$$

Where Td is the Derivative Time Constant.

The controller is schematically represented as follows:

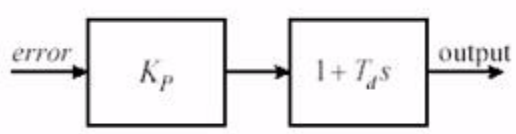


Figure 2-20 Derivative Controller Implementation

The response of a second order system to a step change in input under both proportional and PD control is shown in figure 2-20. In each case the value of the proportional gain is kept constant.

Since there is typically still an offset with PD control, Proportional plus Integral and Derivative control is often used. This is known as PID or 3-term control.

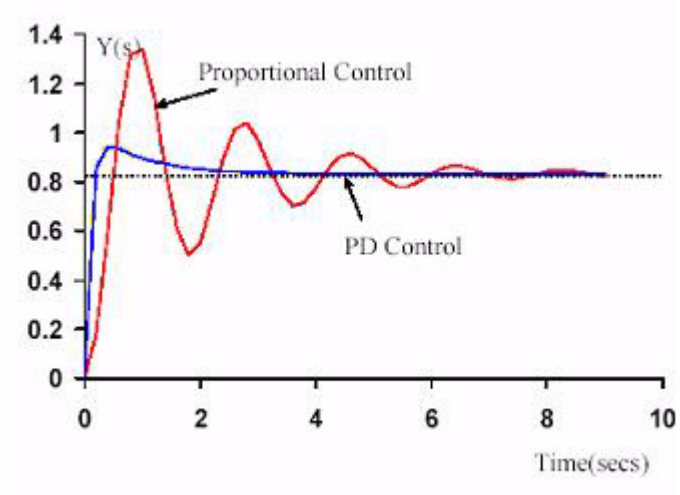


Figure 2-21 Effect of variations of derivation time

2.17.8. Proportional plus Integral plus Derivative (PID)Control

The output of a PID controller is determined as follows:

$$Output = u(t) = K_p e(t) + K_i \int_0^t e(t)dt + K_d \frac{de(t)}{dt}$$

$$U(s) = K_p E(s) + \frac{K_i}{s} E(s) + K_d s E(s)$$

$$\frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} + K_d s = K_p \left(1 + \frac{K_i}{K_p s} + \frac{K_d}{K_p s}\right)$$

$$G_c = K_p \left(1 + \frac{1}{T_i s} + T_d s\right)$$

Once the type of controller has been determined for a particular system, the values of the controller constants, the proportional gain, the reset rate and the derivative time, must be determined. The determination of these values is known as Controller Tuning. For a PID controller it is not a trivial task, for example large values of proportional gain may increase the speed of response but they will also reduce the stability of the system. Large values of the reset rate will mean that the steady state offset is eliminated but will also increase the oscillations of the system. A large derivative time will increase the stability of the system but may also produce poor control if the measured signals contain large amounts of noise.

There are a number of different PID controller structures [27]. Different manufacturers design controllers in different manner. However, two topologies are the most often case:

* parallel (non-interactive)
* serial (interactive).

Parallel structure is most often in textbooks, so it is often called "ideal" or "non-interactive" structure because proportional, integral and derivative mode are independent on each other. Parallel structure is still very rare in the market, the reason for that being mostly historical: first controllers were pneumatic and it was very difficult to build parallel structure using pneumatic components. Due to certain conservatism in process industry most of the controller used there are still in serial structure, although it is relatively simple to realize parallel structure controller using electronics. In other areas, where tradition is not so strong, parallel structure can be found more often.

A serial connection of proportional, derivative and integral element is called serial or interactive structure of PID controller. Parallel structure is shown in figure 2-22.

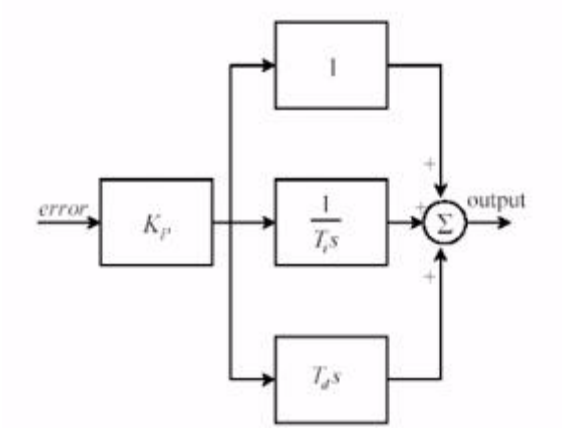


Figure 2-22 PID "Parallel" Structure

2.17.9. PID Controllers and measures of control performances

The elements of a PID controller are summarized in the following table:

Or, with respect to response speed, stability and accuracy:

	VALUE	TOO MUCH	TOO LITTLE
P	Most Stable	Stability Decreases	Larger Offset Slower Response
I	Removes Offset	Stability Reduces Longer Oscillation Period	Slower Return to Set-point
D	Increases Stability Reduces Oscillation Period	Stability Increases Amplifies Noise	Max benefit not achieved

Figure 2-23 PID parameters - 1

Parameter	Speed of response	Stability	Accuracy
increasing K	increases	deteriorate	improves
increasing K_i	decreases	deteriorate	improves
Increasing K_i	increases	improves	no impact

Figure 2-24 PID parameters - 2

And, to further summarize:

mode	benefit	problem
proportional	• rapid adjustment of manipulated variable • speeds dynamic response	• non-zero offset • can cause instability
integral	• produces zero offset	• slow dynamic response • can cause instability
derivative	• provides rapid response to control variable changes	• non-zero offset • sensitive to noise in controlled variable

Figure 2-25 PID parameters - 3

The values for the three PID parameters are typically chosen to minimize a performance criteria.To perform an optimization, we need quantitative measures of control performance, like the following standard measures for the controlled variable:

Name	Description	Equation
ISE	Integral of Square Error	$\int_0^T e^2(t)\,dt$
IAE	Integral of Absolute Error	$\int_0^T \lvert e(t)\rvert\,dt$
ITAE	Integral of Time-multiplied Absolute Error	$\int_0^T t\,\lvert e(t)\rvert\,dt$
ITSE	Integral of Time-multiplied Square Error	$\int_0^T t\,e^2(t)\,dt$

Figure 2-26 Errors minimization

IAE, ISE, and ITAE increase as the controlled variable spends time away from the set point.

The ITAE weights the error within time and hence emphasizes the error values later on in the response rather than the initial large errors.

IAE gets the absolute value of the error to remove negative error components, and is usually used for simulation studies.

The ISE squares the error to remove negative error components, discriminating between over-damped and under damped systems, i.e. a compromise minimizes the ISE:

2.18. PID tuning

Tuning means setting the adjustable parameters of a controller to give best performance.

PID is defined entirely from process dynamic concepts – notions of error and transient response - not specific to any technology or branch of engineering; it's a general purpose control algorithm.

One prime difficulty is that each of the parameters (gain, reset, and derivative time) affects more than one characteristic of the response. For example, both the amplitude and frequency of an underdamped response respond to any of the parameters. The settings must be adjusted in concert for best performance [26].

The notion of optimization is to adjust parameters for best possible result. In a curve fit, we adjust fitting parameters to get smallest error between data and model. In control, similarly, we could adjust the 3 knobs of the PID algorithm to make, for instance, the IAE a minimum. In an analytical calculations of closed loop performance of idealized processes, one could easily find out how to minimize IAE – simply set the gain to infinity. For more realistic processes, we could solve the equations numerically.

The real world is more complicated, however. In fact [46]:

1. Our models will not be perfect, so we would have an optimum tuning of a non-existent process;

2. We optimized for one operating condition, but didn't consider how well the controller would work for other legitimate conditions (such as new set points);

3. Our optimization objective function was defined entirely in terms of the controlled variable, irrespective of what the manipulated variable had to do.

The classic trade-off is between tight control (implying aggressive movements of manipulated variable, closer approach to the instability boundary) and loose control (milder use of manipulated variable, more variation in the controlled variable).

Tight control tends to reduce variation in the controlled variable at the expense of greater variation in the manipulated variable. It also tends to be more sensitive to variations in process parameters (such as time constants that are not really constant). Looser control may be more forgiving, or robust, to such variations. Tuning, then, is an attempt to achieve an acceptable combination performance and robustness [26].

Tuning might be done by:

1. Doing detailed simulations with rigorous process models, seeking to optimize quantitative measures of the deviation.

2. Using correlations to choose tuning parameters, based on general process models.

3. Turning knobs in the control room, observing the results and making further adjustments.

A number of analytical and numerical methods have been proposed for tuning this kind of controller since 1940; these methods are usually different in complexity, flexibility and in the amount of process knowledge used. Nevertheless, there is no generally accepted design method for them. Therefore, the design of PID controllers still remains a challenge before researchers and engineers.

A large number of industrial plants can approximately be modelled by a first order plus time delay (FOPTD) transfer function as follows:

Eq. (2.29)

$$G(S) = \frac{Ke^{-\tau_d S}}{TS+1}$$

To design PID controllers for this important category of industrial plants, various methods have been suggested during the past sixty years.

Traditionally, PID controllers have been tuned empirically, e.g., by the first method of Ziegler and Nichols, called “continuous cycling method”, widely known as a fairly accurate heuristic method to determine good settings for controllers for a wide range of common industrial processes. It also has the advantage of requiring very little information about the

process; however, it requires knowledge about the ultimate data which are obtained by destabilizing the system under proportional feedback. Moreover, the method inherently leads to an oscillatory response in the face of a change in the setpoint.

Similar to the Ziegler and Nichols methods, Cohen and Coon technique sometimes brings about oscillatory responses, because it was designed to provide closed loop responses with a damping ratio of 25% [16]. However, Ziegler and Nichols methods are still widely used, either in their original form or with some modifications [36].

The factors that affect practical tuning of closed loop control systems are as follows:

1) Steady state gain;

2) Load inertia and capacity of system to overcome (capacity response rate);

3) "Rigidity" of coupling through system & transportation delays (Time Constants of system elements);

4) Deadbands (and Backlash);

5) Susceptibility to internal noise.

Often, the Derivative component is unnecessary in critically damped, linear systems and can be wound out/ disconnected. Exceptions may occur in:

a) systems that are intentionally non-linear,

b) overdamped linear systems with excessive deadbands or transportation delays and

c) some positioning systems.

In linear systems, careful scrutiny of the degree to which transportation errors are present and the level of steady state gain needed should be made.

In systems of all types, Reset Derivative (Phase Advance), for example, is often used to mask electrical and/or mechanical coupling shortfalls and one has to decide if these are prompting an apparent but incorrect need for the Derivative component. Reference Derivative has a legitimate role to play in overdamped linear systems providing the need for overdamping has been verified.

Practical systems are usually comprised of elements capable of a resolution of only one decade better than the steady state accuracy the process requires. (Generally, there is no point of trying to set a loop to achieve stability and accuracy at 1/10.000 when the set point can only be set to 1/100). The higher the system gain, the more critical will be the coupling requirements and the more difficult the loop will be to stabilize. The response of a system to a stepped change provides information about the system constants. Initially there is little change in response (indicating the degree to which transportation delays and deadbands exist).

The tangent of the steepest slope of the response curve is taken to be the initial slope of a single exponential curve. The time interval between the application of the step change and the point where the tangent of the steepest slope intercepts with the initial value of the output is taken as the transport lag of the system. "Ziegler and Nichol's charts" show the transient response to a step change and controller settings can be based on readings from these charts.

Care must be taken to ensure that, if used, analogue amplifiers do not enter saturation and that error signals are not being limited or clipped. There is only one good solution to induced noise - eliminate the source. Loops nested to levels greater than 3 may need special methods.

2.18.1. Internal Model Control (IMC)

Using an on-line parameter estimation algorithm to identify the parameters of the model, the parameters of most linear model based controllers can be adjusted in line with changes in process characteristics. Although great strides have been made in resolving the implementation issues of adaptive systems, for one reason or other, many practitioners are still not confident about the long term integrity of the adaptive mechanism. This concern has led to another contemporary topic in modern control engineering: robust control [46].

Robust control involves, firstly, quantifying the uncertainties or errors in a 'nominal' process model, due to nonlinear or time-varying process behaviour for example. If this can be accomplished, we essentially have a description of the process under all possible operating conditions. The next stage involves the design of a controller that will maintain stability as well as achieve specified performance over this range of operating conditions. A controller with this property is said to be 'robust' [30].

A sensitive controller is required to achieve performance objectives. Unfortunately, such a controller will also be sensitive to process uncertainties and hence suffer from stability problems. On the other hand, a controller that is insensitive to process uncertainties will have poorer performance characteristics in that controlled responses will be sluggish. The robust control problem is therefore formulated as a compromise between achieving performance and ensuring stability under assumed process uncertainties. Uncertainty descriptions are at best very conservative, whereupon performance objectives will have to be sacrificed. Moreover, the resulting optimisation problem is frequently not well posed. Thus, although robustness is a desirable property, and the theoretical developments and analysis tools are quite mature, application is hindered by the use of daunting mathematics and the lack of a suitable solution procedure.

Nevertheless, underpinning the design of robust controllers is the so called 'internal model' principle. It states that unless the control strategy contains, either explicitly or implicitly, a description of the controlled process, then either the performance or stability criterion, or both, will not be achieved. The corresponding 'internal model control' design procedure encapsulates this philosophy and provides for both perfect control and a mechanism to impart robust properties (see Fig. 2-27).

If the process model is invertible, then the controller is simply the inverse of the model. If the model is accurate and there is no disturbance, then perfect control is achieved if the filter is not present.

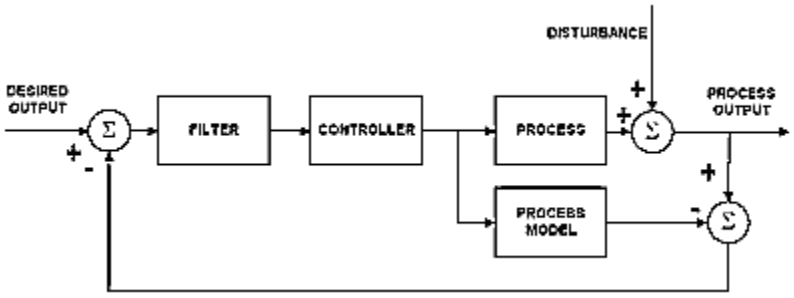


Figure 2-27 Schematic of Internal Model Control Strategy

This also implies that if we know the behaviour of the process exactly, then feedback is not necessary. The primary role of the low-pass filter is to attenuate uncertainties in the feedback, generated by the difference between process and model outputs and serves to moderate excessive control effort. The strategy and the concept that it embraces are clearly very powerful. Indeed, the internal model principle is the essence of model based control and all model based controllers can be designed within its framework.

2.19. Control of Multiple-Input, Multiple-Output Processes

Many systems show a strong interaction among different control loops; it means that one controlled variable can affect other variables in the neighboring loops, originating what it is called a Multiple Input-Multiple Output (MIMO) system.

We consider decentralized control of a general square multivariable n×n system G(s), i.e., having n inputs and n outputs. Decentralized control implies that the overall system is decomposed into a number of interacting subsystems for which individual controllers are designed. Such a decomposition of the control problem is often preferred due to the robustness of such structures, both with respect to model uncertainty and sensor/actuator failures, and the ease of (re)tuning compared to the case with full multivariable controllers.

However, the potential cost of using a limited controller structure is reduced closed loop performance due to the presence of interactions among the subsystems. The main tasks involved in the design of a decentralized control system, as defined, e.g., in [24], are:

1) selection of controlled outputs
2) selection of manipulated variables (inputs) and measurements
3) selection of a control configuration, or structure of interconnections
4) selection of controller types and tuning of the single controllers.

The most commonly used tool for control structure selection for single loop controllers is still the Relative Gain Array (RGA), introduced by Bristol (1966) [39].

Some important advantages of the RGA are that it depends on the plant model only, that it is scaling independent and that all possible configurations can be evaluated based on a single matrix.

Dealing with MIMO systems implies also:

1) Modeling the interactions
2) Decoupling strategies
3) Multiloop-loop model-predictive
4) Handling constraints.

In practical control problems there typically are a number of process variables which must be controlled and a number which can be manipulated . In figure 2-28 different types of control problems are shown (note "process interactions" between controlled and manipulated variables). If process interactions are significant, even the best multiloop control system may not provide satisfactory control. In these situations there are incentives for considering multivariable control strategies .

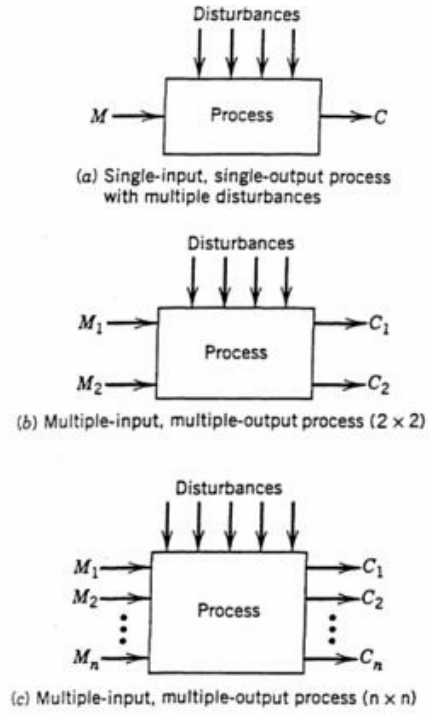


Figure 2-28 Process Interactions

2.19.1. Multiloop Control Strategy

This is a typical industrial approach, consisting of using n standard PID controllers, one for each controlled variable. In this respect, a control system design will consist of:

1. Select controlled and manipulated variables.
2. Select pairing of controlled and manipulated variables.
3. Specify types of FB controllers.

Example: 2×2 system [32]

Figure 2-29 2×2 system

There are two possible controller pairings (M is the manipulated variable, C the controlled variable):

* M1 with C1, M2 with C2 or
* M1 with C2, M2 with C1

For nxn system there are n! possible pairing configurations.

Let us consider the case of two controlled variables and two manipulated variables, which requires four transfer functions:

Eq. (2.30)

$$\frac{C_1(s)}{M_1(s)} = C_{P11}(s), \quad \frac{C_1(s)}{M_2(s)} = C_{P12}(s)$$
$$\frac{C_2(s)}{M_1(s)} = C_{P21}(s), \quad \frac{C_2(s)}{M_2(s)} = C_{P22}(s)$$

Thus the input-output relations for the process can be written as:

Eq. (2.31)

$$C_1(s) = G_{P11}(s)M_1(s) + G_{P12}(s)M_2(s)$$
$$C_2(s) = G_{P21}(s)M_1(s) + G_{P22}(s)M_2(s)$$

Or in vector-matrix notation as:

Eq. (2.32)

$$\underline{C}(s) = \underline{\underline{G_P}}(s)\underline{M}(s)$$

where

$$\underline{C}(s) = \begin{bmatrix} C_1(s) \\ C_2(s) \end{bmatrix}, \quad \underline{M}(s) = \begin{bmatrix} M_1(s) \\ M_2(s) \end{bmatrix}$$

are vectors, and

$$\underline{\underline{G_P}}(s) = \begin{bmatrix} G_{P11}(s) & G_{P12}(s) \\ G_{P21}(s) & G_{P22}(s) \end{bmatrix}$$

is the Transfer Function matrix for the process.

Process interactions may induce undesirable interactions between two or more control loops. For example, in a 2×2 system there are control loop interactions due to the presence of a third feedback loop. Moreover, closed -loop system may become destabilized, and the controller tuning becomes more difficult.

2.20. The relative gain array (RGA)

Single loop controllers in large scale decentralized control systems often are tuned in an independent way. This means that the surrounding controllers often are put in manual mode, before a certain controller is tuned. In this way the loop to be tuned is not influenced by the other loops. However, after the tuning is completed and the surrounding controllers are taken into service again these loops will interact with the newly tuned loop. These interactions will in general affect the performance of this loop and thus a pairing which minimizes the interactions between single loops and the remaining system should be preferred [32] [34].

The relative gain of a scalar subsystem is the ratio between the open-loop transfer-function and the transfer-function resulting when all other outputs are perfectly controlled. The common rule is to pair on subsystems with relative gains close to 1, based on the assumption that the interactions then will have a small effect on the subsystems behaviour. Since the frequency corresponding to the desired bandwidth is the most important, the focus is usually on the relative gains around this frequency. An often cited advantage of the RGA is that it is controller independent, due to the assumption of perfect control. However, this assumption is at the same time an important shortcoming of the RGA. The effect of interactions is most important around the desired bandwidths of the subsystems, and at these frequencies the gain changes due to finite bandwidth control can be highly different from the changes predicted by the RGA [46].

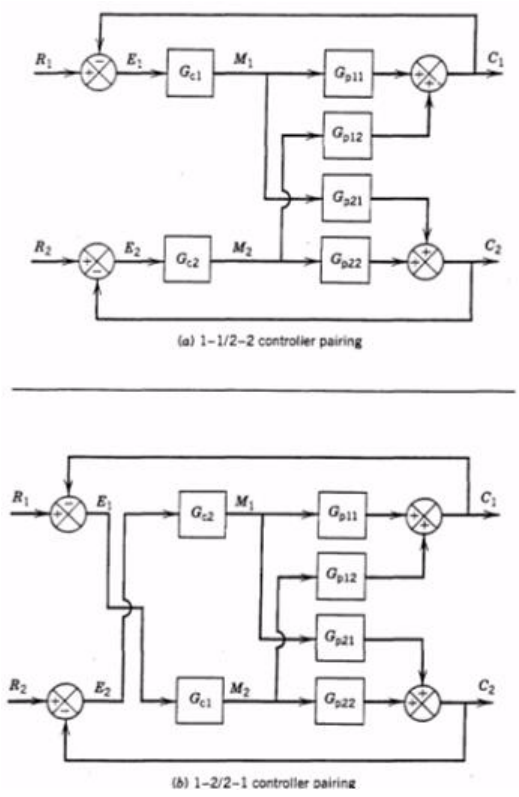


Figure 2-30 Block diagram for conventional 2×2 multiloop schemes

This is to some extent due to the fact that the individual loops have non-perfect control at this frequency, but is primarily due to the fact that the remaining (n-1)×(n-1) system, considered perfectly controlled when calculating the relative gain for a specific subsystem, in general will have a performance which differs significantly from the performances of the individual loops (due to the presence of interactions within that subsystem). This may be one reason why the RGA, while usually working well for 2×2 systems, is less effective for systems with n > 2.

The RGA provides two useful types of information:

1) Measure of process interactions

2) Recommendation about best pairing of controlled and manipulated variables.

It requires knowledge of steady state gains but not process dynamics. For example, the case of a 2×2 system gives:

Steady-state process model,

$$C_1 = K_{11}M_1 + K_{12}M_2$$
$$C_2 = K_{21}M_1 + K_{22}M_2$$

Then, the RGA is defined as:

$$RGA = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix}$$

where the relative gain, λ_{ij}, relates the ith controlled variable and the j^{th} manipulated variable. For a 2×2 system,

Eq. (2.33)

$$\lambda_{11} = \frac{1}{1 - \frac{K_{12}K_{21}}{K_{11}K_{22}}}, \qquad \lambda_{12} = 1 - \lambda_{11} = \lambda_{21}$$

One can follow the guidelines provided by Marlin [46] to decide about the interaction level based upon λ_{ij}:

$\lambda_{ij} < 0$ Open- and closed-loop process gains are of different signs. In a 2×2 process, if the single-loop controller gain were positive for stable feedback control, the same controller gain would have to be negative for stable multiloop feedback control. THus, the sign of the controller gain to retain stability would depend on the mode of other controllers in the multiloop system -not a desirable situation.

$\lambda_{ij} = 0$ This situation can occur when the open-loop process gain is zero, which indicates no steady-state relationship between the input and output variables. This is a situation not generally desirable, since a controller with this pairing can function only when other controllers are in automatic.

$0 < \lambda_{ij} < 1$ The steady-state loop process gain with the other loops closed is larger than the same process gain with the other loops open.

$\lambda_{ij} = 1$ There is no transmission interaction, but this not preclude the possibility that the manipulated variable might affect another variable (one-way interaction).

$\lambda_{ij} > 1$ The steady -state loop process gain with the other loops closed is smaller than the same process gain with the other loops open.

$\lambda_{ij} = \infty$ When the process gain is zero with the other loops closed, it is not possible to control the variable in a multiloop system.

3 Thermal screen constraints and specifications

The main task of the Thermal Screen is to perform the dynamical thermal insulation between the "cold" (temperatures around -10 °C) Tracker and the "warm" (temperatures around +18 °C) Electromagnetic Calorimeter environments (see Figure 3-1).

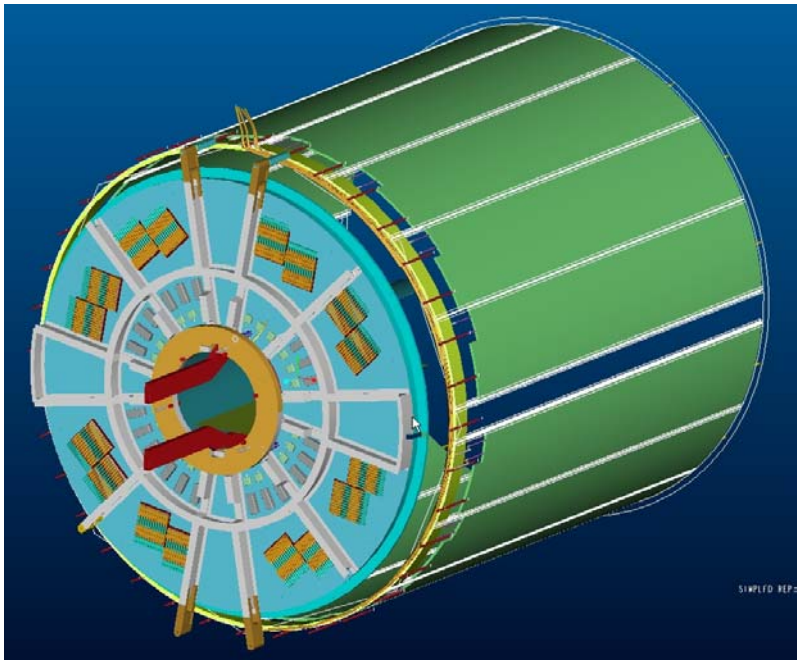


Figure 3-1 Tracker - External rails for the thermsl screen

In normal operational conditions the temperature on the outer skin of the Tracker supporting structure must be coincident with the average environmental temperature inside the experiment, i.e. +18 °C. In order not to induce distortions in the Tracker Support Tube, this equilibrium temperature must be reached on the outer skin of the Thermal Screen. The inside skin of the Screen, facing the cold Tracker volume, must preserve the maximum efficiency of the cooling system of the detector electronics. Thus, it has to provide a thermally

neutral interface to the cold volume, i.e. a wall temperature not higher than -10 °C. This fixes to an average value of 28 °C the thermal gradient to be achieved through the thickness of the screen.

The requirements on the Thermal Screen performance are the following [8]:

1) Permanently regulate the temperature of the layer outside the Tracker volume, thus making it acceptable for the operation of the "warm" Electromagnetic Calorimeter. Failure possibilities of the Thermal Screen system have to be minimized following the methodology explained in the standard and they have to be taken into account in the general operation of the CMS detector.

2) Maintain the correct subzero temperatures of the Tracker volume in the outer Tracker layer (which coincides with the Thermal Screen cooling panel). To add extra guarantees against failures of the Tracker Cooling system, the Thermal screen coolant shall be supplied by an individual cooling plant. Failure possibilities of the Thermal Screen system have to be minimized following the methodology explained in the standard and they have to be taken into account in the general operation of the Tracker. In case of failure of the Tracker's Cooling system all the detector electronics has to be switched off and the Thermal Screen must ensure the maintaining of the required temperature conditions inside the cold volume. In this particular case, the temperature uniformity requirement between the inside skin and the outer skin of the Supporting Tube will be 4 °C.

3) Guarantee the integrity of the Tracker Support Tube. In order not to induce distortions to the Tracker Support Tube, temperature equilibrium must be reached on the outer skin of the Thermal Screen.

The inside skin of the Thermal Screen has to provide a thermally neutral interface to the cold volume, i.e. a wall temperature not higher than -10 °C. This fixes the average value of the thermal gradient to be achieved through the thickness of the screen. The required temperature uniformity both on the inside skin and the outer skin of the Supporting Tube is 2.5 °C. All the mentioned operative requirements must be ensured in all the standard operational conditions of the Tracker, i.e. with the detector cooling system on and the electronics both partially and fully on and off. The possibility and consequences of a general failure of the Tracker cooling system must be minimized following the methodology explained in the standard and they have to be quantified and taken into account in the design of the operation of the Thermal Screen.

In normal operational conditions the temperature on the outer skin of the Tracker supporting structure must be coincident with the average environmental temperature inside the experiment, i.e. +18 °C. In order not to induce distortions in the Tracker Support Tube, this equilibrium temperature must be reached on the outer skin of the Thermal Screen. The inside skin of the Screen, facing the cold Tracker volume, must preserve the maximum

efficiency of the cooling system of the detector electronics. Thus, it has to provide a thermally neutral interface to the cold volume, i.e. a wall temperature not higher than -10 °C. This fixes to an average value of 28 °C the thermal gradient to be achieved through the thickness of the screen.

The required temperature uniformity of both the inside cold skin and the outer skin of the Supporting Tube is 2.5 °C. All the mentioned operative requirements must be ensured in all the standard operational conditions of the Tracker, i.e. with the detector cooling system on and the electronics both on and off.

In case of a general failure of the Tracker cooling system, all the electronics of the detector is switched off and the Thermal Screen must ensure the maintaining of the required temperature conditions inside the cold volume. In this particular case, the temperature uniformity requirement can be released to 4 °C and the average value of the temperature on the outer skin of the Tracker Supporting Tube can be temporary lowered to +16 °C.

3.1. General Description

The Thermal Screen is composed of a heating and a cooling surface, separated by a layer of insulating material, each one able to adapt its temperature to the value desired for the facing volume. The heat flow, naturally due through the surface separating the two volumes at different temperatures, is artificially produced on one side of the Screen and removed on the other side; this removes any thermal boundary layer in the regions facing the separating surfaces and confines the thermal gradient within the thickness of the Screen. The heat is produced on one side of the Screen through a series of electrical resistances and is removed on the other side by a conductive plate kept cold by flow rate of sufficient thermal capacity.

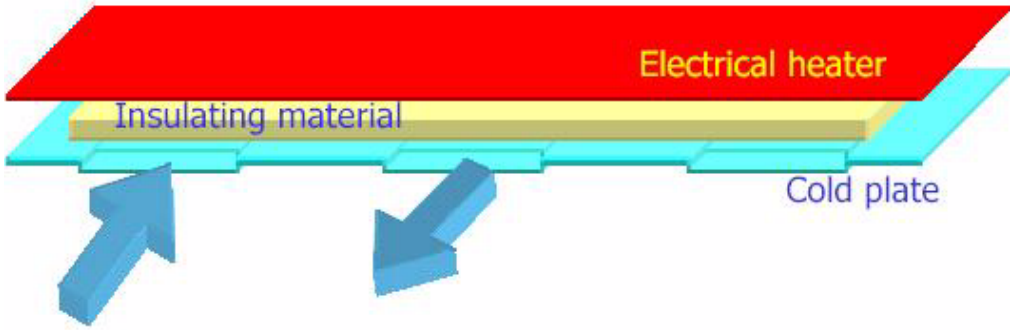


Figure 3-2 Thermal Screen's Three Layers

The presence of the insulating material allows for a reduction of the total amount of heat to be produced and removed, thus limiting the needed heating power and cold flow rate.

The coverage of the surface of the Support Tube is achieved by 32 panels, divided in two rows, each one spanning half-length of the Tracker. Sixteen panels form each row, each one

covering a circular sector of 21°. Through all the length of the Tracker openings are left at 3 o'clock and 9 o'clock positions, in order to allow for the insertion of the Tracker sub-detectors via a system of fixed rails.

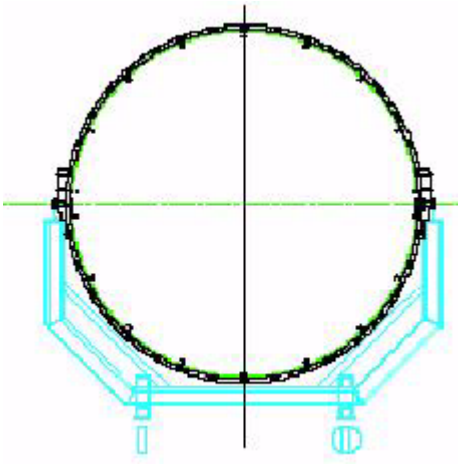


Figure 3-1 Tracker Support Tube (Front View)

The assembly of a curved cold plate, of a layer of insulating material and of a thin heating foil basically constitutes a screen panel. Each panel is independently inserted inside the Support Tube by sliding on dedicated Glass Fibre U-profiles attached to the inner skin of the Tube. In order to avoid damages to the ECAL hardware there must be no water inside the ECAL volume; besides, not to compromise data quality, the local temperature must be of 18 °C with a tolerance of ±1 °C.

Figure 3-2 Tracker Support Tube

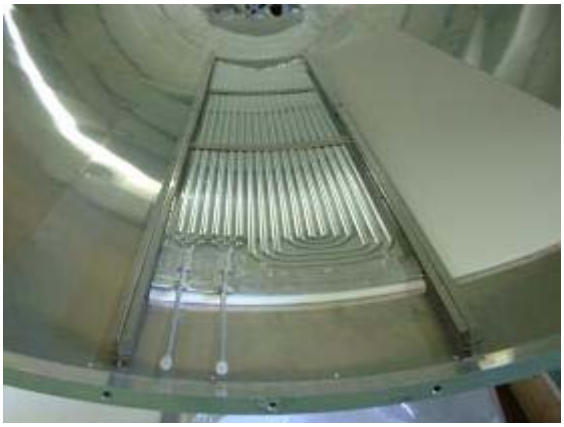


Figure 3-3 A Thermal Screen Panel Installed

3.1.1. Modularity

The modularity of the Thermal Screen is dictated by several needs:

- > Match the multiplicity of the service channels connecting the Tracker sub-detectors to their powering, cooling and read-out systems;
- > Limit the size and weight of each module to dimensions that can be reasonably handled two persons;
- > Fix the surface of each module in such a way that the local temperature variation due to a possible failure of a module does not induce deformations of the Tracker Support Tube exceeding 0.1 mm.

3.1.2. Accessibility and Maintenance

The Thermal Screen is mounted in the Support Tube prior to the installation of the Tracker sub-detectors and is then blocked by the fanning out services of the sub-detector. For this reason, the screen must be designed to be essentially maintenance-free. Nevertheless, in case of a failure of a module, it must be possible to access it by de-cabling only one service sector and to extract and substitute it without removal of the Tracker sub-detectors. As this operation is only possible during a winter shutdown, the feeding system of the cold part of the screen must include the possibility of excluding a module from the circuit in case of failure.

3.1.2.1. The Heating Foils

The heating foils are standard Kapton foil FEP 205FN029) with an etched circuit in Inconel 600, 25 μm thick. The total foil thickness is typically 0.1 mm. These heaters are radiation-hard till above 106 rads and, in the required temperature range, can easily accept up to 45 $W/_{m}2$.

Electrical specifications:
Dielectric strenght (V * t)= 1000 V * 1 min.
Electrical insulation > 500 MΩ.
Dielectri rigidity= 1250 Vac with a threshold of 5 mA.

These panels are produced by Rica Industries in two sizes to allow for a fine tuned control:

- Size 1: Three panels measuring 443*435 mm, with an electrical resistance of 79.7 Ω ±5 % at 22 °C.

- Size 2: Two panels measuring 560*435mm, with an electrical resistance of 63 Ω.±5 % at 22 °C.

Leads are made of 4 m long Polyimide AWG26 cables according to ESA/SCC 3901/001-24, capable of 3.61A ±5 % @ 48 V Max and with a resistance of 1.02 Ω ±5 %.

Every foil is equipped with a Pt 1000 RTD (Resistance Thermal Device) temperature sensor embedded in the capton. The sensor specifications are the following:

Manufacturer: Sensotherm GmbH
Type: PTFB 102
Size: 2.0 × 10.0 × 1.4 mm
Nominal Resistance at 100 °C: 1000 Ω
Temperature Range: -50 °C - +300 °C

The Heating Foils have been glued together to form the thermal screen panel at CERN and, after the acceptance tests, will be kept in stand-by for a period of ~2 years, during the assembly of the experiment. Following this phase, the operational life of the foils, in a dry Nitrogen atmosphere (dew point = –35 °C), is estimated to be not less than 10 years of continuous operation.

Throughout the whole operational life the Heating Foils will be monitored in temperature. The temperature signal will be used as feedback for the control of the power supplies, which will regulate the supply voltage such as to match the room temperature of the environment of the experimental cavern (18 ± 5 °C). In any case, the allowed range of temperature for the operation of the foils should be not less than –30 °C(min)/+40 °C(max)

The maximum supply voltage to the Heating Foils will not exceed 48 Volts.

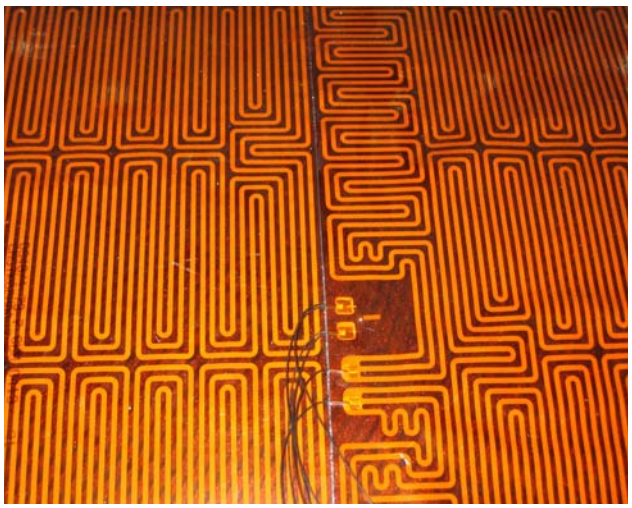


Figure 3-4 Heating foil - RTD detail

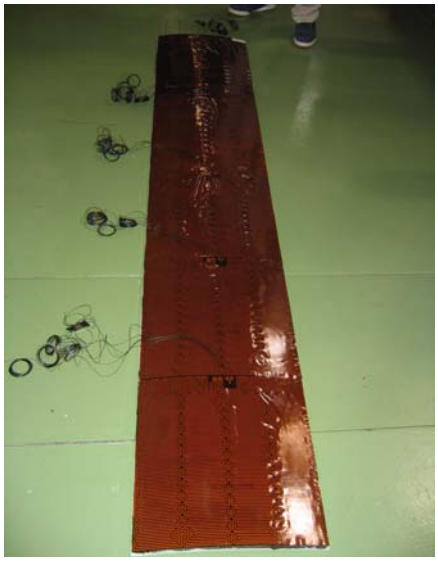


Figure 3-5 A thermal screen panel laying on the floor

The environment surrounding the Heating Foils will be subject to a 4 Tesla magnetic field acting along the axis of the Support Tube for 6 months per year and will become progressively more radioactive during operation. The estimated integrated dose absorbed by the Heating Foils at the end of the design operational life is about 5 kGy.

3.1.3. The Insulating Layer

The use of a Rohacell 51IG foam (k = 0.033 W/mK), allows for a material whose interesting structural properties would naturally improve the stiffness of the Screen panels. The main concerns related to the use of this material come from its flammability, its sensitivity to moisture absorption and from the unavailability of an exhaustive documentation on its long-term behaviour under radiation.

Rohacell is a closed-cell rigid expanded plastic material for lightweight sandwich construction. It has excellent mechanical properties, high dimensional stability under heat, solvent resistance and, particularly at low temperature, a low thermal conductivity. The strength and moduli values are the highest are the highest for any foamed plastic in its density range. Rohacell is manufactured by hot forming of methacrylic acid/methacrylonitrile copolymer sheets. During foaming this copolymer is converted to polymethacrylimide.

The main characteristics of Rohacell are specified in the following table:

Physical Properties	Metric	Comments
Density	0.0513 g/cc	ASTM D 1622-63
Water Absorption	17.4 %	Equilibrium at 98% RH
Moisture Absorption at Equilibrium	1.3 %	15% RH
Moisture Absorption at Equilibrium	2.6 %	30% RH
Moisture Absorption at Equilibrium	4.2 %	50% RH
Moisture Absorption at Equilibrium	5 %	65% RH
Water Absorption at Saturation	15 %	Volume %; 50 days immersion at 20*f*C.
Moisture Expansion	Max 2 %	50 days immersion at 20*f*C.
Mechanical Properties		
Tensile Strength, Ultimate	1.86 MPa	ASTM D 638-68
Elongation @ break	4 %	ASTM D 638-68
Modulus of Elasticity	0.0686 GPa	ASTM D 638-68
Flexural Yield Strength	1.57 MPa	ASTM D 790-66
Compressive Yield Strength	0.883 MPa	ASTM D 1621-64
Shear Modulus	0.0186 GPa	ASTM C 273-61
Shear Modulus	0.0206 GPa	ASTM D 2236-69
Shear Strength	0.786 MPa	ASTM C 273-61
Electrical Properties		
Dielectric Constant	1.06	10.0 GHz
Dielectric Constant	1.07	2.0 GHz
Dielectric Constant	1.09	5.0 GHz
Dielectric Constant	1.11	26.0 GHz
Dissipation Factor	0.0002	2.0 GHz
Dissipation Factor	0.0004	5.0 GHz
Dissipation Factor	0.0011	10.0 GHz
Dissipation Factor	0.005	26.0 GHz

3.1.4. The Cold Plates

The cold plate is made by two thin aluminium sheets (0.7 mm ± 0.1 thick each) joined together such as to produce a spiralling triple channel (2.5 x 20 mm cross section), through which the cooling liquid flows. The technique used is the "hot-roll-bonding" or co-lamination at the Pechiney Rhenalu workshops under control of DATE company. The spiralling path is designed to always have neighboring channels in counter-current flow. Also, the distance between two adjacent channels is dimensioned for a thermal efficiency of the connecting plate always higher than 95%, so that the temperature non-uniformity on the cold plate surface is kept at a minimum.

Figure 3-6 Cold plate pipes schematic

The cold plate is 2500 mm long and 420 mm wide, bent to a radius of curvature of 1175 mm.

Quality assurance include tests at:

1. 2 bar (operational pressure): no plastic deformation must be induced;
2. 6 bar (safety factor=3): no plastic deformation must be indiced to modify the hydraulic behaviour.
3. 20 bar: no leaks must be induced in the circuit.

Moreover, every plate has undergone a temperature cycling test (from -30 $^{\circ}C$ to 20 $^{\circ}C$ for thirty times and with twelve assembling/disassembling phases.

3.1.5. The assembled panel

The assembled Thermal Screen Panel is obtained by the simple superposition of the cold plate, the insulating layer and the heating foils. Two heating foils will be used to cover the panel surface, in order to allow for an easy and quick adaptation to different boundary conditions in

different zones of the panel. In addition, thin U-shaped glass fibre profiles are added to the sides for a safe and correct sliding of the panel onto the U-profiles attached.

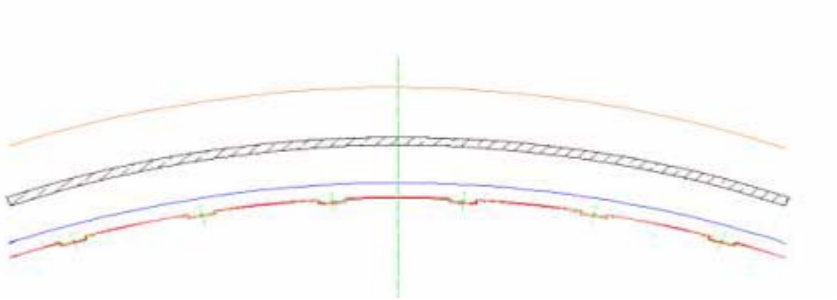


Figure 3-7 Section view of one panel

3.2. Cooling

The stringent reliability requirements of the Thermal Screen impose a conservative choice of the cold flow feeding system. For this reason, it has been decided to feed the cold plates by a double redundancy system. The main concept is that of using two independent systems, each one feeding 16 of the 32 panels on 8 separate lines, and to always have adjacent panels fed by different systems. This provides a fine granularity and allows for an acceptable global thermal behaviour of the screen also in case of failure of one feeding system. The standard Tracker cooling pipes (12 mm) are also used to feed the Thermal Screen panels, with one independent pipe servicing two cold plates.

The design assumptions are C_6F_{14} as cooling fluid and a flow speed in the feeding pipes of about 1 m/s. Under these conditions, the expected distributed pressure drops in the feeding pipes are expected to be not higher than 0.02 bar/m, i.e. a total pressure drop of 1 bar over the whole 50 m length of the pipes. The total required pressure is of 4 bar, for a volume of liquid inside the screen of 32 l (volume in supply and return pipes: 144 l). The flow per loop is of 0.11 l/sec., the pressure drop in each panel is 1,5 - 2 bar, and the power dissipation in the whole screen is typically 5 kW.

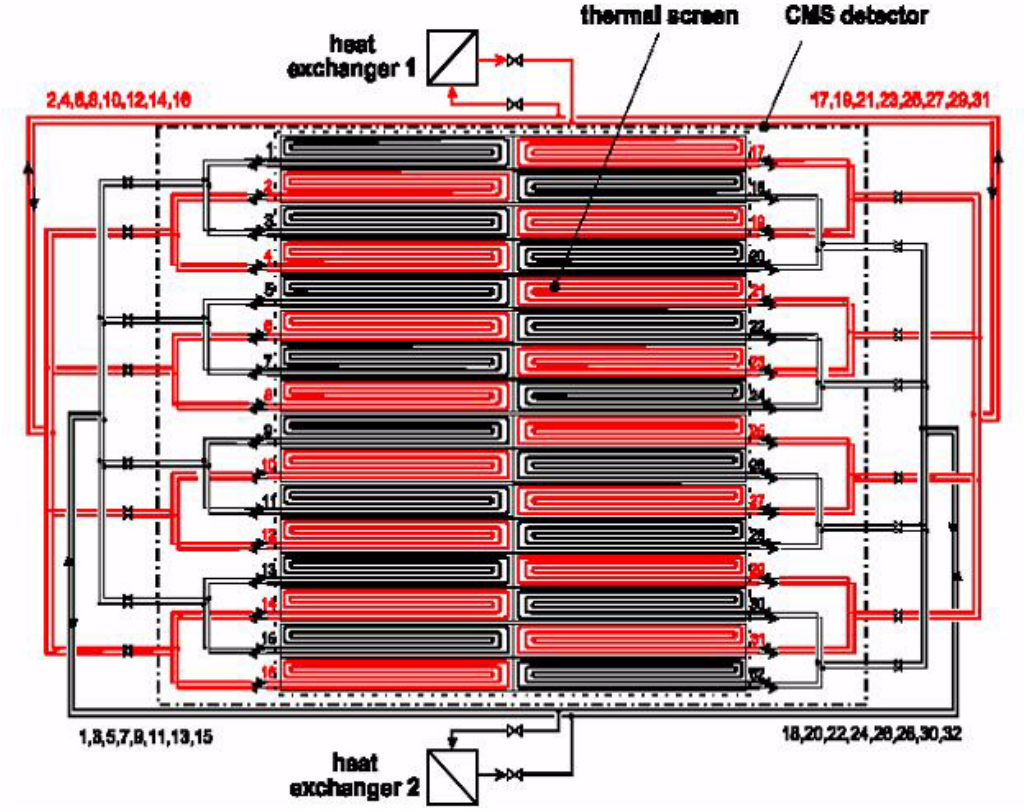


Figure 3-8 Thermal screen cooling plant

3.3. Heating

The cold plate of each Thermal Screen panel is in contact with layers of services of the End-Cap and of the Barrel detectors for an approximate length of 1.5 m. The rest of the cold plate faces the volume of the Barrel detectors, where no routing of packed services is foreseen. During the operation of the Tracker, the power dissipated by the low voltage cables of the detectors, only partially taken by the side-running cooling pipes, will modify the inside boundary conditions of the interested part of the cold plate. To keep this into account, and to increase the system realiability, it has been decided to use five independent heating foils to cover the two regions of each panel.

Modulating the power provided to the heating foil covering the service region (that is, controlling each module separately) will automatically adapt the outside boundary conditions of the panel, thus keeping unchanged the resulting temperature. It has been decided to power the heating foils by 48 V lines, with a maximum calculated current intensity of 0.79 A. This will allow the system to be classified as Low Voltage application, avoiding special safety measures otherwise needed.

3.4. Thermal model

The thermal screen is conceived as composed by two cylindrical shells separated by a layer of insulating material, as shown in the central part of Figure 3-9. The internal shell is in thermal contact with the silicon detector volume (SIL), while the external shell is separated from the ECAL volume by the cylindrical support tube (CST). The ECAL volume, which is in contact with the CST, has to be maintained at a temperature of 18 °C, value assumed as reference temperature Text for the ECAL environment. The SIL volume has to be maintained at a temperature of -10 °C, value assumed as reference temperature Tint for SIL environment. A gap of about 2 mm, filled by a gas at atmospheric pressure, separates the thermal screen from the CST and from SIL. A gaseous gap of about 5 mm separates the CST from the ECAL. Both the internal and the external shells are constituted by aluminium plates with channels for the circulation of the working fluids respectively in vaporization (cold side), and in condensation (warm side).

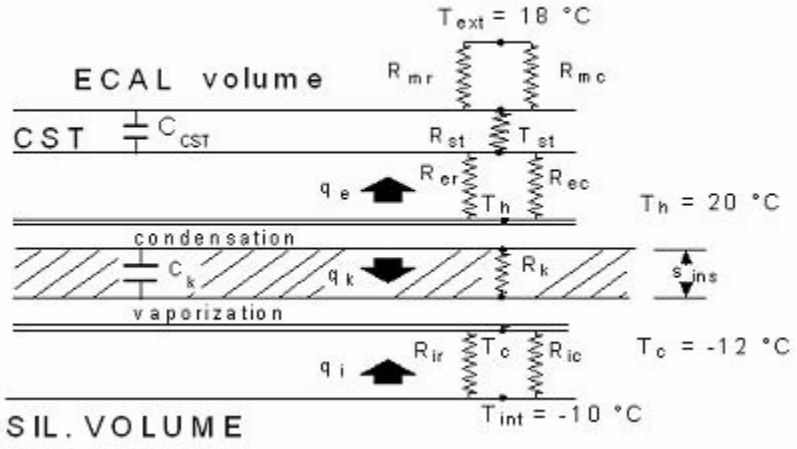


Figure 3-9 Thermal model

The thickness of each plate is of the order of 2 mm, while the layer of the insulating material has a thickness ranging between 4 and 6 mm. The thermal conductivity of this material is assumed to vary in the range 0.02-0.04 W/mK, depending on the available material.

The thickness of the screen in the radial direction could be thus maintained in about 10 mm. As a consequence of the large ratio between the radius of curvature and the Thermal Screen thickness, a plane geometry is considered for this model of the screen. The temperature of the working fluid in evaporation is –1 2 °C (Tc), while the condensation is assumed to occur at 20 °C (Th).

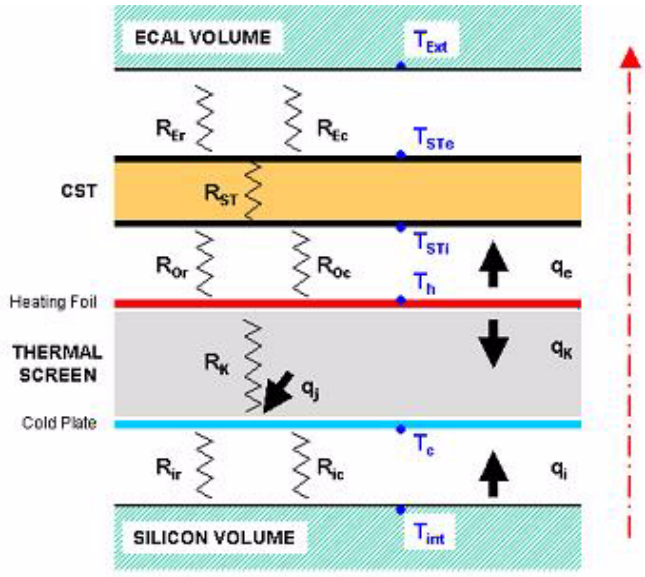


Figure 3-10 Thermal model II

3.5. Thermal resistances and heat fluxes evaluation

The thermal resistances controlling the heat transfers between the warm and the cold sides of the screen are indicated in Figure 3-10:

* REr e REc are the radiative and convective resistances that regulate the thermal interactions between CST and ECAL.

* RST is the thermal resistance of the CST structure (honeycomb in between two carbon fiber layers).

* ROr e Rir are the radiative resistances pertaining respectively to the thermal screen's external and internal sides, and depend on the surface emissivity of CST, of the thermal screen itself and of the tracker internal volume.

* ROc e Ric are, in the same volumes, resistances due to convection phenomena (because of the nitrogen filling the gaps).

* RK is the conductive resistance inside the insulating material.

* CK and Cst are the thermal capacitances related to the insulating layer and the support tube.

The equivalent thermal resistances Rext, Rint and Rm due to radiation/conduction acting respectively on THS warm side, THS cold side and on CST-ECAL are influenced by the surface emissivity of THS (within the range 0.1 ÷ 0.5) and CST (reasonably close to 1), their values span in the following ranges:

$$Rext = (1/Rec + 1/Rer) - 1 = [0.056 \div 0.074\ [m^2\,K/W] \qquad \text{Eq. (3.1)}$$

$$Rint = (1/Ric + 1/Rir) - 1 = 0.056 \div 0.074 \ [m^2 K/W] \quad \text{Eq. (3.2)}$$

$$Rm = (1/Rmc + 1/Rmr) - 1 = 0.1 \ [m^2 K/W] \quad \text{Eq. (3.3)}$$

The smallest values pertain to the maximum THS emissivities.

The conductive resistance of the insulating material depends upon the thermal conductivity of the available material. In the range of the expected values ($0.02 < kins < 0.035$) it results:

$$Rk = sins/kins = 0.14 \div 0.25 \ [m^2 K/W] \quad \text{Eq. (3.4)}$$

The CST thermal resistance may be evaluated by the knowledge of the honeycomb apparent density (90 kg/m3) and of the carbon fiber conductivity (~1 W/mK):

$Rst = 0.0066 \ [m^2 K/W]$

As for the thermal capacitance, we calculated it both for the Rohacell layer and the CST, using their specifications:

$$C_{CST} = 7877.44 \ J/m^3 K$$

$$C_{Rohacell} = 2808 \ J/m^3 K$$

(Carbon Fiber specific heat @ 20 C: 0.17 Cal/g °C
Nomex specific heat @ 25 C: 0.30Cal/g °C)

3.6. Heat fluxes evaluation

Referring to Figure 3-10, the warm side of the screen may provide two heat fluxes: qk towards the cold side through the insulating material, and qe towards the CST through the thermal resistances Rext, Rst and Rm.

The two heat fluxes may be expressed as follow:

$$qk = (Th - Tc)/ Rk = 228 \ W/m^2 \quad \text{Eq. (3.5)}$$

$$qe = (Th - Text)/(Rm + Rst + Rext) = 12.3 \ W/m^2 \quad \text{Eq. (3.6)}$$

The values of these thermal fluxes as well as of the others involved in this study are referred to Thermal screen unit area.

The total heat flux that must be available during the condensation of the working fluid at -20 °C is: $qc = qk + qe = 246 \sim 240 \ W/m^2$.

The cold side of the screen, where the working fluid vaporizes at –12°C, is the heat sink of other two heat fluxes, qk already evaluated and qi pertaining to the SIL volume. This flux, according to the decreasing temperatures, is directed towards the vaporization channel. In this

analysis is not considered the heat flux produced from the electrical cables present in the gap between the tracker volume and thermal screen owing to the incertitude of its estimation.

The vaporization zone is thus interested by the following thermal fluxes: qv = qi +qk. The heat flux qi depend on the temperature difference (Tint – Tc), and the thermal resistance Rint; considering Rint= 0.053 m^2 K/W, we have: qi = (Tint – Tc)/ Rint= 35.7 W/m^2 .

The total heat flux qv is then equal to: qv = qi +qk = 264 W/m^2 .

The vaporization and the condensation channels (cold and warm sides of the screen) are interested by the following heat fluxes: qv =qi +qk= 264 W/m^2 qc =qe +qk= 240 W/m^2 .

4 The thermal screen control system

4.1. A systems engineering approach

For centuries, people have been designing and building systems; large and small, simple and complex. They have developed techniques to co-ordinate and execute their development activities. Modern systems are more complex than ever before, and have forced a quantum leap in the sophistication of these techniques.

Originally driven by the needs of aerospace and defence projects, the culmination of these techniques is called Systems Engineering (SE) [56]. Systems engineering work begins by identifying the needs of stakeholders, users, and operators, and transforms these needs into a responsive, operational system design and architecture. This deliverable system must conform to the demands of the marketplace as well as to the initial set of functional and non-functional requirements.

The International Council on Systems Engineering (INCOSE) official definition of SE states that "Systems Engineering is an interdisciplinary approach and the techniques to enable the realization of successful systems" [56]. Put more usefully, SE is an organized approach to problem solving, by experienced engineers and physicists with a broad, system-wide overview towards solving the problem, weighing options and evaluating risks and constraints. SE is responsible for integrating all the technical backgrounds, subject matter experts and speciality groups into a team development effort. It starts with defining customer needs and required functionality early in the life cycle, managing requirements, and proceeding into design synthesis and system verification and validation.

An effective systems engineering approach must perform, at a minimum, the following activities to produce an optimal system:

* Accurately assess the available information and find what is missing
* Identify performance or effectiveness measures that define success or failure
* Manage and analyse all source requirements that depict user needs
* Conduct systems analysis to formulate a behavioral design that meets all functional and performance requirements
* Allocate functional behavior to the appropriate physical architecture
* Perform trade-off analysis to support decision making concerning alternative designs or architectures
* Create executable models to verify and validate system operation

The project should develop a Systems Engineering Management Plan (SEMP) which defines and captures those project management processes that address planning and organization, monitor and control, and project assessment. With enough detail in the SEMP, team members are provided with an effective model for improving communications and exchanging information.

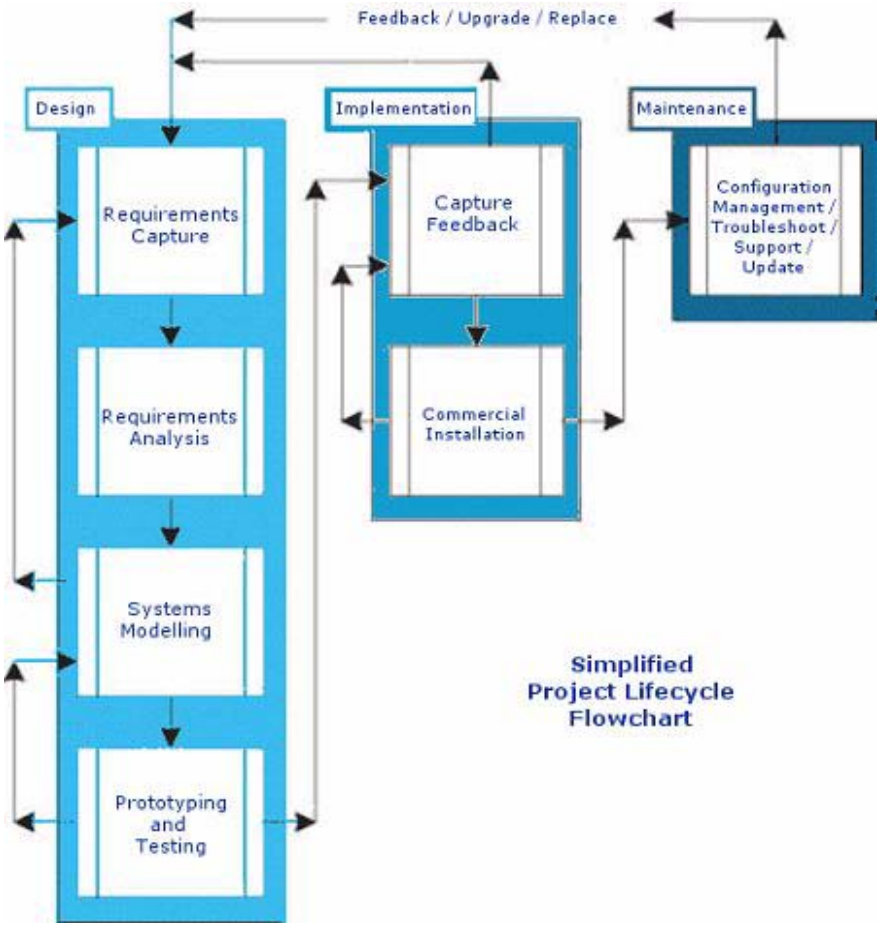


Figure 4-1 A systems Engineering process

A systems engineering process is a formalized, planned, reviewed, and continually evolving series of activities which:

- Involves multiple engineering disciplines working co-operatively together
- Involves the customer as intimately as possible in all phases
- Needs sophisticated communication of desires, expectations, needs, and engineering data
- Needs formalized notations for the individual problem domains within the project and, particularly, the interfaces between problem domains and between engineering disciplines

- Allows the engineering of a need into a reality
- Can be thought of as the engineering of a requirements set into a finished system

The activities in the systems engineering process are focused on four major technical phases:

- Requirements management
- System behavioral analysis
- Architecture analysis
- System verification and validation.

The first phase in such a process translates end-user, marketing, or customer source statements into formal requirements. In principle this is confined to a requirements capture exercise at the start of a project. In reality, it occurs throughout the lifecycle, drawing requirements from many sources. Typical requirements analysis tasks include:

- Identify source material
- Identify stakeholder needs
- Identify initial set of requirements (top-level functional, non-functional, performance and interface requirements)
- Establish design constraints
- Define effectiveness measures
- Capture issues/risks/decisions.

The main objectives of the next phase, the system functional or behavioural analysis, are to:

- Describe the problem defined by the requirements analysis in clearer detail
- Identify and describe the desired functional behaviour of each system element or process.

This functional analysis is typically performed without consideration of a specific design solution. Key tasks of the functional analysis phase include:

- Define operational scenarios
- Derive system behavior model (and other models as needed) to reflect control and function sequencing, data flow and input/output definition
- Derive functional and performance requirements, and allocate to behaviour model.
- The physical architecture analysis phase performs system synthesis by assigning functions to identified physical architecture elements (structures, subsystems, and components).

This synthesis activity is performed for the purpose of defining design solutions and identifying subsystems and components that will satisfy the established requirements. Candidate tasks include:

- Group and allocate functional behaviour among the system's component parts
- Define functional interfaces between system component parts
- Identify physical interfaces
- Ensure that performance requirements are preserved
- Allocate non-functional requirements
- Assess failure modes, effects and criticality
- Identify make-or-buy alternatives.

The technical evaluation phase of systems engineering demonstrates that the system performs as needed and as specified.

The initial task is a comprehensive verification and validation analysis, which may include static and dynamic logic analyses, functional performance assessments, system simulation, hardware-in-the-loop testing, and so on. Other key tasks include:

- Test planning
- Compliance and cost assessment
- Best design solution selection
- Automatic report and specification generation.

The nucleus for any sound systems engineering process is the Systems Engineering Management Plan (SEMP). This plan will describe the systems engineering strategy to be used for project planning, controlling and execution. The SEMP will identify all project deliverables to be produced, and the individual systems engineering tasks necessary to produce each of them. To accurately reflect an accepted systems engineering process, the SEMP follows current industry standards and methodologies.

These include IEEE 1220-1998, Standard for Application and Management of the Systems Engineering Process; Electronic Industries Alliance (EIA) ANSI/EIA-632-98, Processes for Engineering a System; Systems Engineering Capability Maturity Model (SE-CMM), Carnegie Mellon Software Engineering Institute; and the soon to be released International Standards Organization (ISO) 15228, Systems Engineering Life Cycle Management.

4.2. Risk management

All projects involve risk; the risk that something will go wrong [64].

This is not necessarily a bad thing, as no progress is made without taking some risk; however, there is a difference between unmanaged risk, e.g., gambling, and managed risk where the probabilities are well understood, and contingencies made. Risk is only a bad thing if risks are ignored and they become problems.

Risk management consists in assessing which risks are most likely to apply to the project, deciding a course of action if they become problems, and monitoring projects to give early warnings of risks becoming problems.

A working efinition of risk can be "A measure of the uncertainty of attaining a goal, objective, or requirement pertaining to technical performance, cost, and schedule", and it is always present in the life cycle of a system [64]. The system may be intended for technical accomplishments near the limits of the state of the art, creating technical risk. System development may be rushed to deploy the system as soon as possible to meet an imminent threat, leading to schedule risk.

Risk has two components - the likelihood that an undesirable event will occur and the consequence of the event if it does occur. The likelihood that an undesirable event will occur is often expressed as a probability. The consequence of the event is expressed in terms which depend on the nature of the event (e.g., increased costs, performance loss, delayed delivery).

Air transport provides two examples of events and their consequences - the event of arriving at the destination 15 minutes late usually has benign consequences, while the event of an airplane crash has harsh consequences and possible loss and injury. Most people would judge both of these events to have low risk; the likelihood of arriving 15 minutes late is high but the consequences are not serious. On the other hand, the consequences of a crash are very serious but are offset by the low likelihood that such an event will occur (which is why airlines still have customers).

Risk management, in the context of systems engineering, is the recognition, assessment, and control of uncertainties that may result in schedule delays, cost overruns, performance problems, adverse environmental impacts, or other undesired consequences [6].

There are two main branches of risk management:

* Program risk management: the management of technical risks and task performance uncertainties associated with systems engineering and development programs, in order to meet performance, cost, and schedule objectives.

* Environmental risk management: the management of environmental, health and safety risks associated with the production, operation and disposal of systems, in order to minimize adverse impacts and assure sustainability of these systems.

The functions of a risk management program are to:

1. Identify the potential sources of risk and identify risk drivers.

2. Quantify risks, including risk levels, and assess their impacts on cost (including life-cycle costs), schedule, and performance.

3. Determine the sensitivity of these risks to program, product, and process assumptions, and the degree of correlation among the risks.

4. Determine and evaluate alternative approaches to mitigate moderate and high risks.

5. Take actions to avoid, control, assume, or transfer each risk.

6. Ensure that risk is factored into decisions on selection of specification requirements and design and solution alternatives.

The objective of risk management is to ensure the timely delivery of a system and its associated processes that meet the customer's need.

Risk management involves five processes - planning, identification, assessment, analysis, and mitigation.

* Risk planning is the process of deciding (in advance) how risk will be managed, including the specification of the risk management process and organizational responsibilities.

* Risk identification is the process of recognizing potential risks and their root causes as early as possible, and setting priorities for more detailed risk assessment.

* Risk assessment is the process of characterizing or quantifying those risks which merit attention.

* Risk analysis is the process of evaluating alternatives for handling the assessed risks. This includes performing "what if" studies.

* Risk handling, finally, is the process of dealing with a risk by choosing a specific course of action. Risk can be mitigated by choosing to avoid the risk (perhaps with a change in the design), to control the risk (perhaps with additional development resources), to assume the risk (expecting that the event can be overcome in normal development), or to transfer the risk (for example, with special contract provisions).

4.3. Safety systems

The term "Safety Instrumented Function" or SIF is becoming common in the world of Safety Instrumented Systems (SIS).

The definition of a SIF as provided in IEC standard 61511 "Functional safety: Safety Instrumented Systems for the process industry sector" [19] describes a safety instrumented function as a "safety function with a specified safety integrity level which is necessary to achieve functional safety. A safety instrumented function is further defined in 61511 as a "function to be implemented by a SIS, other technology safety-related system, or external risk reduction facilities, which is intended to achieve or maintain a safe state for the process, with respect to a specific hazardous event."

The standard 61511, however, uses the terms SIS and SIF somewhat interchangeably in places.

From this definition we can also see that there are two types of safety instrumented functions. The first is a safety instrumented protection function, which is a safety instrumented function operating in the "on demand" mode. The second is a safety instrumented control function, which is a safety instrument function operating in the continuous mode.

Two examples mucht be an automatic fire sxtinguisher system, which (hopefully) only operates in the event of fire, and the smoke detector controlling it, which (hopefully) scans the air continuously looking for the presence of smoke.

From [21] we might define the SIF as an identified safety function that provides a defined level of risk reduction or safety integrity level (SIL) for a specific hazard by automatic action using instrumentation.

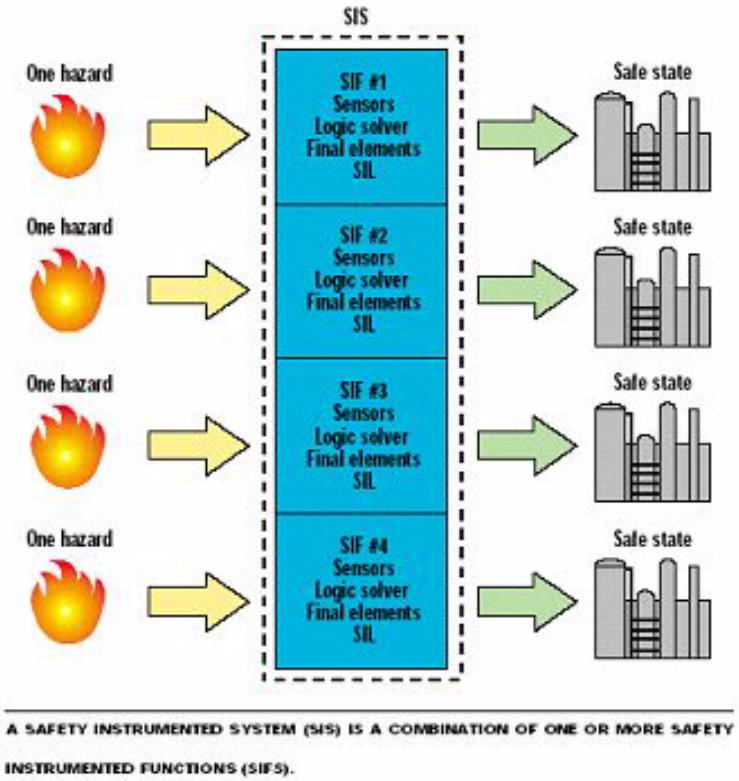


Figure 4-2 Safety Instrument Functions

A SIF is made up of sensors, logic solver, and final elements that act in concert to detect a hazard and bring the process to a safe state. Some examples of SIFs are:

• High pressure in a vessel opens a vent valve: The specific hazard is overpressure of the vessel. The high pressure is detected by a pressure-sensing instrument, and logic (PLC, relay, hardwired, etc.) opens a vent valve, bringing the system to a safe state.

• High temperature in a furnace that can cause tube rupture shuts off firing to furnace: The specific hazard is tube rupture. Instrumentation automatically causes a main fuel trip that removes the heat, bringing the system to a safe state.

• Flame-out in an incinerator that can lead to a release of toxic gas causes process gas feed to be shut off: the specific hazard is a flame-out. The automatic instrument protective action is to close process gas feed to the incinerator, which stops any toxic gas release bringing the system to a safe state.

• Flame-out in an incinerator that could cause fuel gas accumulation and explosion causes a main fuel gas trip: the specific hazard is a flame-out. The automatic instrument protection action is a main fuel gas trip, which cuts off the fuel and prevents fuel gas accumulation, bringing the system to a safe state.

There are functions that may seem like a SIF or part of a SIF, but are not. A SIF is normally associated with life-and limb protection. If you have identified an instrumented protection function and the consequence of the hazard could be killing or injuring, the function is a potential SIF (pending SIL analysis—there may be adequate layers of protection so that identification as a SIF is not required).

However, when a SIF operates, there may be related actions that occur at the same time that place portions of the process in desirable operating states to minimize startup time, loss of inventory, process equipment problems, etc. Operating companies sometimes fall into the trap of considering these related actions as part of the SIF and can increase the difficulty of achieving the target SIL.

This can lead to increased and unnecessary cost, burden, and complexity. Equipment or asset protection functions also are not SIFs. Every plant has protective functions that protect the plant's equipment and assets. This is primarily a commercial or money issue. If there are no safety aspects to these protective functions, they are not SIFs. But since there are few to no standards in this area, some people do assign an asset integrity level (AIL) to these protection functions and treat these systems like safety instrumented systems.

For example, if high-high level in a knockout drum to a compressor shuts it down to protect it from mechanical damage due to liquids, and there is no anticipated safety issue (such as rupture of the compressor case), then this is not a SIF but rather an equipment protection function.

A key to SIL selection is to correctly identify the safety instrumented functions for a facility. Failure to identify true SIFs leads to less safety. Conversely, identifying things as SIFs that are not leads to unnecessary cost, burden, and complexity. By definition, each SIF must have a SIL based on how much risk reduction the SIF must provide to help reduce the risk of a particular hazard to an acceptable level when considered with the rest of the protective layers that reduce the risk of that particular hazard. The SIL is selected based on the risk posed by the hazard the SIF is protecting against.

This risk is composed of a consequence (what bad things that can happen) and a pre-safeguard frequency (how often the hazard is expected to occur if no protections—SIS or non-SIS— are provided). However, while there is a single hazard (and generally a single

consequence) associated with a SIF, you can have multiple initiating causes, each with its own frequency of occurrence.

For example, overpressure of a vessel due to loss of cooling (with a consequence of vessel rupture and fire/explosion) could be caused by loss of cooling water supply, loss of cooling water pump(s), temperature control loop failure, plugging of tubes, etc. Each of these initiating causes can have a different frequency of occurrence, and thus different risks (consequence x frequency) for the same SIF.

4.4. A common design methodology

The Tracker control system main scope is producing a highly safe and recyclable control system capable to interlock the Low Voltage-High Voltage power lines in case of faults related to humidity and/or temperature raising, as well as of other nature. This system also includes the thermal screen, which is the ultimate resource in order to protect the detector from temperature raises in case of a failure of the main cooling system.

Therefore, not only the thermal screen control system must keep the temperature constant, but must provide inputs to the Tracker interlocks.

For this reason, it was envisaged to use a common methodology to design, build, commission and test the thermal screen control system (which is more oriented towards the regulation) and the tracker interlocks [43].

Another reason is dictated by the fact that the tracker community needs highly uniform control systems to be used and debugged in the Institute home labs before the final installation in the CMS cavern. The same systems will accompany the detectors during their pre-installation phase at CERN and, eventually, the very same systems will be used in the final control system.

As a first classification, every process can be split into three layers, as already shown in Chapter 2:

- Process Layer: Temperature sensors (RTD, Thermistors); humidity sensors; flow meters; pressure meters; heaters.
- Control layer: Siemens PLCs (3 families: S7- 200, 300, 400); fieldbus (if needed).
- Supervision layer: Siemens native SCADA Suites (web servers available).

The Tracker control system has also interfaces with other control systems, namely Cooling and Ventilation (CV), Detector Safety System (DSS) and LHC Safety System:

The methodology hereby proposed (which has eventually been adopted) is rigorous (because all the steps are clearly stated) and flexible at the same time (one can use these criteria for no matter what process control).

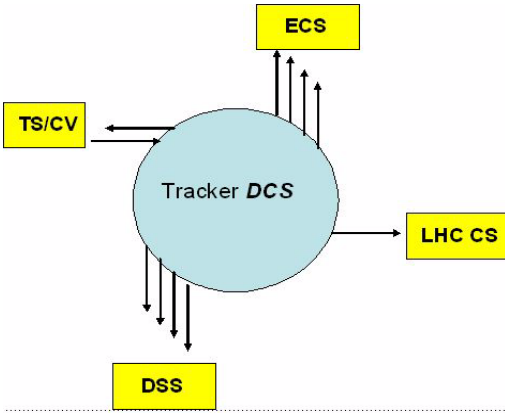


Figure 4-3 Tracker DCS interactions

4.5. Performance specifications

There are two fundamental specifications for the thermal screen: safety and reliability. These ultimate requirements can be further broken down to derive a control-system specification that quantifies the target to be met by controller-design:

* Safety: The safety requirement translates rather straightforwardly into a constraint on the level of temperature inside the tracker volume. Clearly, the tracker must never be in danger of getting warm or to reach condensation level, because either case would result in a drop of performance and in a probability of short circuits, with disastrous consequences.

* Reliability: The system needs to be operated in a completely autonomous fashion, being not accessible to human intervention during operations. Aspects which contribute to this requirement include the following:

* Control quality: It turns out that not only the screen has to keep the tracker cold, but the Ecal environment has not to be cooled down.The controller will have to take into account both the requirements of the two detectors, managing to satisfy them at the same time. Thus, control quality demands steady control of the temperature on the two sides.

* Maintenance: During maintenance periods, the controller can be operated manually, or can be simply be turned off.

4.6. Overview

Each panel (see Figure 4-2) has on one side Capton heating foils with diffused resistors on a kapton, and on the other a radiator for coolant flow. Between the two sides there is an insulating layer. These panels are installed on the inner surface of the CST, and each panel has its own power supply which is controlled by the PLC. Each panel also has 5 temperature sensors on the heating foil side, and these can be fed as inputs for the controlling PLC.The heating foils are the active component of the thermal screen system, and the temperature is maintained by the PLC by controlling the power dissipated in the heating foils of each panel. The coolant flow on the inner surface of the panels is relatively constant, and the flow rate is not actively controllerd; furthermore, sixteen of the panels are supplied from one of the cooling plants, and the other sixteen from the second cooling plant. This is done to ensure that, should one cooling plant fail, the remaining panels may maintain the temperature levels for at least a short period, and hopefully long enough to restart the problematic plant.

The following table shows the control objectives and the related process variables and the sensors used to monitor them.

control objective	process variable	sensor
1) Safety	Maintain Tracker and Ecal isolated	RTD (Pt1000, four wires)
2) Environmental protection	Prevent tracker heating	RTD (Pt1000, four wires)
3) Equipment protection	Ensure that voltage is applied to heating foils	Voltage input to PLC
4) Smooth plant operation and production rate	When possible, make slow adjustment to voltage	Voltage input to PLC
5) Product quality	Maintain the temperature gradient below 2.5 degrees	RTD (Pt1000, four wires); Hygrometrix RH sensor 411 series.
6) Profit optimization	Maximize the use of heating foils	Voltage input to PLC RTD (Pt1000, four wires)
7) Monitoring and diagnosis	Provide alarms and warnings	RTD (Pt1000, four wires); Hygrometrix RH sensor 411 series.

In figure 4-4 the process is shown with its inputs/outputs. MV is the Manipulated Variable (the voltage applied to the heating foils); DV is a Disturbance Variable: it cannot be controllable, and comes both from the Tracker side (variations in the inside temperature) or the ECAL side (temperature oscllations external to the thermal screen). The ouput is the Controlled Varible CV, namely the temperature facing ECAL on the Tracker Support Tube.

Figure 4-5 shows one panel as it looks like in the ISA symbols representation, along with PLC, cooling plant and power supply:

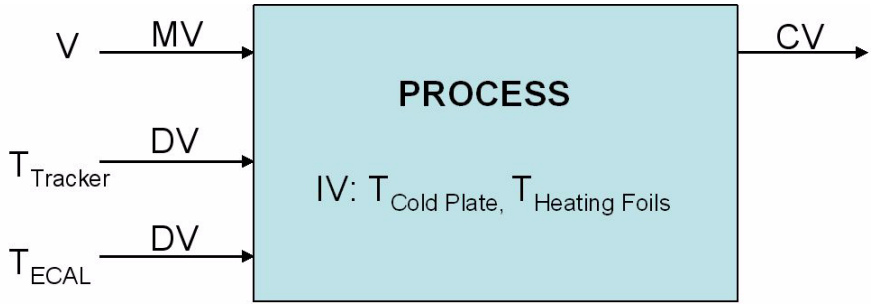


Figure 4-4 Thermal screen process representation

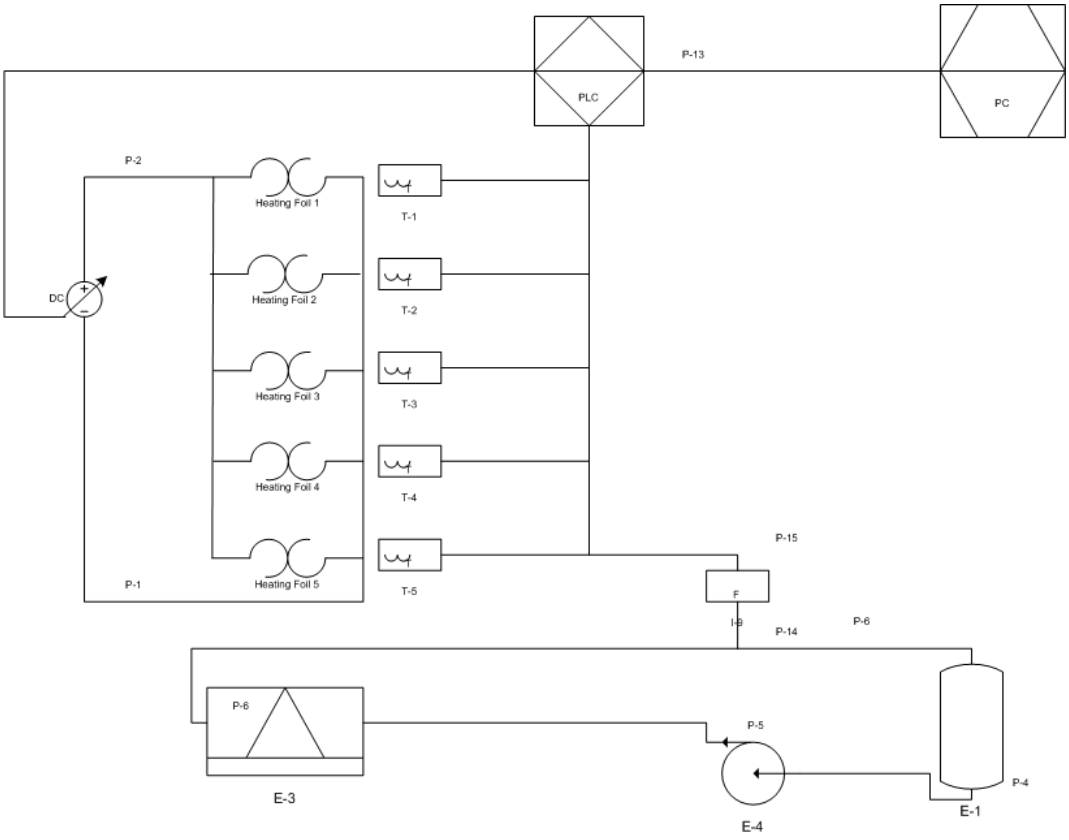


Figure 4-5 ISA drawing of a single panel circuit

4.7. System modeling

Traditionally, in electro-thermal modelling, the RC equivalent of a material sample is calculated using the equivalence between the fundamental laws of heat and electricity, such that heat transfer problems can be associated with electrical circuits [33]. Consider a heat-conducting solid, the ends of which are maintained at different temperatures T1 and T2. The Fourier law relates the heat flow to the gradient of the temperature, q= -*k* grad(*T*) J/m2-sec.,

where q is the heat flow per second per unit area and k is the thermal conductivity. The form of this law is identical with ohm's law for the flow of electric charge, which relates the current density with the gradient of electrical potential. For a sample of length L (m) and cross sectional area A (m2) normal to the direction of heat flow Q (J/sec.), Eq. (10) can be expanded as Q/A = k(T1 - T2)/L.

Q = (kA/L)(T1 - T2)

supplied to a system raises the internal energy, according to the relation Ct(dT/dt) = Q (J/sec.), where Ct is the thermal capacitance (J/deg), and can be written directly as

Ct = MCp (M being the mass of the block (kg), and Cp is the specific heat at constant pressure (J/deg-kg).

The thermal, electrical equivalent values of the resistance (R) and capacitance (C) of mild steel sample (Fig.2(a)) are calculated using Eqs. (12) and (13),

R = L /(k*A) (14)

$C = \rho \cdot c \cdot L \cdot A$ (15)

ρ being the density of the material (kg/m3).

Table 4-1.

Thermal Quantity	**Electrical Quantity**
Heat Flow Q	Current Source i
Temperature Difference θ	Voltage u
Thermal Resistance R_θ	Resistance R
Thermal Capacitance C_θ	Capacitance C

asd

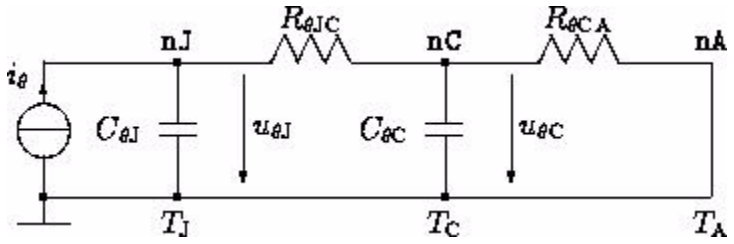


Figure 4-6 Thermal network describing heat conduction from junction temperature through the case to ambient temperature. Current source equals the power dissipation.

Heat conduction in an electrical component is usually modeled using a ladder circuit shown in Fig. 4-6. However, the ladder topology is applicable also in modeling any kind of heat transfer. In Fig. 4-6, the current source i_θ describes the power dissipation in the junction, and node voltages u_θ represent the differences between the actual and the ambient temperature.

Current flows from junction temperature Tj through thermal resistances $R_{\theta JC}$ and $R_{\theta CAA}$ and to the ambient T_A. $C_{\theta J}$ and $C_{\theta C}$ are the thermal capacitances of the junction and case, respectively [25].

In the thermal screen case, the temperature inside the tracker volume has been considered as a constant voltage source, while the power generated by the heating foils is a variable voltage source [54].

Therefore, the following is the electrical network representation of the system:

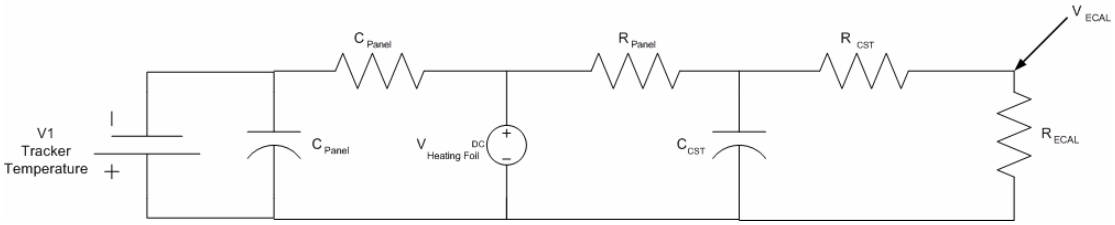


Figure 4-7 Electro-thermal model

The circuit is resolved to find the voltage V_{ECAL} as a function of the current supplied to the heating foils; it results in the transfer function:

$$G(S) = \frac{V_{Ext}}{I_{Heat}} = \frac{-0.236}{109.794 \cdot S + 0.4847}$$

The system is a first order without delay. It can be studied in the time and frequency domain modeling the whole control chain in the MatLab Simulink environment:

Here, the system is fully represented along with the PID controller in parallel mode; the preliminary study does not include the PID effect, which will be taken into account at the tuning phase.

In figure 4-8 the Simulink diagram of the control system is shown, along with the MatLab Optimization Toolbox module.

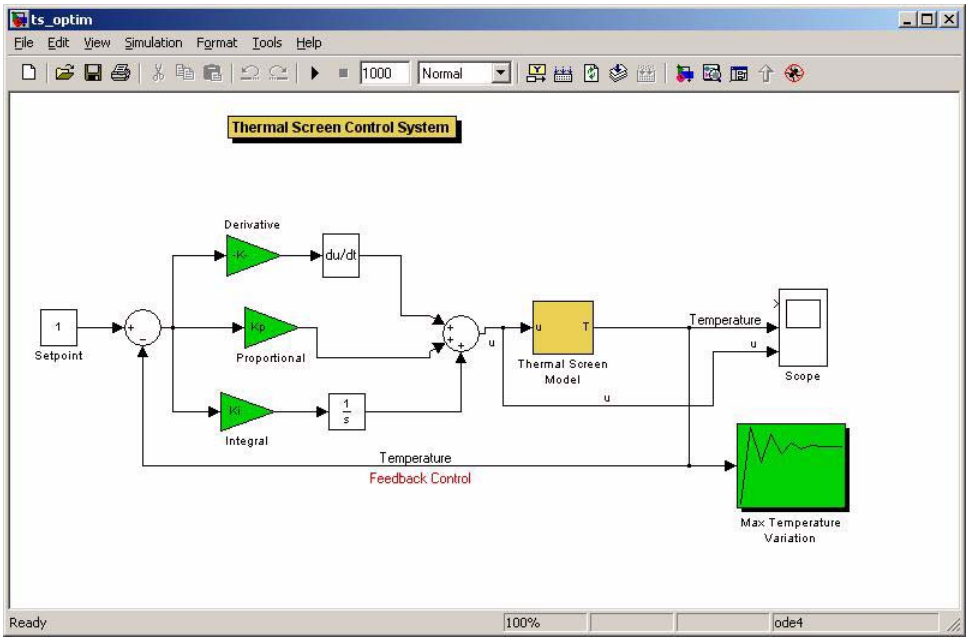


Figure 4-8 Simulink diagram

In the following diagrams (Figure 4-9 and 4-10) the Bode diagram and the step response are shown. The process has a time constant of 5000 sec.

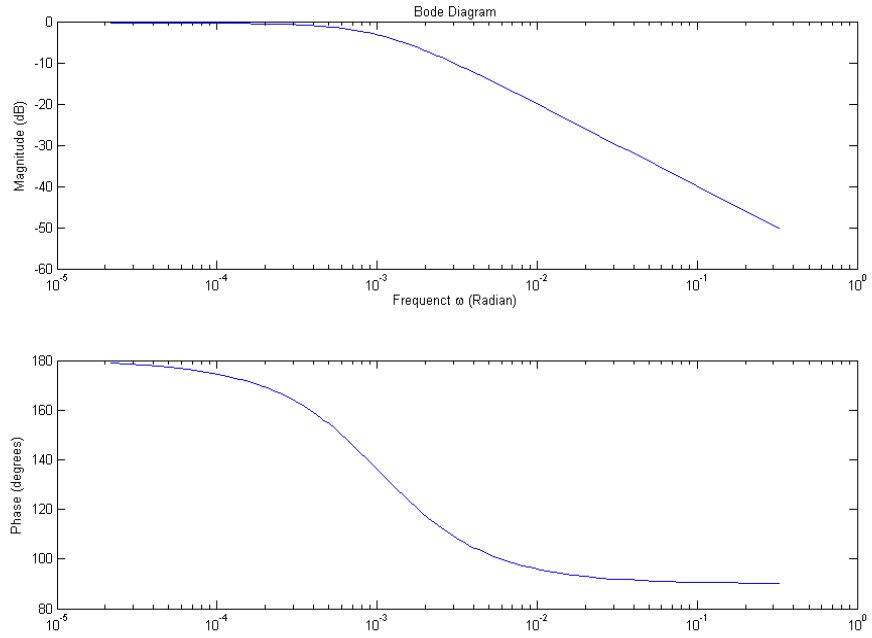


Figure 4-9 Bode diagram

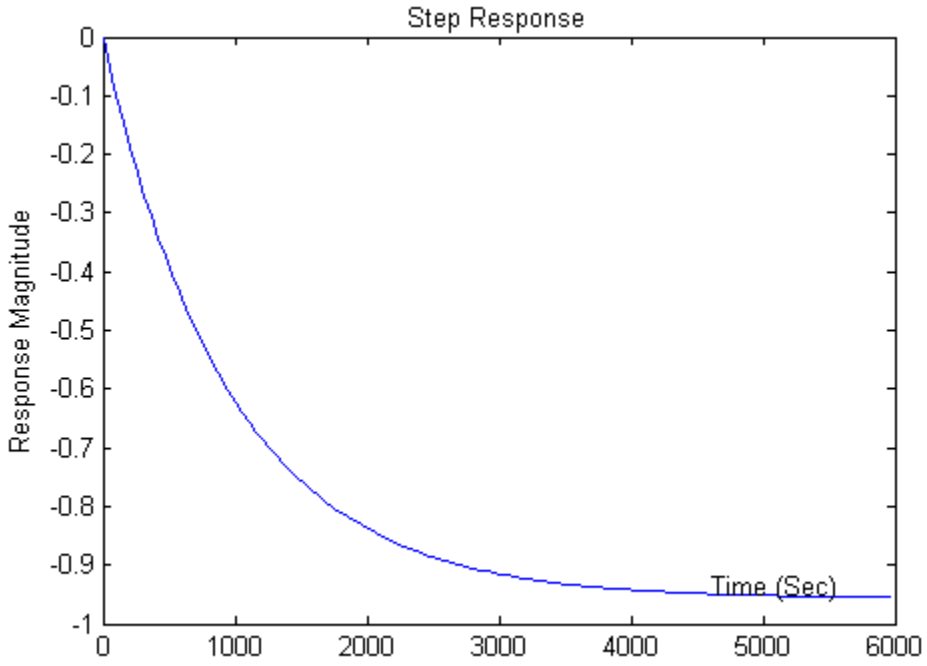


Figure 4-10 Step Response

4.8. Controller design

Matlab optimization toolbox provides an environment to assist in time domain based control design. The software allows for the following constraint selection: rise time, settling time, overshoot; the lower and upper constraint bounds define a channel between which the system's step response should fall, as in Fig. 4-11.

The routine adjusts the tunable variables in an attempt to better achieve the constraints on systems signal defined by the main interface. The optimization problem is solved using a sequential quadratic programming (SQP) algorithm and quasi-Newton gradient search technique.

For the following set of constraints:

Raise time= 60 sec., Settling time=50 sec., Overshoot=1.8

the tuning parameters will be:

P: 621.24

I: 206.35

D:0.

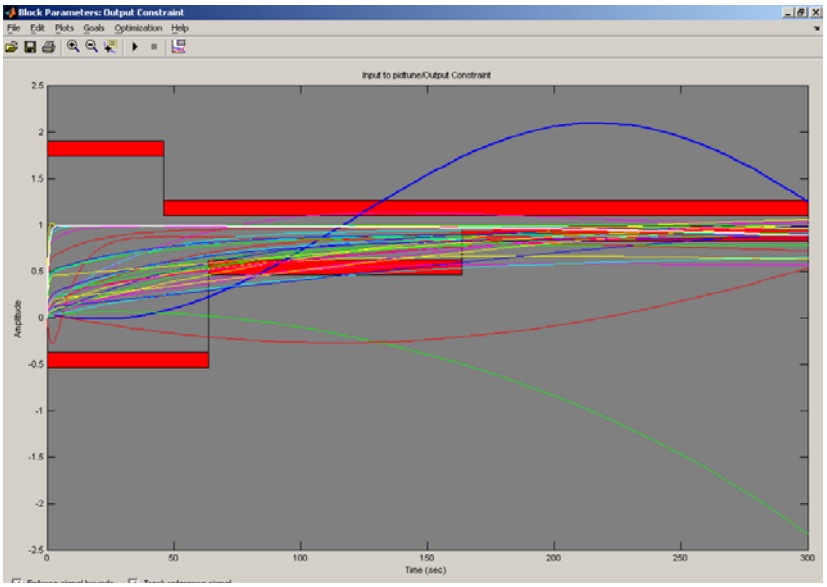


Figure 4-11 MatLab Optimization toolbox output

Another tuning method is based upon a modification of Skogestad's routines. One change is that the transfer functions are simplified based upon numerically fitting a function to the data that are generated from the original model. This way it is also possible to tune a controller if only data is available, but no transfer function representation has been determined.

Transfer function:

Eq. (4.1)

$$G = \frac{(0.236 \cdot s)}{109.794 \cdot s + 0.4847}$$

Retuning:

Desired closed-loop time constant: $\tau_c = 1$

Controller: $G_c = P + \frac{I}{s} + D \cdot s$

P = 12.90

I = 51.24

D=0

Eq. (4.2)

$$G_c = 12.904044 \cdot \left(1 + \frac{1}{0.251788 \cdot s}\right)$$

IAE = 1.12; ISE= 0.64; ITAE = 1.04

Gain margin = 17.45; Phase margin = 84.26

Gain crossover frequency = 15.62 rad/s; Phase crossover frequency = 0.92

4.9. The MIMO problem

Each thermal screen panel will have, in the final installation, three neighbours: two panel facing it on the long side, and one on the short side, as in Figure 4-12.

The interactions on the short side are considered less important than those on the long side;

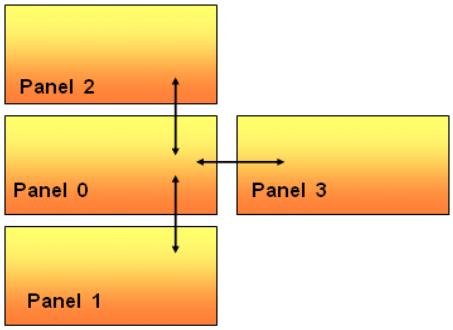


Figure 4-12 Panels mutual interactions

thus, they will be ignored. Yet, the two adjacent panels pose a clear problem of MIMO. In case of non negligible interaction, we will have to decouple the control loops.

In fact, the single-loop controllers are completely independent algorithms that do not communicate directly among themselves, while the manipulation made by one controller can influence other controlled variables.

In order to determine the degree of interaction of the panel, a model has been conceived, based on the fact that the two adjacent panels will not be in contact, but separated by a two mm gap of air inside the tracker.

Moreover, we kept studying the system from the thermo-electrical analogy point of view, thus taking into account the effect of a panel simply by adding a current generator and a resistor to the original circuit, as in Figure 4-13.

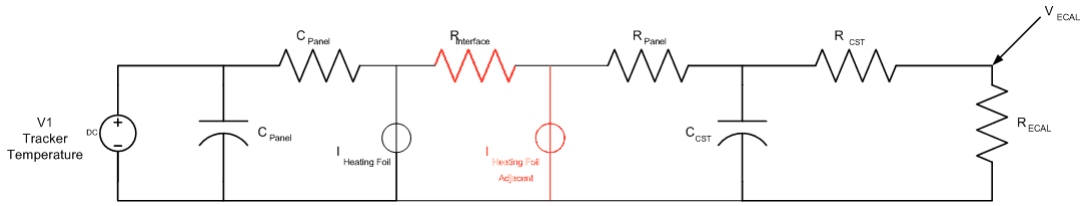


Figure 4-13 MIMO electro-thermal model

The additional current generator intensity can be calculated from the power dissipated by each panel (140 W) through a resistance of 61 Ω. THe resistance between the two panels is 0.1 Ω (same resistance as a thin layer of air).

Solving the system for V_{ECAL} with respect to $V_{HeatingFoil}$ we get the transfer function:

$$G_{01} = \frac{0.056}{1 + 0.00156 \cdot S}$$ Eq. (4.3)

which represents the interaction between the panel n. 0 and the adjacent panel n. 1.

If we call the voltage applied on the two panels 0 and 1, respectively M1 and M2 (manipulated variables) and the same for the controllers (C1 for the controller of the panel 1 loop and C2 for the controller of the panel 2 loop), we can say that there exist two possible controller pairings: M1 with C1, M2 with C2 or M1 with C2, M2 with C1.

This kind of representation provides two useful types of information: a measure of process interactions and recommendations about best pairing of controlled and manipulated variables.

Eq. (4.4)

$$C_1 = K_{11}MV_1 + K_{12}MV_2$$
$$C_2 = K_{21}MV_1 + K_{22}MV_2$$

The relative gain array relates the ith controlled variable and the jth manipulated variable:

Eq. (4.5)

$$RGA = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix}$$

where the relative gain itself, λ_{ij}, relates the ith controller and the fth manipulated variable:

Eq. (4.6)

$$\lambda_{ij} = \frac{\text{open - loop gain}}{\text{closed - loop gain}}$$

In this case, we have:

$$K = \begin{bmatrix} 0.0180 & -0.056 \\ -0.056 & 0.0180 \end{bmatrix}$$

Det (K)=-0.0028

$$(K^{-1})^T = \begin{bmatrix} -6.401 & -19.91 \\ -19.91 & -6.401 \end{bmatrix}$$

$$\lambda_{11} = \frac{1}{1 - \frac{K_{12}K_{21}}{K_{11}K_{22}}} = 0.243$$

$$RGA = \begin{bmatrix} \lambda_{11} & 1-\lambda_{11} \\ 1-\lambda_{11} & \lambda_{11} \end{bmatrix} = \begin{bmatrix} 0.243 & 0.756 \\ 0.756 & 0.243 \end{bmatrix}$$

Since $0 < \lambda < 1$, the system is controllable, nd we can conclude that mutual interaction between adjacent panels is negligible [Marlin].

These results allows for the design of independent, single-loop PID controllers without any specific decoupling, which will be implemented in a PLC.

4.9.1. PLC Programming

Programmable Logic Controllers provide a replacement for hard wired relay and timer logic circuits found in traditional control panels. The PLC offers much more flexibility in process control since its behaviour is based on executing simple programmed logical instructions. As a result installation is more straightforward and amendments are easier to implement [14]. Most modern PLCs offer internal functions such as timers, counters, shift registers and special functions making sophisticated control possible using even the most modest unit. The PLC offers standard input and output interfaces that will suit most process plant equipment and machinery. Standard input interfaces are available that permit direct connection to process transducers. Standard output interface circuitry will usually permit direct connection to contactors that energize process actuators such as pumps or valves.

The job of the PLC is to monitor inputs from the process under control and, based on the program being executed in program memory, energize the appropriate outputs. The control program, which determines the behavior of the PLC, can be easily modified thus permitting the entire operation of the external hardware to be altered without the need to disconnect or re-route a single wire.

In summary, PLC systems offer many advantages over hard wired systems and has specific features which suit them to the industrial control environment :

- Modular construction, allowing easy replacement and addition of units.
- Standard input/output connections.
- Ease of programming and re-programming when maintenance and/or enhancement become necessary.
- Noise immunity of inputs and outputs.

The rapid utilization by industry of PLCs has generated a demand for a range of PLC systems from the small system with only twenty digital input/output (I/O) ports and a program memory capable of holding 1000 steps or less to the large system with hundreds of I/O ports and memory capable of holding more that 10000 steps.

The image map below illustrates the basic structure of most PLC systems [52].

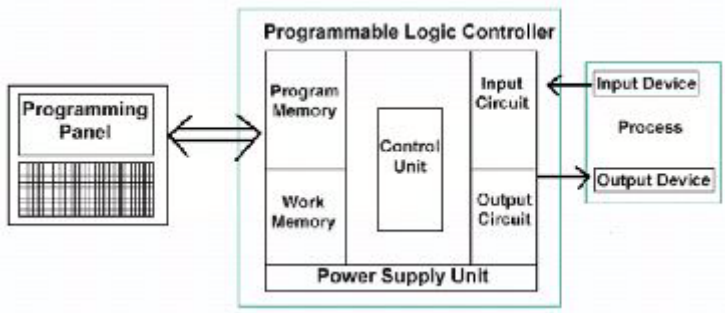


Figure 4-14 Block diagram of a typical PLC

We will now consider each of the elements in the above block diagram.

Control Unit

This section of the PLC is responsible for scanning repetitively all the input states, executing the program steps and setting or clearing the output or auxiliary relays in response to program logic and important states. The scan rate is determined by the microprocessor clock rate and the program size and is in the order of 1 to 10 ms per kilobyte of program. The control unit will also manage the timer/counters and any other special functions.

Program Memory

This memory will either be battery backed CMOS RAM during program development or some form of Read Only Memory (typically FLASH) that does not lose its contents when switched off. This is the area where the program steps are stored, and it is important that this area be non-volatile since in the event of a power failure it must retain its content. Typical backup battery lifetime is measured in years, typically about 5, and there will be a 'battery low' warning indicator either on the status panel of the PLC or as an auxiliary contact. Additionally, many systems offer some sort of backup capability, usually under the form of Compact Flash memory cards.

Working Memory

In many cases this area of CMOS RAM will be at least partially battery backed to prevent loss of its contents in the event of a power loss. This is the area where the control unit will store auxiliary relay states, timer/counter states, various I/O states and arithmetic/logic calculations results as well as tables of constants and variables.

Many PLC operating systems include special functions to read/write special add-on modules such as Analog to Digital and Digital to Analog Converters.

Input/Output Circuit

Inside the PLC there are a few options that are available to the user . For a digital port these are:

* TTL-compatible low-voltage inputs
* 24V opto-isolated inputs
* 240V opto-isolated inputs.

The above circuits interface directly to the PLC internal logic circuitry. The internal circuit will also be optically isolated from the PLC external input connection. This isolation provides freedom from external noise and in particular noise generated by switches connected to the PLC that are switching heavy currents or high voltages.

The internal PLC circuitry is interfaced to either electro-mechanical or solid state relay contacts. The choice available usually includes the following:

* 24 Volt 100mA switched opto-isolated
* 240 Volt 1A ac (triac solid state switch)
* 240 Volt 2A ac (relay switched).

The output will always be either optically isolated from the external circuitry or isolated by virtue of the use of an electro-mechanical relay. This provides greater freedom from externally generated noise for the PLCs internal circuitry.

4.10. Operation and programming

A PLC program consists of a series of steps. These steps represent logical operations performed on the inputs and internal variables. The outputs can be set (energized) or cleared (de-energized) in response to the logical outcome of the program steps. This requires that the microprocessor control unit scan the inputs, execute the stored program and set or clear the outputs in a defined sequence. This takes time, and how the time is divided between tasks depends on the way that the PLC scans the input contacts. In fact, the control unit (microprocessor) might scan the input contacts as they are called for by each instruction in the program. Alternatively, the state of all inputs might be read and latched at fixed intervals, and the program accesses this stored "snapshot". Finally, changes of input levels might directly cause the execution of specific actions of the stored programs.

The output port is then driven directly on encountering an OUT instruction following a logical operation. The outputs are latched until the next scan cycle. The input scanning uses a built-in delay of around 3ms to prevent contact bounce affecting the input state read. This method applies mainly to small PLC systems.

Typical timings are:

* Instruction fetch and execute: 5ms
* Scan/actuate necessary contacts: 3ms (debounce)

4.10.1. Programming standards: IEC 1131

The IEC 1131 standard is an effort towards standardizing PLC oriented control [52] [20]. This is more necessary now that personal computers and software are starting to open the PLC market. These standards were not designed to enforce a rigid style, so the different PLC vendors will still have programming environments that vary, but the conceptual elements will be consistent [37]. The most notable differences between different PLC impementations will be addresses of outputs, inputs, internal memory, etc.

There are a few components to the standard:

* IEC 1131-3 Data types and programming
*IEC 1131-5 Communications
* IEC 1131-7 Fuzzy control.

IEC 1131-3 defines the basic programming languages.:

1. IL (Instruction List) - This is effectively mnemonic programming
2. ST (Structured Text) - A BASIC like programming language
3. LD (Ladder Diagram) - Relay logic diagram based programming
4. FBD (Function Block Diagram) - A graphical dataflow programming method
5. SFC (Sequential Function Charts) - A graphical method for structuring programs.

4.11. Industrial communication

With Siemens PLC systems, the communication from the process control to the field level is possible over the SIMATIC networks PROFIBUS, ETHERNET or TCP/IP. The family SIMATIC NET family contains an array of products with different performance characteristics: data exchange is possible over various levels, between various automation stations or various devices.

4.11.1. OPC

In the past manufacturers of SCADA systems were required to build specific drivers to connect to a whole host of different front-end equipment as each manufacturer of such equipment developed his own proprietary communications mechanism.

In order to try to reduce diversity and standardize the interface between the supervisory level and lower levels of the control system, the OLE for Process Control (OPC) Foundation was set-up in 1995 as collaboration between a number of leading automation hardware and software suppliers and Microsoft.

This initiative has led to the production of an open and flexible interface standard which is now supported by the majority of the manufacturers in this domain. This allows a SCADA system to connect to a multitude of different front-end systems via a common interface (API, Application Program Interface).

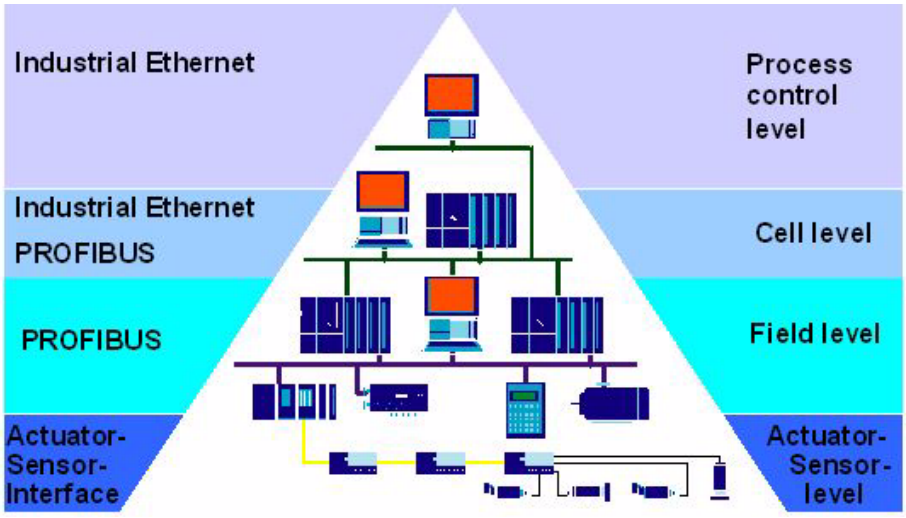


Figure 4-15 Industrial communications layers

This OPC Data Access (DA) mechanism supports three kinds of access, namely:

• Synchronous read or write

• Refresh whereby the data is read but in an asynchronous manner All OPC data items include a time stamp and a quality flag.

• Subscription whereby data is only sent on change to interested clients

Communication between PLC modules is usually accomplished via proprietary bus.

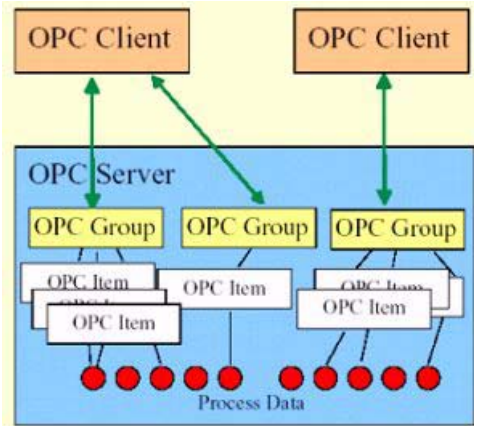


Figure 4-16 OPC communications

4.12. The thermal screen PLC

The PLC used to control the thermal screen is manufactured by Siemens and belongs to the S7-300 family (medium to large applications).

Its main specifications are the following:

* Program memory up to 85 K instructions
* Up to 1024 digital input or output lines
* Multipoint capable interface for the configuration of smaller networks.
* Quick execution time (the CPU executes 1024 instructions within 1 ms).
* Modular configuration and quick enhancements are possible through interface modules with an integrated backplane bus.
* Modular extensibility through an extensive program module of digital, analog, simulation and function modules to communications and other types of modules.
* Integrated PROFIBUS interface by the 300 2-DP Series.

Special care has been taken in choosing the 315-series CPU, since it will have to bear the computational load of thirty-two separate PID loops.

In figure 4-17 the final configuration of the system is shown: inputs to the PLCs are temperature and humidity sensors, as well as digital signals indicating the status of other interlock lines. Outputs are voltages to the heating foils and alarms to the tracker itself. An HMI (Human-Machine Interface) programmed in the Siemens ProTool/Pro environment is the interface with the operator (see Chapter 5).

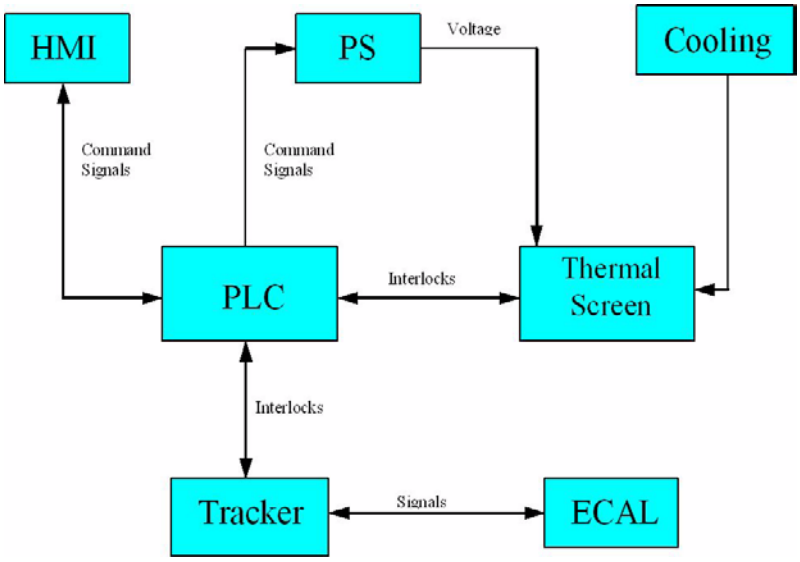


Figure 4-17 Thermal screen control system schematics

4.13. Power supply

The heating foils are powered with power supplies manufactured by Xantrex-Quality Source.

Model XPD 120-4.5 allows for a maximum output of 120 VDC at 4.5 A (540 W of output power), which is more than enough for the heating foils. Moreover, this model is rackable and remotely controllable: a PLC analog output (a voltage between 0 and 10 VDC) drives proportionally the power supply output, with a remote programming accuracy of 1% of full scale output for the default range.

A photo of the unit is shown in Fig. 4-18.

Figure 4-18 Xantrex Power Supply

4.14. PID programming

The software product "Siemens Standard PID Control" essentially consists of two function blocks (FBs) which are program areas that contain the algorithms for generating control and signal-processing functions for continuous or step controllers. It is a pure software control in which a standard function block incorporates the functionality of the controller [53].

The behavior of the controller itself and the properties of the functions in the measuring and adjusting channel are realized by means of the numeric algorithms of the function block. The data required for these cyclic calculations are saved in control-loop-specific data blocks. An FB is only required once to create several controllers.

Every controller is represented by an instance data base (DB) which must be created application-specifically.When the "Standard PID Control Tool" is used, this DB is created implicitly.

This means that the design of a specific controller is limited to specifying the structural and value parameters in the editing windows of the user interface [11]. The instance DB is created by the configuration tool.

The calculation of the algorithms for a certain controller is carried out in the processor of the S7 automation system in the set time intervals (sampling times). The calculation results and thus the updated values of the input and output variables (measuring and manipulated

variables) and status signals (limits) are stored in the corresponding instance DB or transferred to the process periphery.

In order to process several control loops which are to be executed with the same period but different delay, a controller call distribution function (Loop Scheduler = LP_SCHED) is available through which the configuration of extensive plant controls becomes structured and thus simple. In addition, even utilization of the CPU is ensured.

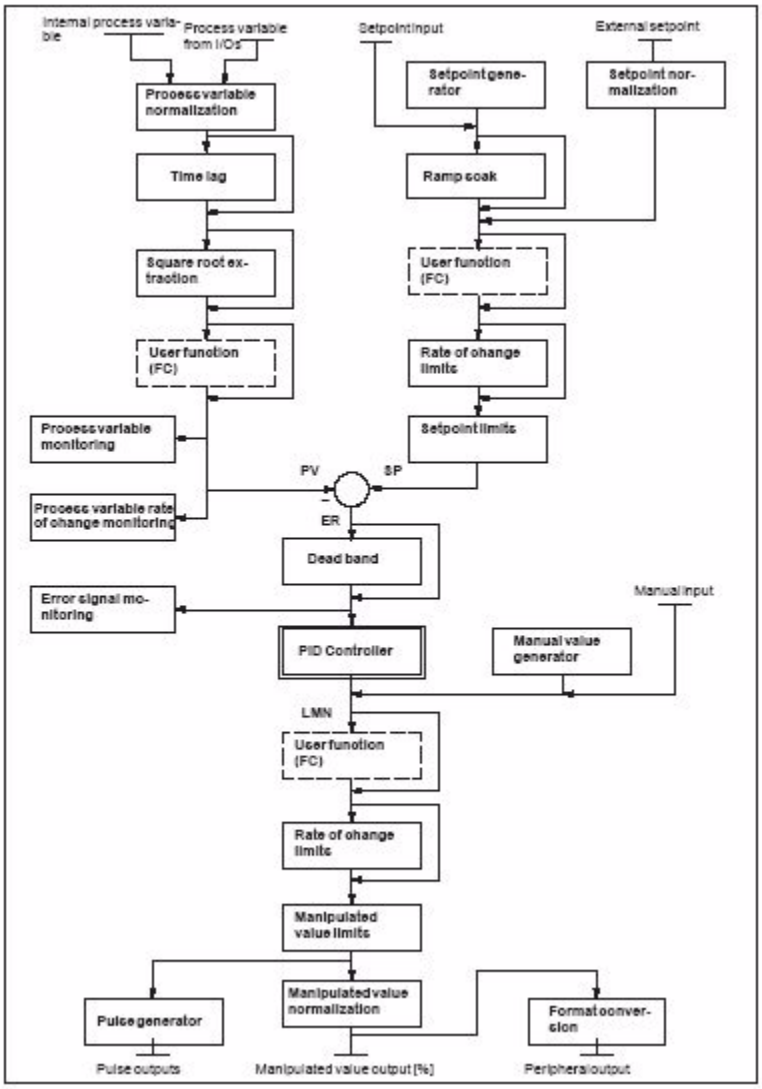


Figure 4-19 Sequence of functions of the standard PID controller.

Continuous Control with FB41 "CONT_C" FB "CONT_C" is used on SIMATIC S7 programmable controllers to control technical processes with continuous input and output variables.

During parameter assignment, the user can activate or deactivate subfunctions of the PID controller to adapt the controller to the process; the controller can be used as a PID fixed setpoint controller or in multi-loop controls as a cascade, blending or ratio controller. The functions of the controller are based on the PID control algorithm of the sampling controller with an analog signal, if necessary extended by including a pulse generator stage to generate pulse duration modulated output signals for two or three step controllers with proportional actuators.

Apart from the functions in the setpoint and process value branches, the FB implements a complete PID controller with continuous manipulated variable output and the option of influencing the manipulated value manually.

Detailed description of the subfunctions:

Setpoint Branch

The setpoint is entered in floating-point format at the SP_INT input.

Process Variable Branch

The process variable can be input in the peripheral (I/O) or floating-point format.

The CRP_IN function converts the PV_PER peripheral value to a floating-point format of -100 to +100% according to the following formula: Output of CRP_IN = PV_PER

The PV_NORM function normalizes the output of CRP_IN according to the following formula: Output of PV_NORM = (output of CRP_IN) _ PV_FAC + PV_OFF.

PV_FAC has a default of 1 and PV_OFF a default of 0.

Error Signal

The difference between the setpoint and process variable is the error signal. To suppress a small constant oscillation due to the manipulated variable quantization (for example in pulse duration modulation with PULSEGEN), a dead band is applied to the error signal (DEADBAND). If DEADB_W = 0, the dead band is switched off. PID Algorithm.

The proportional, integral (INT), and derivative (DIF) actions are connected in parallel and can be activated or deactivated individually. This allows P, PI, PD, and PID controllers to be configured. Pure I and D controllers are also possible.

Manual Value

It is possible to switch over between a manual and an automatic mode. In the manual mode, the manipulated variable is corrected to a manually selected value. The integrator (INT) is set internally to LMN - LMN_P - DISV and the derivative unit (DIF) to 0 and matched internally. This means that a switchover to the automatic mode does not cause any sudden change in the manipulated value.

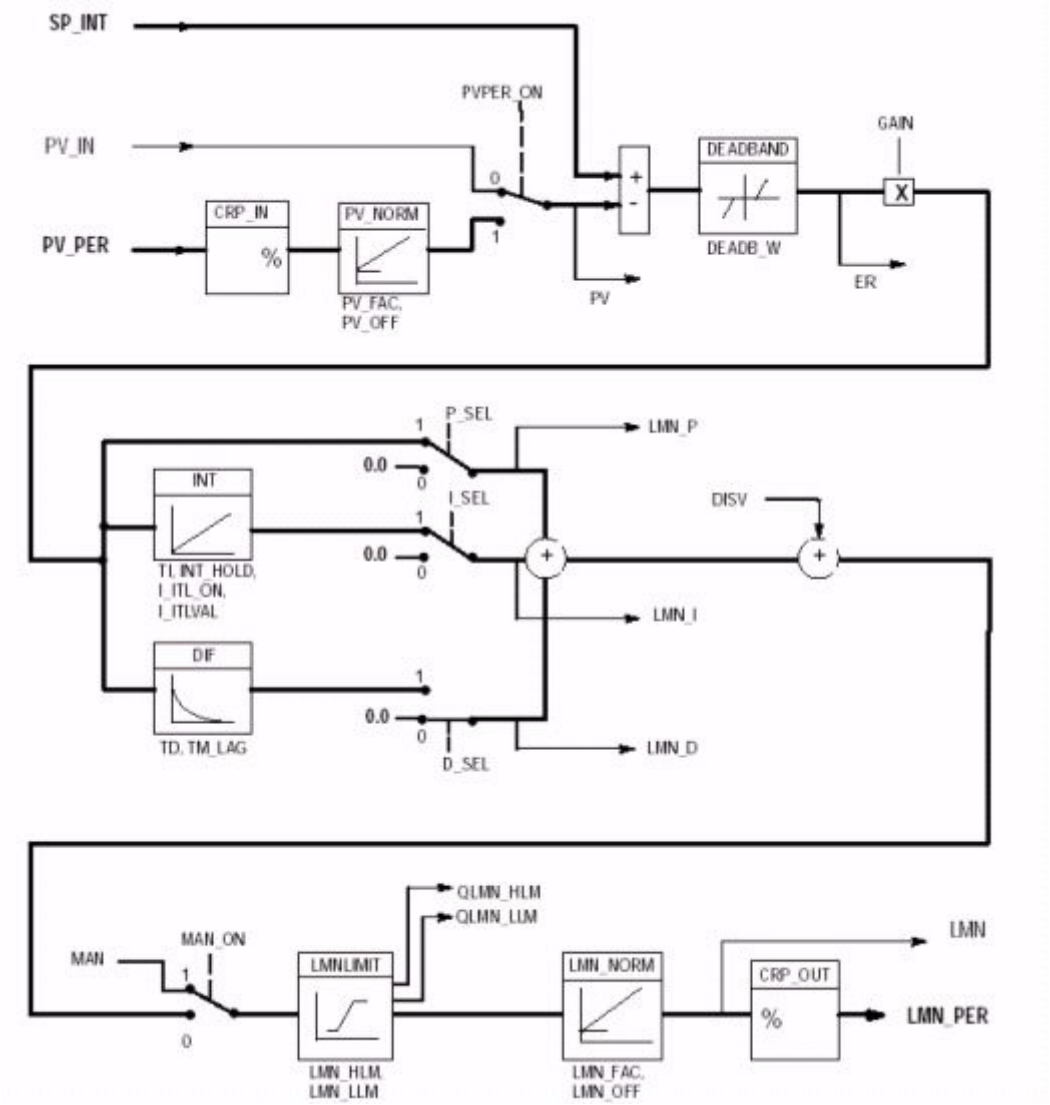


Figure 4-20 PID parameters selection points

Manipulated Value

The manipulated value can be limited to a selected value using the LMNLIMIT function. Signaling bits indicate when a limit is exceeded by the input variable.

The LMN_NORM function normalizes the output of LMNIMIT according to the following formula: LMN = (output of LMNLIMIT) _ LMN_FAC + LMN_OFF LMN_FAC has the default 1 and LMN_OFF the default 0.

The manipulated value is also available in the peripheral format. The CRP_OUT function converts the floating-point value LMN to a peripheral value according to the following formula: LMN_PER = LMN.

Feedforward Control

A disturbance variable can be fed forward at the DISV input. Complete Restart/Restart FB41 “CONT_C” has a complete restart routine that is run through when the input parameter COM_RST = TRUE is set.

During startup, the integrator is set internally to the initialization value I_ITVAL. When it is called in a cyclic interrupt priority class, it then continues to work starting at this value. All other outputs are set to their default values. The block does not check for errors internally. The error output parameter RET_VAL is not used.

4.14.1. Human Machine Interface

The components of the level SIMATIC HMI (Human Machine Interface) serve as an interface between machine and user. Functions, switches or process values are visualization on operator or touch panels [12].

With the help of this visualization, error messages or measured values for the user can be easily represented. An optical detection of processes lightens operation to the user where he/she can quickly learn the external effects of his/her actions [4]. The devices are configured with the configuration software Pro Tool that is available in three different performance variants suitable to device class. The Human Machine Interfaces can be connected directly over MPI or PROFIBUS-DP to the automation systems. With configured function switches, command buttons or display elements one can receive direct access to the CPU.

It must be pointed out, though, that there exist two kind of interface: a PLC proprietary one, shown in figure 4-36, accessible only from the Simatic environment and meant to be used by an experienced operator to tune the controller and set the main critical parameters (P,I,D, sampling time, dead band etc.).

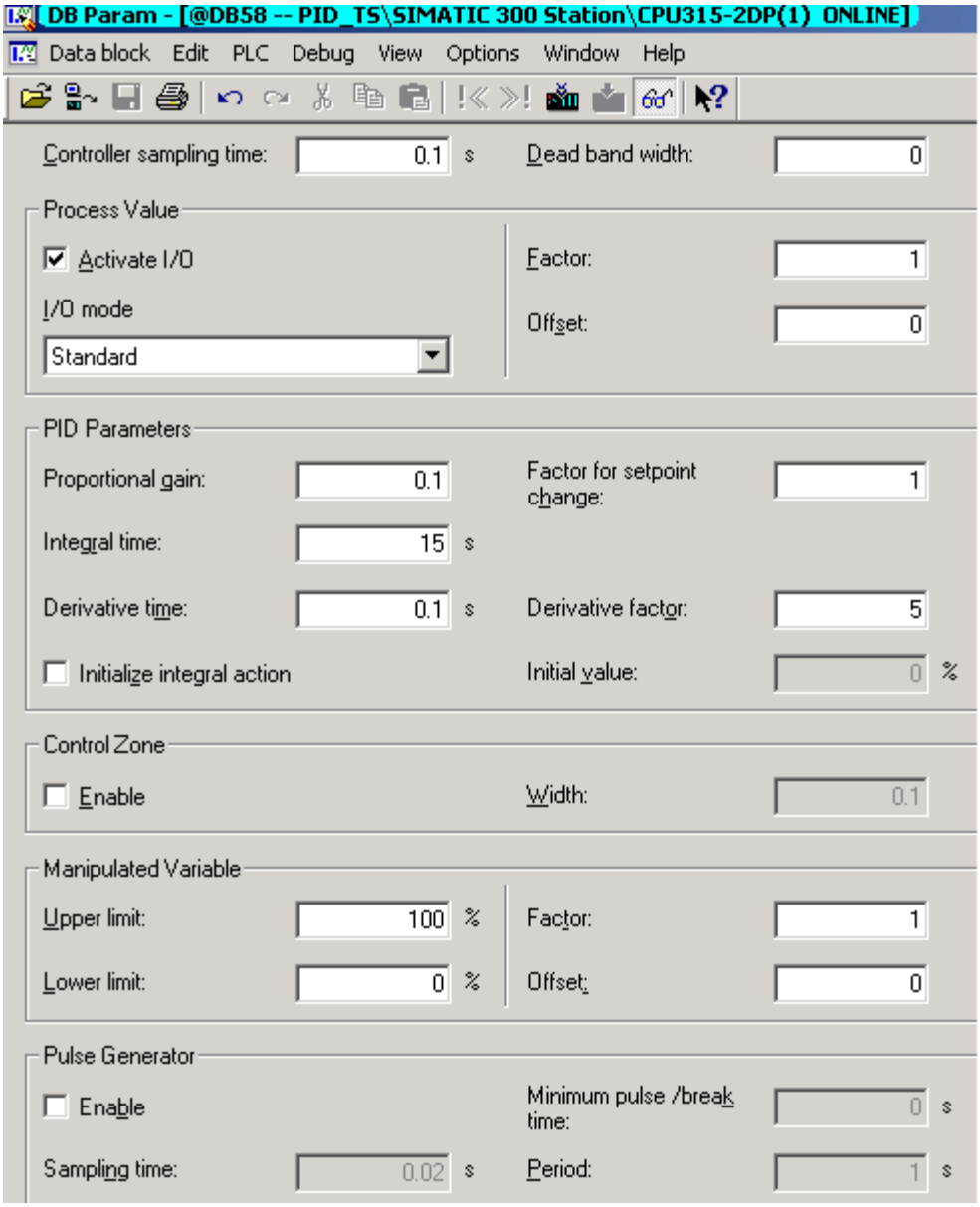


Figure 4-21 Simatic PID HMI

Another HMI has been designed in the ProTool/Pro environment; this is the interface that will eventually be displayed in the CMS control room to monitor and control the thermal screen. From here the operator has a limited access to the process paraameters, yet can see and archive all the temperatures, voltages, humdities of the panels. Different access levels are foreseen, password protected. Also, different levels of details are offered by means of different screens: the shifter will not need to know the PID parameters which, instead, are of vital importance in the debugging and maintenance phases.

Therefore, the general screen in Figure 4-23 gives the maximun of information: the user is presented with synoptic diagram of the process, where the two main parameters (temperature of the carbon support tube and of the cold plate) are in evidence, along with the trend lines. Then, for each and every heating foil a temperature and its trend is indicated. Finally, relative humidity at the two sides of the detector are also present. This is all the operator needs in order to make sure that everything is fine. On the other hand, whenever a set point crosses a fixed threshold, a warning or an alarm is issued.

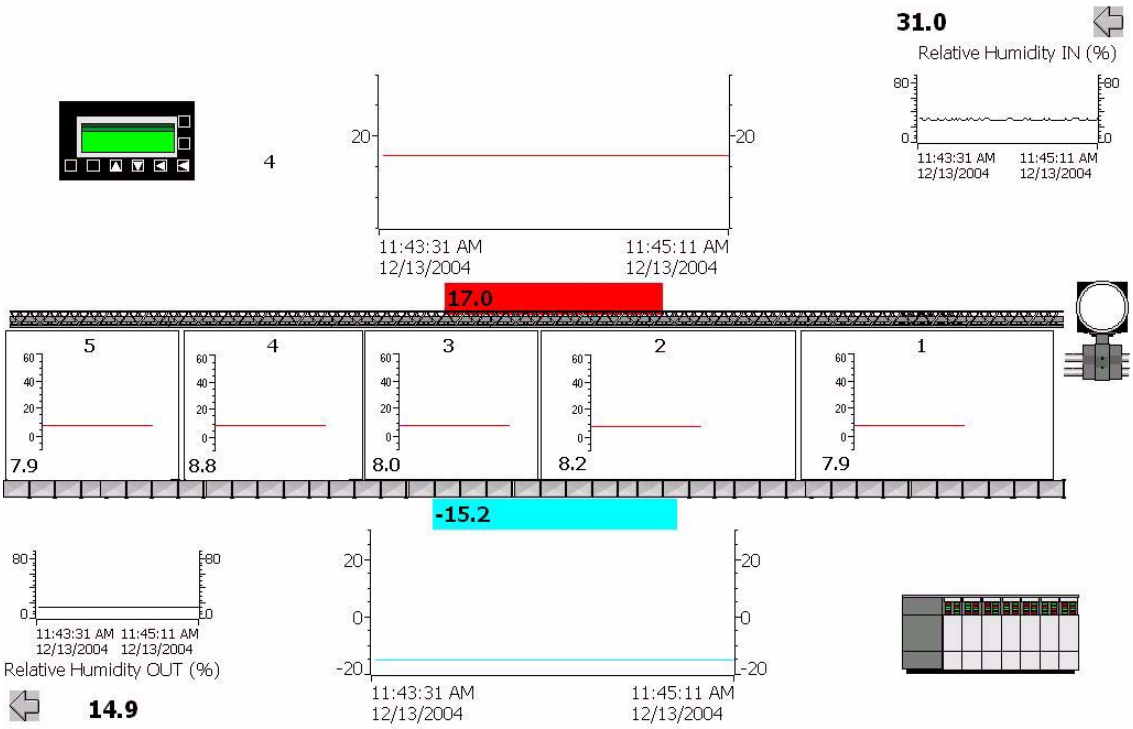


Figure 4-22 Thermal Screen HMI

The interfaces hereby presented are also used to control the test installation that will be described in chapter 5. There, there are two more screen with detailed views of the PID parameters, as in Figure 4-23 and 4-24.

The screen shown in Figure 4-23 is directly accessible from the main, and displays values and trends of the applied voltage and of the percentage of the local manipulated variable (LMN) output. The PID parameters can be controlled dynamically.More in details, Figure 4-25 shows on the same trend line the dynamic values of the PID parameters in real time; this is a useful tool to understand the P, I and D contributions while following the evolution in time of the process.

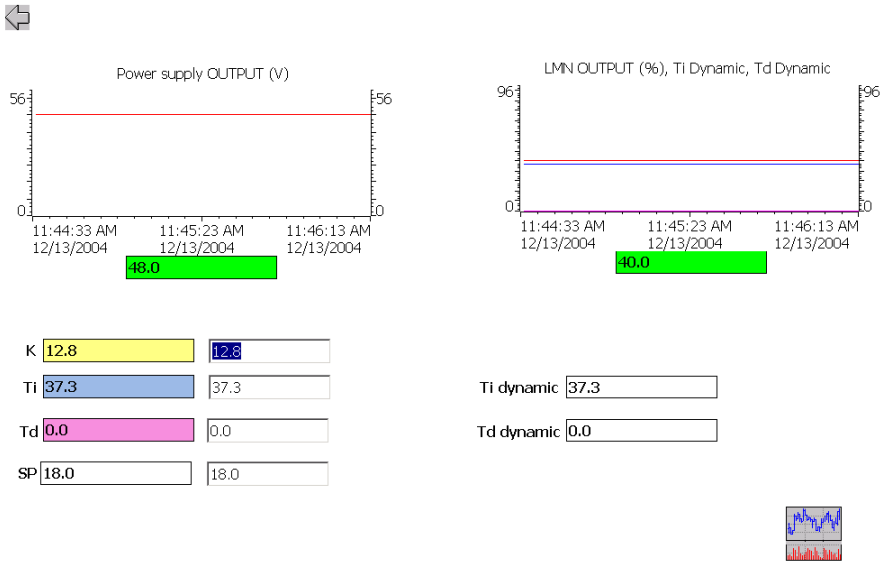


Figure 4-23 Detailed PID HMI

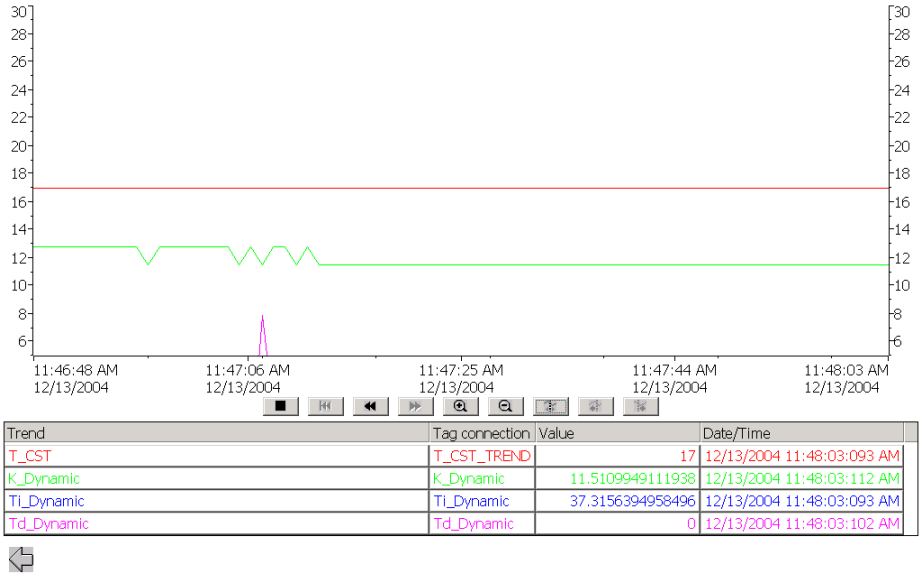


Figure 4-24 PID Dynamic Trends

4.15. Commissioning

The control hardware has been installed and cabled in a LHC rack.

The following figures 4-25, 4-26 and 4-27, represent the three rails that host the PLC modules on the rack.

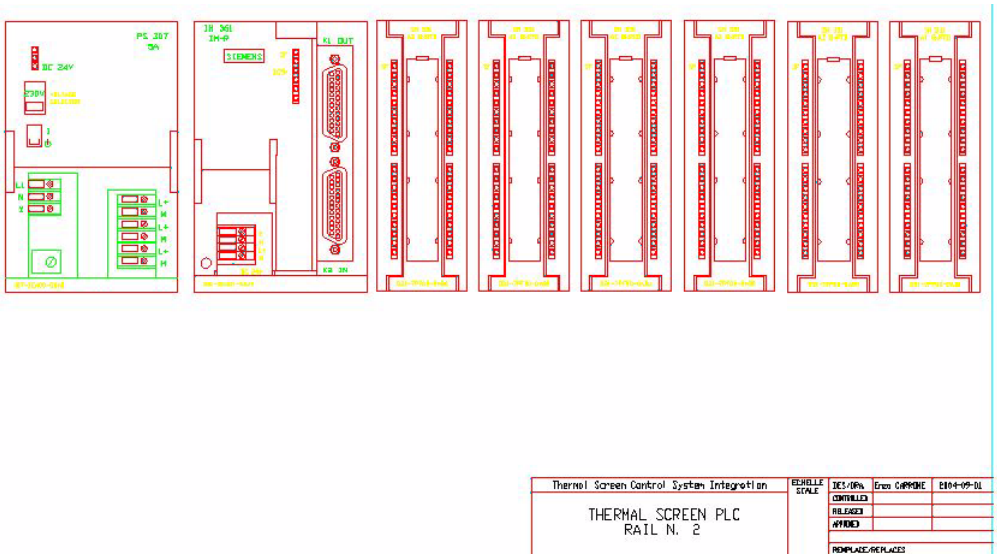


Figure 4-25 First PLC Rail

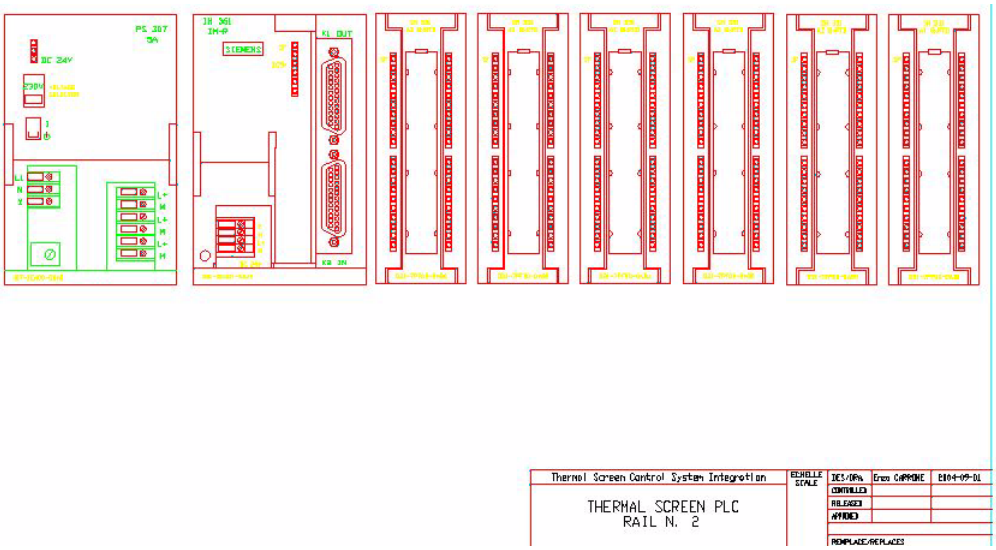


Figure 4-26 Second PLC Rail

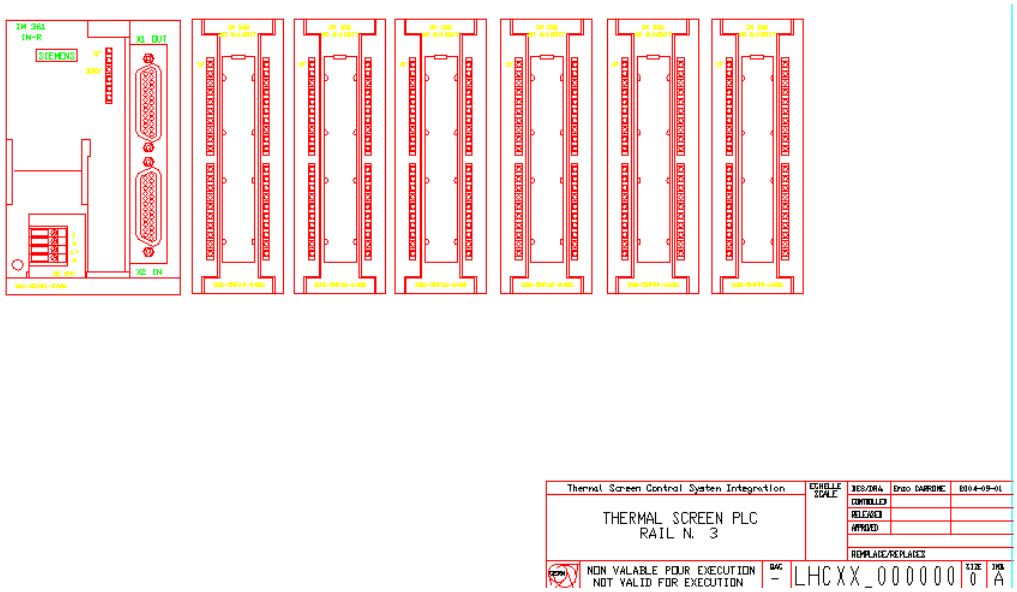


Figure 4-27 Third PLC Rack

The rack is also equipped with a patch panels where all the connectors are located. In fact the following input/output lines are foreseen (Figures 4-27 and 4-28):

- 32 Power lines connecting the power supplies to the heating foils
- 32 Signal lines, each one allowing for two temperature sensors from each panel

The backplane of the rack is shown in figure 4-32, where the almost 600 cables are appropriately wired to the connectors.

The cable length from the PLC to the patch panel is of ten meters.

The patch panel is detachable from the rack on order to provide more flexibility when the rack will be installed in the service cavern of CMS.

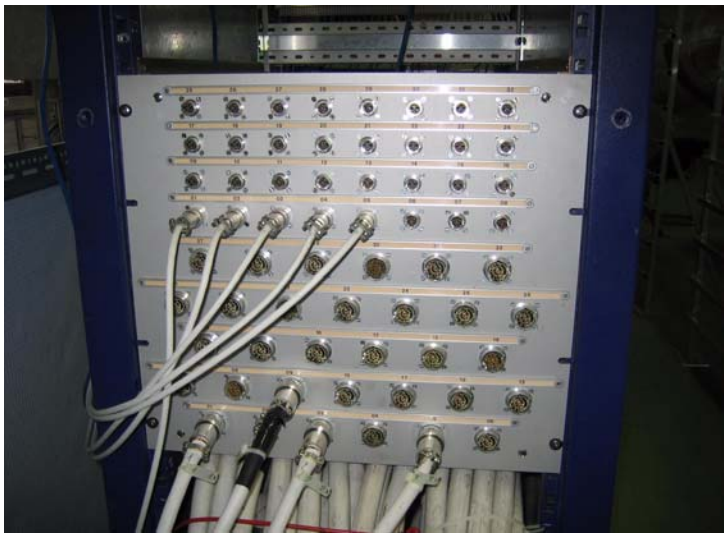


Figure 4-28 Patch Panel

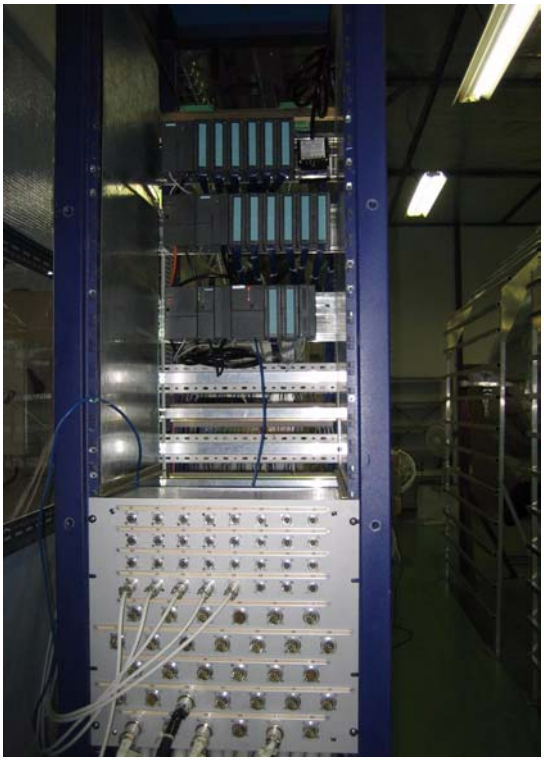


Figure 4-29 LHC Rack

Figure 4-30 Siemens PLC

Figure 4-31 Rack Backplane

4.16. Risk assessment

It was found that the best approach for the Risk Assessment of the Thermal Screen, given the available information on its constituent components, was to employ a qualitative approach. The technique used was HAZOP, or Hazard And Operability Studies. This analysis yielded a useful insight into the potential failure mechanisms of the Thermal Screen, and will thus enable better preparation for the prevention of failures within the system, during operation.

In a system as complicated and advanced as the CMS detector, there are numerous ways that failures and faults could occur. This is particularly the case when new technologies and innovative designs have been employed, as is the case with the CMS. A good way of assessing and pre-emptying such failures or faults is to carry out risk analysis. For this to be feasible at all, the system has to be broken down into smaller subsystems, and each of these then needs to be evaluated independently. One such subsystem is the Thermal Screen. There is a great deal of literature available about different Risk Analysis methodologies, and some approaches are better established than others. The main distinction between the different approaches is whether they are qualitative or quantitative. The level of involvement and accuracy then depend on the particular method. For the Thermal Screen, a technique had to be selected that would be compatible and possible to implement with the information available about the constituent components.

For a quantitative analysis, good and reliable information is required about failure and reliability rates for components, otherwise the implementation of the analysis is futile. For a qualitative approach, a thorough understanding of the technologies involved is required and a good level of experience, so that useful estimations can be made. The Risk Analysis methods that were considered initially covered the whole spectrum of risk analysis techniques, ranging from the highly involved and quantitative methods to the lower input qualitative approaches. The methods that were considered are listed below.

The analysis methods are listed in order of increasing level of involvement and time commitment required, and correspondingly the quality and thoroughness of the analysis achieved.

- Hazard/ Risk Checklist — Qualitative
- HAZOP - Hazard and Operability Studies — Qualitative
- LOPA - Layers Of Protection Analysis-Semi — Quantitative
- Event Tree Analysis — Quantitative
- Fault Tree Analysis — Quantitative

It must be pointed out, however, that even though the more quantitative methods, like Fault Tree Analysis, will tend to yield a more in-depth insight into the risks of a system, they can often be excessive for simpler systems. Therefore, the best approach depends heavily on the system that is being analyzed. For the Thermal Screen, it was eventually decided that Event and Fault Tree analysis would be inappropriate given the information available for the constituent components of the system. As a result, a combination of the Risk Checklist, HAZOP and LOPA approaches was selected for the Risk Analysis.

4.16.1. Initial Risk Assessment

As a starting point for all of the techniques mentioned above, the first method in the list – the Hazard/ Risk checklist – is often employed. This enables the Risk analysts to brainstorm and come up with as many potential faults and failures as can be predicted for the given system. In order to carry out this procedure, access to good and extensive information is required about the components and sub-systems of the system being studied.

The first step was to identify general ways in which faults and failures could occur within the TS system during operation. This was done by following the checklist procedure. There are distinct types of failures associated with different parts of the TS, and these can be generalized by the following categories: Electrical & Circuit Failures This class of faults and failures includes power, circuitry and PLC failures.

Potential risks were identified as:

- Panel RTD sensor failure - Radiation or thermal degradation.
- Communication and Power wires failure - Radiation or thermal degradation.
- PLC unit failure - Failure of core PLC unit or critical PLC constituent module.
- Power supply & Interlock failure - Failure of individual Panel power supply or of overall power. Interlock failure leading to potential hardware damage or human health risk.
- Interface communications modules failure - Fault with interface between PLC and inputs and output pathways.
- Software Failures - Failure of control or monitoring software.
- Mechanical & Hardware Failures These potential faults include the structural and material failures of components of the TS.
- Heating Element failure - Radiation or thermal degradation, or mechanical break due to panel warping.
- Coolant flow leak in panels - Loss of cooling and flooding of chamber with coolant.
- Heat Exchanger plant failure - Failure of one or two Cooling plants providing chilled coolant to the panels.
- Structural collapse of panels or the carbon support tube - Failure due to temperature warping, mechanical fatigue, or structural weakness present from manufacturing process.
- Thermal warping of panels - Would lead to non–uniform coverage of inner CST surface.
- Condensation problems - Condensation on panels and electrical contacts.
- PID Control Loop Failures - The PID control system which regulates the temperature of the thermal screen is critical in maintaining the two temperature environments. PID instability can occur if fluctuations from the equilibrium state are large due to a fault within the system. Some instability scenarios are listed below, but these are dependent on the number of temperature sensors that are eventually used as inputs to the PLC and PID systems.
- Possible variable flow rates between inflow pipes for coolant to each panel would result in slight temperature differences between panels, and once again PID control of this system may well oscillate and become unstable.

- An averaging error could result from using only one input into the PID for the 5 temperature sensors that are present. Taking an average does not take into account local temperature differences.
- Temperature difference between top panels and bottom panels could affect the PID control, again because of the possibility of constantly changing heat transfer between adjacent panels, which was modeled as being negligible in the control and thermal analysis. Top panels warmer due to convection within internal environment.
- Interactions between heating foils on same panel that are on different power supplies, especially between the TOB and TEC regions, could lead to PID difficulties.
- Human Error - As with any partially human controlled or monitored system, the TS will be susceptible to human errors. In this case, the risks are mainly concerned with the monitoring of the system, and the reaction to potential problem and initiation of alarms and shutdown procedures in case of a major problem.

4.16.2. HAZOP

HAZOP analysis involves the use of certain guidewords in conjunction with a selected system parameter to identify ways in which faults can occur, the resulting consequences and the actions that need to be taken. Appendix 6 shows the HAZOP analysis for the thermal screen, with the main guidewords used being 'No' and 'Less' to describe deviations from normal operation for the parameters shown in Appendix 4.

It is obvious from Appendix 4 that the the thermal screen does not have many safeguards in place, and that many of the potential faults could be problematic and would require repair of the system.

Nevertheless, this analysis gives a good idea of the potential failures and faults that can occur within the thermal screen system. Its is a good springboard for a more in-depth, possibly quantitative study if the appropriate failure rate information could be obtained from manufacturers.

The one system of the thermal screen that was accompanied with detailed manufacturer's failure rate data was the PLC, or rather its components. Siemens provide information on the components used in the PLC and its interface with inputs and outputs, in the form of Mean Time Between Failures (MTBF) in years. This information was used to derive an estimate for the failure rate of the PLC overall.

Components of the PLC with listed MTBFs are shown in table 4-2.

The calculation of the overall MTBF is done by:

Order Number	Type Description	MTBF	Quantity
6ES7307-1EA00-0BA0	PS307, 5A	103.8	2
6ES7315-2AG10-0AB0	CPU 315-2DP	22.2	1
6ES7321-1BL00-0AA0	SM 321, DI 32 x DC24V	30.6	1
6ES7322-1BL00-0AA0	SM 322, DO 32 x DC24V/0,5A	16.5	1
6ES7331-7PF00-0AB0	8 Pt. RTD Input	34.7	8
6ES7332-5HF00-0AB0	SM 332, 8AO	26.3	5
6ES7360-3AA00-0AA0	IM 360	51.9	2

Table 4-2. MTBF for PLC components

Eq. (4.1)

$$\frac{1}{MTBF_{Total}} = \frac{1}{MTBF_{ComponentA}} + \frac{1}{MTBF_{ComponentB}} + \ldots$$

Hence, the overall estimation of the failure rate in MTBF (years) for the PLC, from information for these components, is:

MTBF PLC = 1.647

This is an estimation rather than an absolute value, but it does give a good indication of the reliability of the PLC in operation. The value may be surprisingly low, but it will be a conservative estimate, and some of the component failures will not result in a critical failure. A system with so many electrical components, like the PLC, is susceptible to component failures.

5 Identification and Tests

Once the controller design has been frozen, a fully equipped rack has been commissioned in order to house the PLC CPU and the input/output and communication modules. A detachable patch panel has also been provided with connectors for power and signal cables.

A test bench has been designed and constructed at CERN (Building 186) in order to perform acceptance tests on the panels and to tune the controller on a real system.

A full enclosure is provided by a climatic chamber with room for a single thermal screen panel; the chamber is airtight, and also contains a section of the carbon fiber support tube on top of the heating foils, thus allowing for a setup as similar as possible to the final installation.

The chamber is constantly flushed with dry air at a 200 l/h rate. Two humidity sensors have been installed on both sides of the chamber (dry air input and output); their signals, along with the output o the RTDs sensors are gathered from the PLC.

The cold plate is fed by a perfluorohexane cooling unit which is an exact replica of what will be installed in the final system: it provides freon (C_6F_{14}) at -15 °C, with a flow rate of 10 l/h. The PLC itself is assembled in a PLC standard rack (6 U) in its final, full configuration. It means that the test is not performed on mockups or on reduced-scale models, but on the very same components that will eventually be installed in the cavern. A PC is also connected to the PLC via an OPC server, in order to provide archiving, trending, HMI and PID tuning.

This setup has allowed the system identification, a procedure that needs a real process to work on. In the following, we will show how the system identification has been performed and which results were gathered. Moreover, we will also discuss testing procedures and standards.

Since the thermal screen has to provide a Safety Interlock Level (SIL), the control system has also been tested in the beam test area during detector operations, both in September 2003 and May 2004. There, interlock lines have been connected to the final power supplies, and the PLC has been working permanently in order to protect the detectors under test against temperature and humidity excesses, and power cuts.

5.1. Test bench and climatic chamber

The thermal screen control system is based on a PLC from Siemens, which includes S7-315 CPU, along with the TCP/IP communication module, forty analog outputs, forty analog (RTD) inputs, eight digital outputs and eight digital inputs.

Each thermal screen panel will be powered by an individual power supply; since the temperature will be monitored on the heating foils, it is useful to study the relation between this temperature and that of the carbon support tube. Five temperature sensors are available on the heating foils of each panel, thus requiring a choice in terms of which temperature is most representative of the whole behaviour. The final choice has been checking the temperature directly on the Carbon Support Tube and using this data as the Controlled Variable

The coolant temperature is not -10 ° C as will be in the real installation, but -15 ° C, because of constraints related to the cooling plant. It means that the check on the controls will be performed in even more severe conditions.

Figure 5-1 shows an artistic view of the chamber, while figures 5-2 and 5-3 show the two-dimensional view. There, the red profile of the thermal screen panel is evidenced in red, below the carbon support tube foil. The pipes for the dry air are also shown, along with the flanges for the output cables. The panel is connected to the PLC and to the power supply by means of ten cables for heating foils temperature sensors an power. Temperature sensors are also connected to the support tube and to the cold plate.

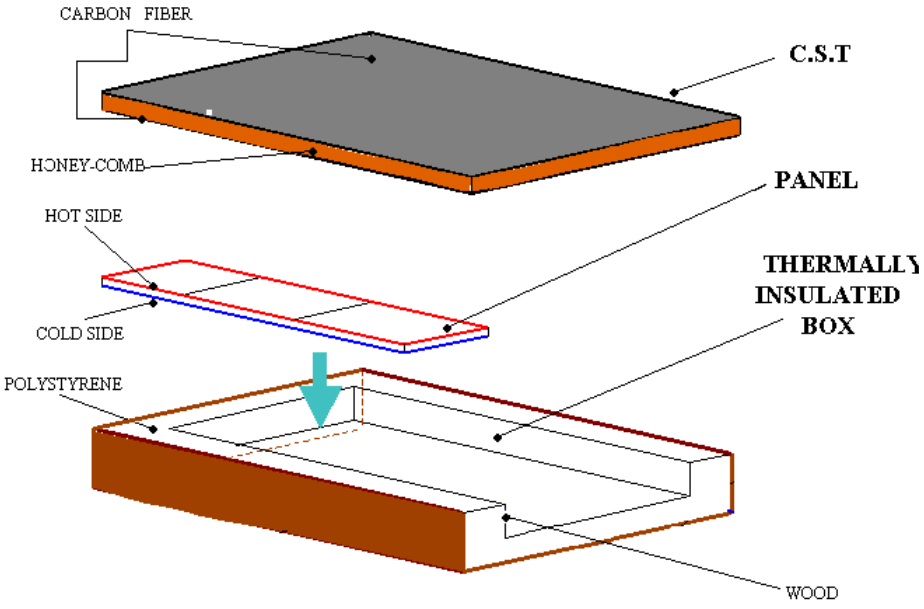
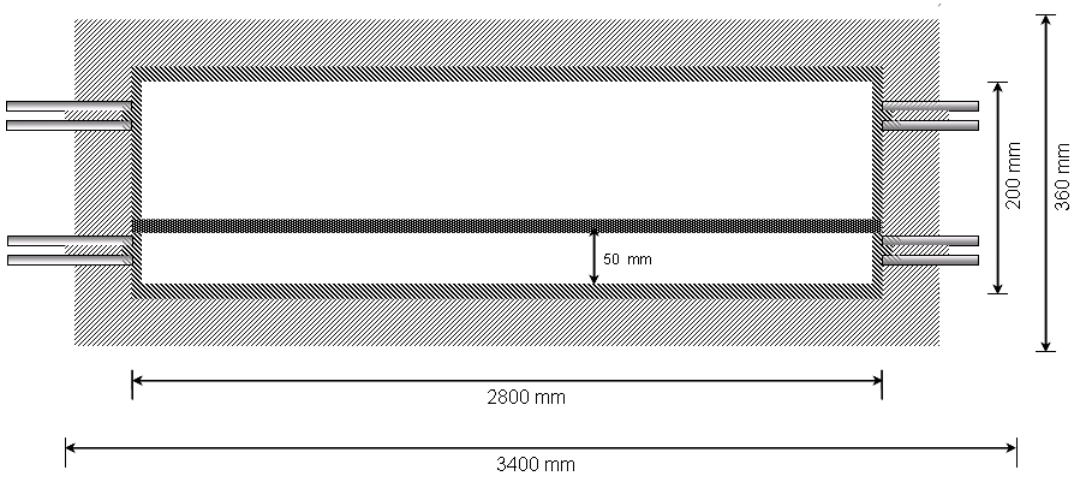


Figure 5-1 Climatic chamber artistic view

Figure 5-2 Climatic chamber - Side view



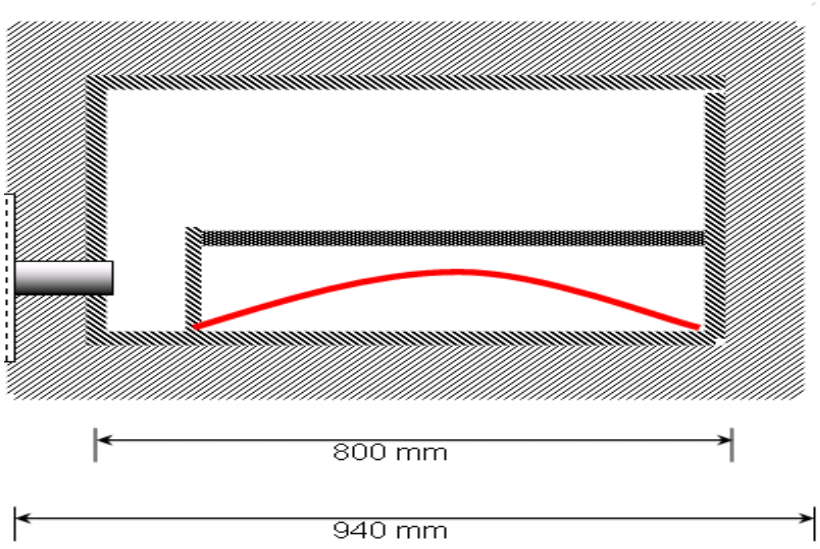


Figure 5-3 Climatic chamber - Front view

Figure 5-4 Climatic Chamber - External View

Figure 5-5 Climatic Chamber - Internal View

5.2. Identification process

A test bench, based on the above mentioned climatic chamber, has been designed in order to perform a system identification of the heating foil. In fact, the behaviour of this active part of the system is difficult to model, but plays a fundamental role in the performance study, due to its inherently non-linear behaviour.

One heating foil has been connected to a power supply (48 V DC), a step input has been applied (thus allowing for a disturbances study and not only transfer function) and its temperature has been monitored over a 15000 sec. time. The data have been collected by the thermal screen Siemens PLC, and analyzed with the MatLab software, System Identification toolbox. A sampling time of 1 sec. has been set.

Using the MatLab Ident toolbox, we are presented with a user interface:

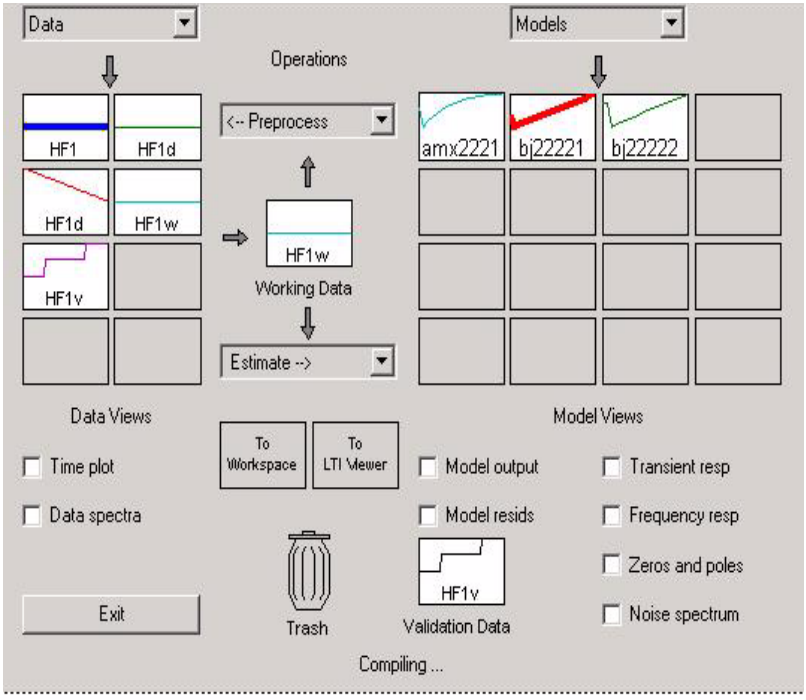


Figure 5-6 Matlab Identification Toolbox GUI

The data to analyze are on the left, while on the right side are the estimated models.

To measure the process response experimentally, we will deactivate the controller (so called manual mode) and apply a step change to the final control element. We will monitor the CV response through the sensor. This means that we will observe the net effect of disturbances, not only just the transfer function.

The data collected from 1 to 2500 sec. are used for model identification purposes, and those from 2500 to 5000 sec. are used for validation.

In order to compute a parametric model, a Box-Jenkins type difference equation model has been selected, with a linear difference equation that relates the input u(t) to the output y(t) as follows:

Eq. (5.1)

$$y(t) = \frac{B(q)}{F(q)}u(t-nk) + \frac{C(q)}{D(q)}e(t)$$

with

Eq. (5.2)

$$D(q) = 1 + d_1 q^{-1} + \ldots + d_{nd} q^{-nd}$$

After the model has been identified, it can be compared to the data-set for validation, in this case, ARX (AutoRegressive eXogenous) and BJ (Box-Jenkins) are shown:

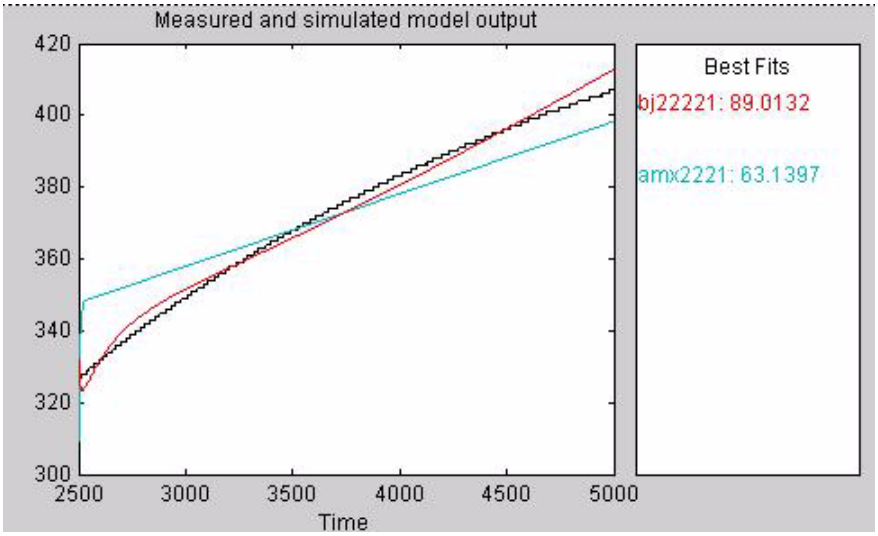


Figure 5-7 Models comparison

The plot shows the simulated (predicted) outputs of selected models. The models are fed with inputs from the Validation Data-Set, while the measured output is plotted in black. A percentage, measuring how closely the prediction of each model matched the "true" values, is displayed of the output variations that is reproduced by the model is displayed at the side of the plot. A higher number means a better model.

The precise definition of the fit is:

Eq. (5.3)

$$FIT = \left[1 - \frac{\|Y - \hat{Y}\|}{\|Y - \bar{Y}\|}\right] \cdot 100$$

where Y is the measured output, $\hat{Y}$ is the simulated/predicted model output and $\bar{Y}$ is the average value.

The BJ22221 model analytical representation is the following:
Discrete-time IDPOLY model (q is the shift operator):

$$y(t) = [B(q)/F(q)]u(t) + [C(q)/D(q)]e(t) \quad \text{Eq. (5.4)}$$

$$B(q) = -3.793e\text{-}005\ q^{-1} + 4.399e\text{-}005\ q^{-2}$$

$$C(q) = 1 - 0.8882\ q^{-1} - 0.004573\ q^{-2}$$

$$D(q) = 1 - 1.863\ q^{-1} + 0.8627\ q^{-2}$$

$$F(q) = 1 - 1.994\ q^{-1} + 0.9936\ q^{-2}$$

Transfer function from input "Voltage" to output "Temperature":

Eq. (5.5)

$$\frac{-4.109\ e^{-5}\ s + 6.088e^{-6}}{s^2 + 0.006375\ s + 8.011e^{-8}}$$

It is also possible to study the time and frequency response of the BJ model (Figure 5-8); it was already clear from the step response that the system is inherently unstable (no steady-state is reached, one real zero).

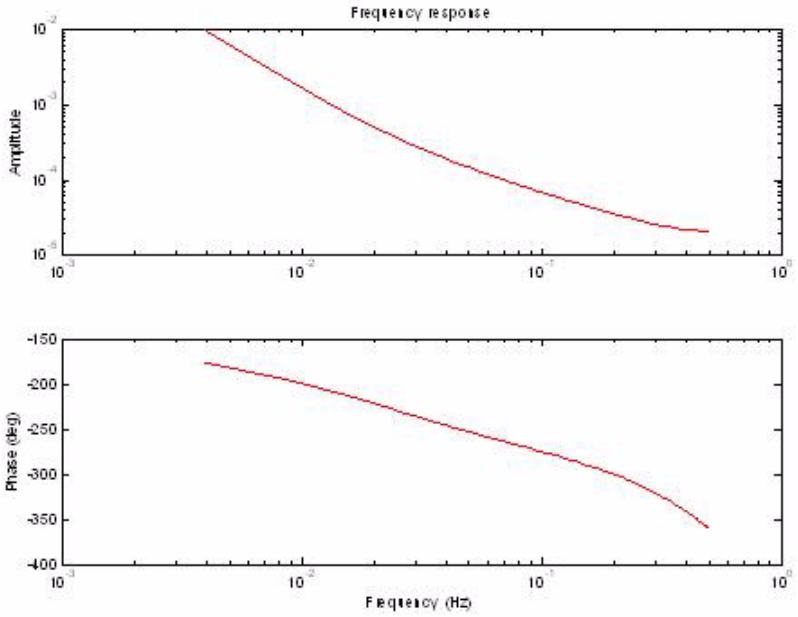


Figure 5-8 Time and frequency response of the BJ model

This is evidence that the thermal screen panel needs a controller in order to behave properly. Even if the design is such that the same amount of power is generated and removed at the same time on both sides of the screen, it is clear that without an appropriate control the action on the heating foils side will be ineffictual.

In order to better understand the behaviour of the system as a whole, an identification has been performed on a full thermal screen panel.

5.3. System Identification of a panel

A further identification on the whole panel has been performed with a digital noise generator: the system was powered by a Wavetek Model 395, 100 MHz synthesized Arbitrary Waveform Generator.

Its specifications are the following:

Sum Input

Level: ±5 V p-p maximum

Impedance: 600 Ω

Bandwidth: >30 MHz

Output

Low level: 0V ±5 V into open circuit

High level: +10V ±0.5 V into open circuit

Impedance: >50 Ω

Rise/fall time: <7 ns

Protection: short circuit without damage.

Period

Range: 655 μ sec. to 10 seconds

Resolution: 4 digits

Accuracy: ±100 ppm

Digital noise

It provides a random (0,1) pattern from the main output.

Clock range: 10 mHz to 100 MHz

Resolution: 4 digits

Accuracy: 100 ppm

Sequence length: 2^n-1; n=6,7,...,16 (It defines the length of the output sequence which will eventually repeat itself).

We selected a sequence length of 65535, a V p-p of 2,2V and a frequency of 200 mHz.
Data have been gathered over a time span of 15000 sec.
The Matlab Ident toolbox is identical to that of the heating foils identification:
We can also plot the input and output signals vs. time:

The models that perform best are the BJ and the first order P1.

- The BJ Model:
Discrete-time IDPOLY model: y(t) = [B(q)/F(q)]u(t) + [C(q)/D(q)]e(t)

$B(q) = -3.427e^{-6}\ q^{-1} + 3.265e^{-6}\ q^{-1}$

$C(q) = 1 - 0.0581q^{-1} - 0.004475\ q^{-2}$ q^-2

$D(q) = 1 - 1.04\ q^{-1} + 0.03964\ q^{-2}$

$F(q) = 1 - 1.523\ q^{-1} + 0.5234\ q^{-2}$

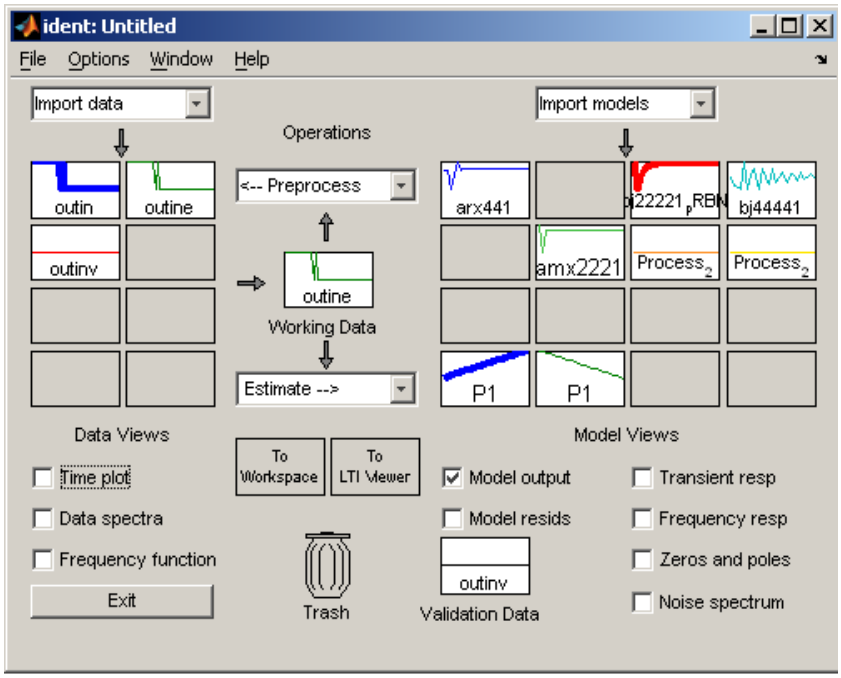


Figure 5-9 IDENT screenshot for the PRBN test

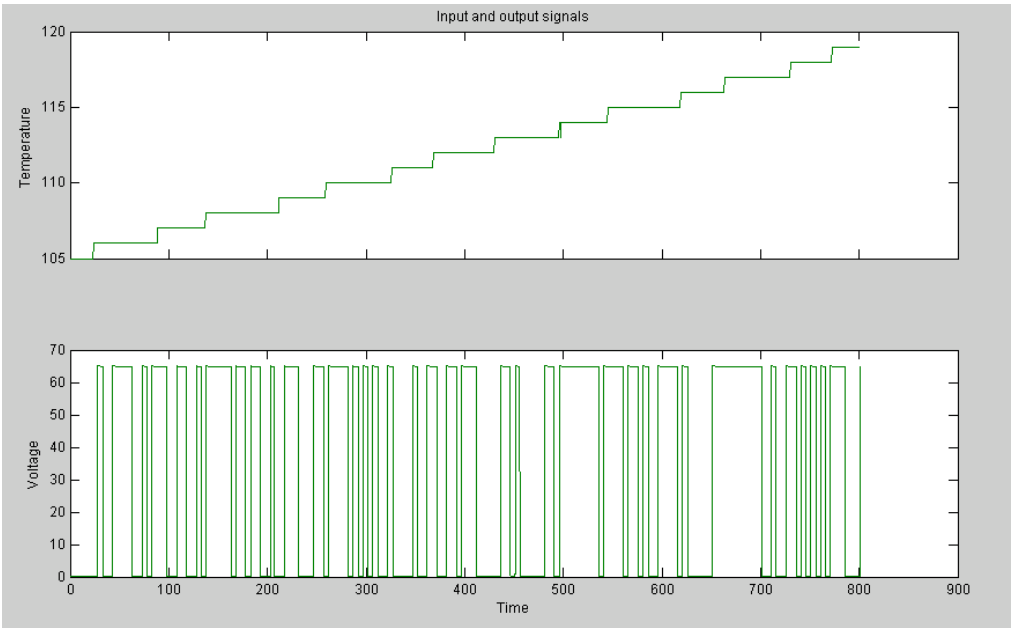


Figure 5-10 Temperature and voltage vs. time diagram

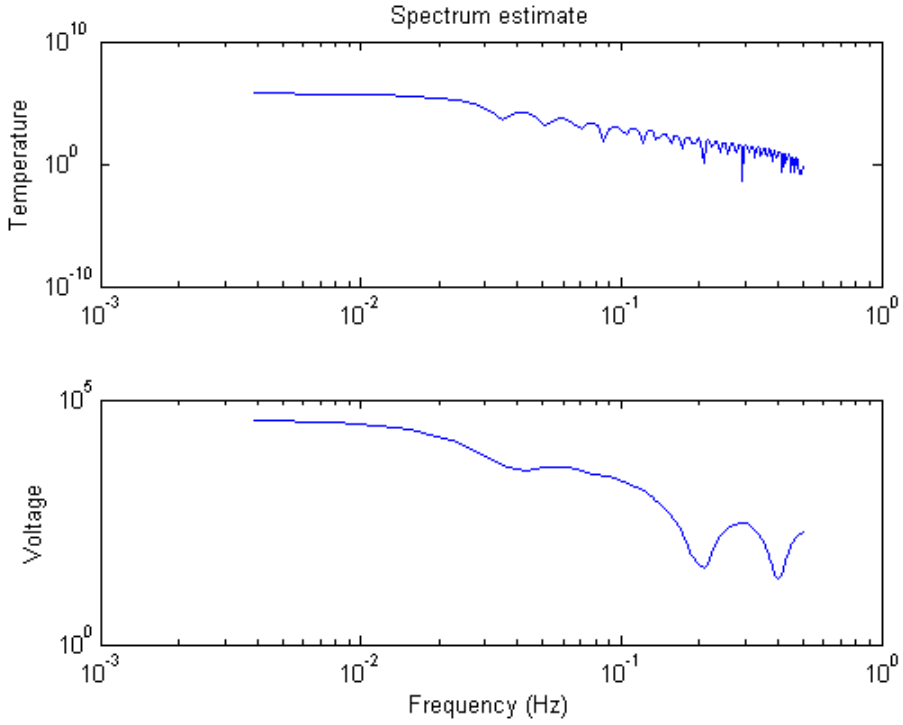


Figure 5-11 Spectrum Distribution

- Process P1 Model:
Process model with transfer function

$$G(S) = \frac{K}{1 + T_{P1} \cdot S}$$

with K = -0.00050863, and Tp1 = 1562.4.

The Bode diagram and the step response of the P1 model are shown in figure 5-12 and 5-143 respectively. These diagrams are in good accordance with the model of the system previously calculated. The difference in terms of time constants can be explained by the following facts:

1. The thermal capacitance of the carbon support tube used in the thermal model calculations is derived from separate experimental tests; therefore, real values could be different.

2. The presence of the thermal mass of the tracker apparatus constantly kept at -10 ° C can not be completely reproduced in the experimental setup.

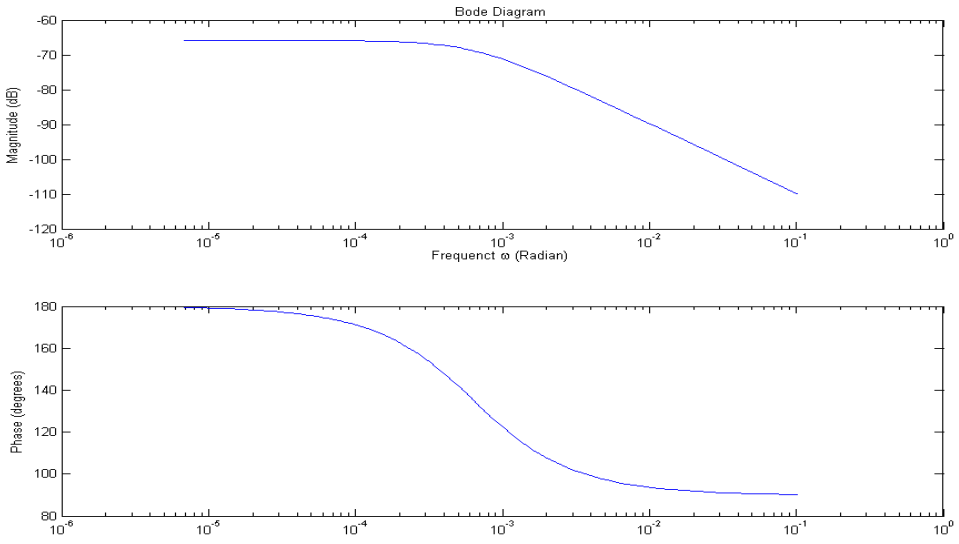


Figure 5-12 Bode diagram of the P1 Model

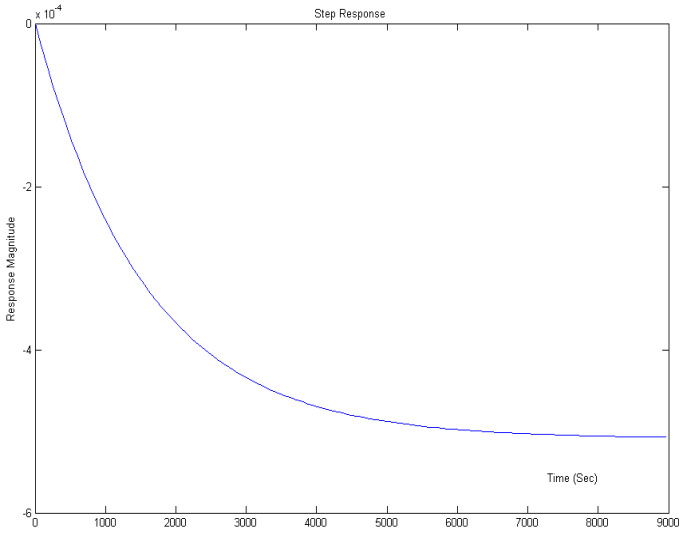


Figure 5-13 Step response of the P1 model

5.4. Controller tuning

Four different tuning methods have been applied to the process.

Let us remember that the thermal screen consists of thirty-two panels, and each panel is provided with five heating foils. The temperature on the tracker support tube is controlled by applying a variable voltage to the five heating foils connected in parallel.

Since there are many different tuning techniques available, the choice has been driven by the following considerations:

- The controller must be easily tunable.

Operations of the thermal screen must be assured over more than ten years' time, and tuning methodologies currently available allow a large degree of complexity; easy tuning techniques will be of great help in maintenance or fine tuning phases, when different operators may need to intervene on the controller parameters.

- The controller must take into account model uncertainty.

Even if a thorough test has been performed using the climatic chamber, the final tracker installation is bound to present some differences, starting from the operational temperature of both the tracker and ECAL installations, to different thermal loads on the two sides of the support tube. Ideally, a controller should be easily tunable even when the model is moderately different from the real operational conditions.

- The controller should allow the designer to trade off control system performances against control robustness.

Tuning is always a matter of trade off among the various process constraints. The thermal screen could require performances that are tougher than expected to be achieved; we presently ask for the temperature on the carbon support tube to span within no more than 1 °C from the set point. Therefore, if needed, it should be possible to sacrifice some performance parameters in order to improve othes (for instance, a slower response "traded" for a higher stability).

The four tuning techniques applied to perform the test are the following: Skogestad (Lambda tuning), IMC (Internal Model Control), IAE minimization, MatLab Optimization Toolbox.

It must be noted that the process was designed to be operated with a coolant temperature of -10 °C, yet these tests have been performed with a coolant temperature of -15 °C, different from the "final" value because of the cooling plant characteristics. Therefore, in order to let the controller and heating foils manage to effectively keep the support tube at the nominal temperature, the set point has bee changed from 18 °C to values between 16 and 17 °C.

This allowed not only to check the efficiency and the performances of the controller, but also to validate these performances under conditions that are different from the operational ones, helping to ensure that the controller works correctly, and with satisfactory results, even under substantial changes in the configuration.

1st setup. Skogestad (Lambda tuning), Figure 5-14.

Lambda=0.1 (very aggressive controller). Kp=12.79, Ti=108.52, Td=10.42.

Set Point=17 °C. Temperature oscillations are not kept within 1 °C as desired, since they span from 16.1 to 17.2 °C (control window of 1.1 °C).

It is evident that, although it is relatively easy to prevent the system from overheating (the maximum temperature reached is only 2 °C more than the setpoint), the cooling curve is steeper, which means that cooling down is faster than warming up; a less steep cooling down behaviour would be desirable. This is also a symptom that a derivative component could be of some help.

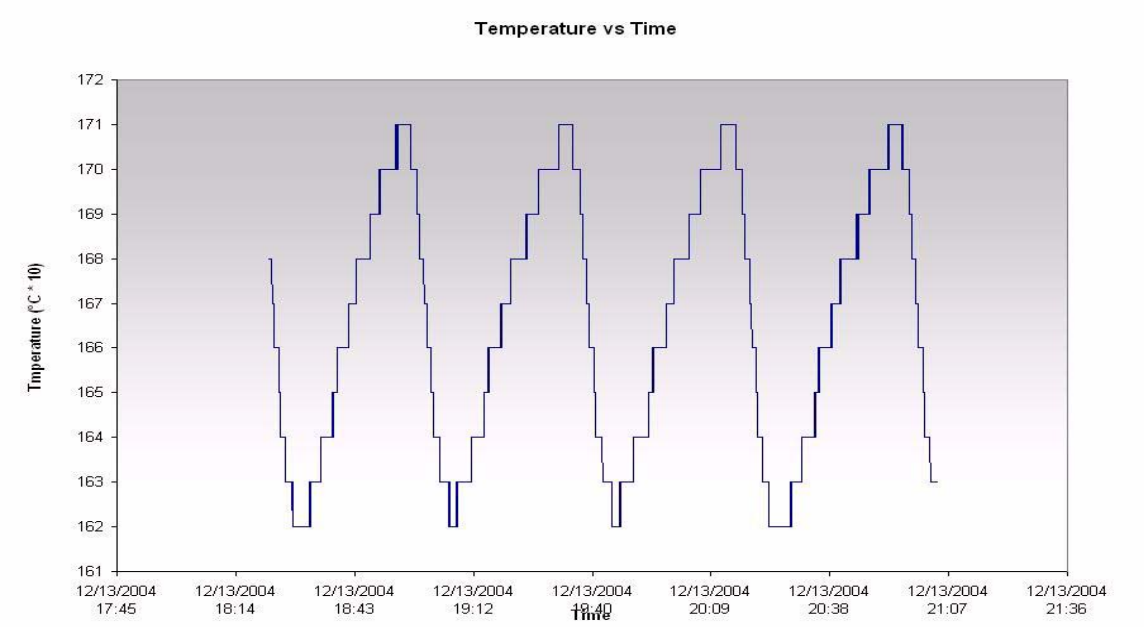


Figure 5-14 Skogestad tuning

2nd Setup. Internal Model Control, Figure 5-15.

Kp=0.465, Ti=231.2, Td=0, Set Point=17 °C. Here the upper and lower limits are 16.4 and 17 °C respectively, which means a control window of 0.6 °C. The system response is smoother, and the temperature is kept within a smaller window (the beneficial effect of the derivative component is clear). Moreover, the maximum temperature is exactly that of th setpoint.

This is a very good result, since it is preferable that the temperature of the carbon support tube oscillates below the set point rather than above; this result is perfectly accomplished.

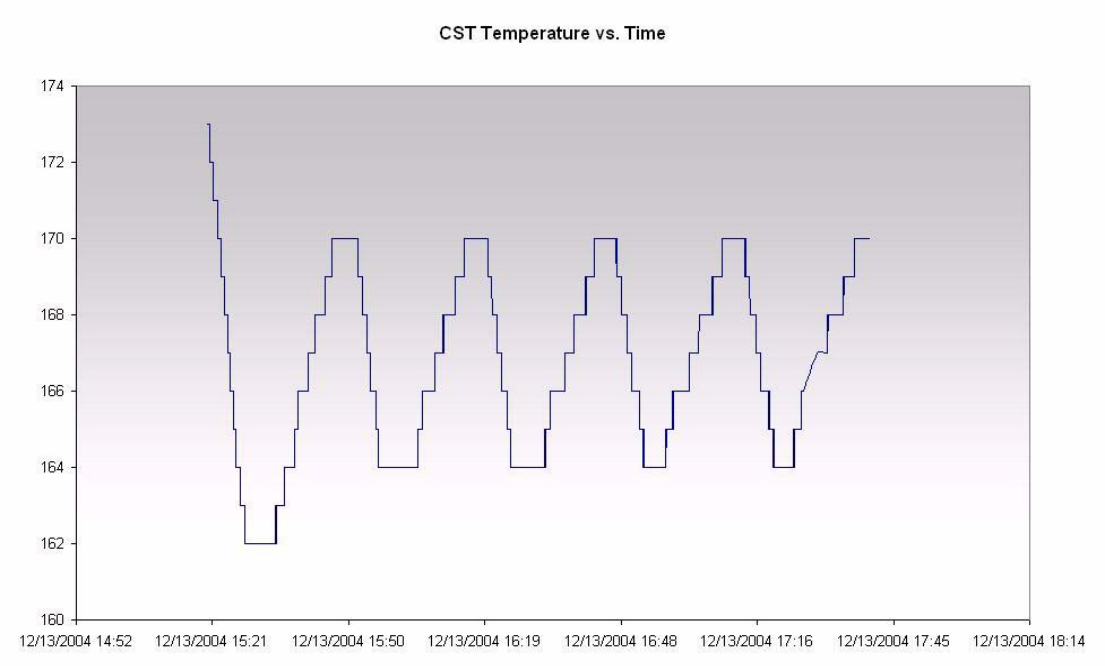


Figure 5-15 IMC Tuning

3rd Setup. IAE minimization, Figure 5-16.

This tuning approach gives the following results, with Kp=0.464, Ti=231, Td=0, Set Point= 15.9 ° C. Temperature spans in a range from 15.4 to 16.1 ° C (window width: 0.7 ° C).

The control window is getting wider, and the controller is becoming less optimally tuned.

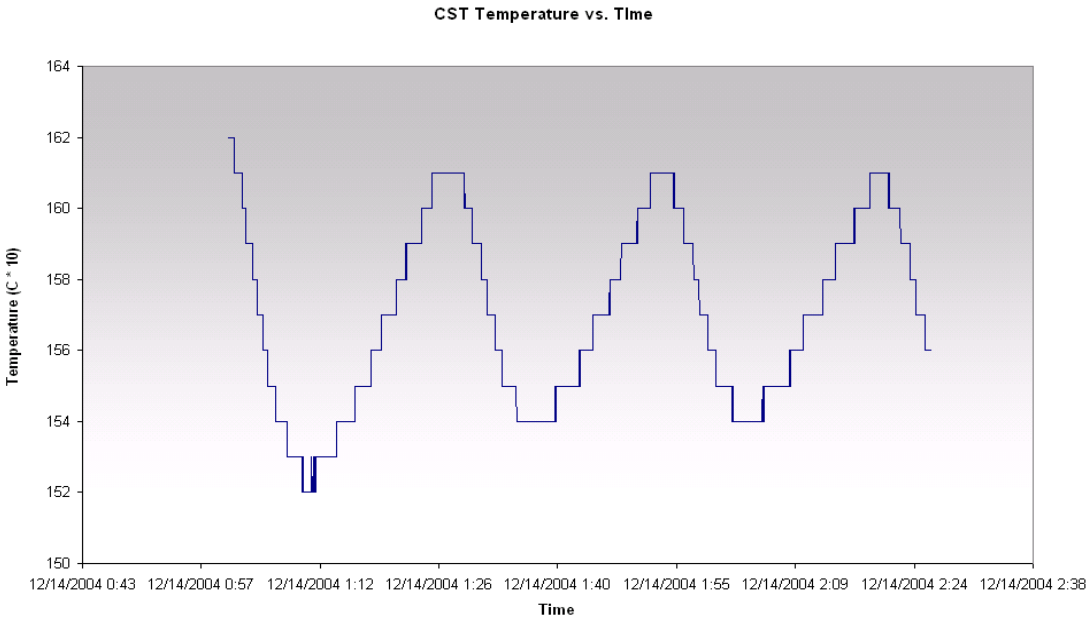


Figure 5-16 IAE Minimization Tuning

4th Setup. MatLab Optimization toolbox, figure 5-17.

We specified a rise time of 80 seconds, settling times of 120 sec., overshoot of 1.4, getting Kp=621.24, Ti=206.35, Td=0, Set Point= 16 ° C. Here the control windows is of 1 ° C, which is still adequate (although wider if compared to the other techniques). The behaviour in terms of temperature rise and fall is quite the opposite with respect to the first tuning (Skogestad): the warming up phase is steeper then cooling down. This kind of behaviour could be a source of overheating in case of strong system perturbations. Therefore, it does not allow for good tuning in the specific application.

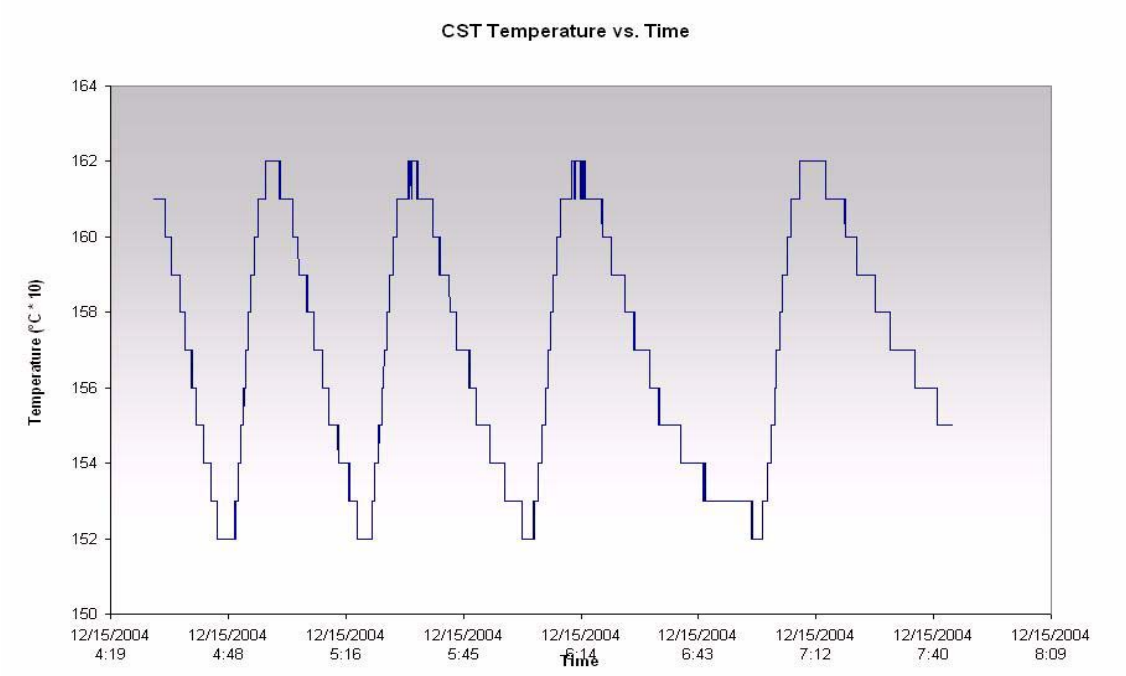


Figure 5-17 MatLab Optimization Toolbox tuning

According to the previous considerations, the IMC method appears to be the most appropriate for the task at hand.

The best performance achieved by this controller is the proof of the utility of developing a model of the process before accessing the hardware setup: the IMC tuning strongly relies upon a model of the system and constantly compares the predicted behaviour to the real one. This is also the evidence that the extra effort done to model the system pays off in terms of performances and of quality of the control.

All the tuning rules adopted guarantee the system stability. The differences are not even in terms of capability to follow the set point, where all the setups hereby described proved to succeed, but rather in providing a fast tracking of the controlled variable with a smooth behaviour and keeping the temperature closer to the lowest limit, which could be tricky under large system perturbations.

The behavior of the thermal screen panel has also been tested with respect to the individual heating foils (each panel is made out of five heating foils in groups of two sizes: two foils measuring 560 * 435 mm, three foils measuring 443 * 435 mm).

We monitored every single temperature on every foil while tuning with the IMC technique, to check against divergent behaviors. It turns out, as shown in figure 5-19, that all the five foils follow the same trends imposed by the applied controlled voltage, spanning in a temperature range from 6 to 10 °C.

This is quite a satisfactory result, since it assures the uniformity of the temperature over the whole thermal screen panel.

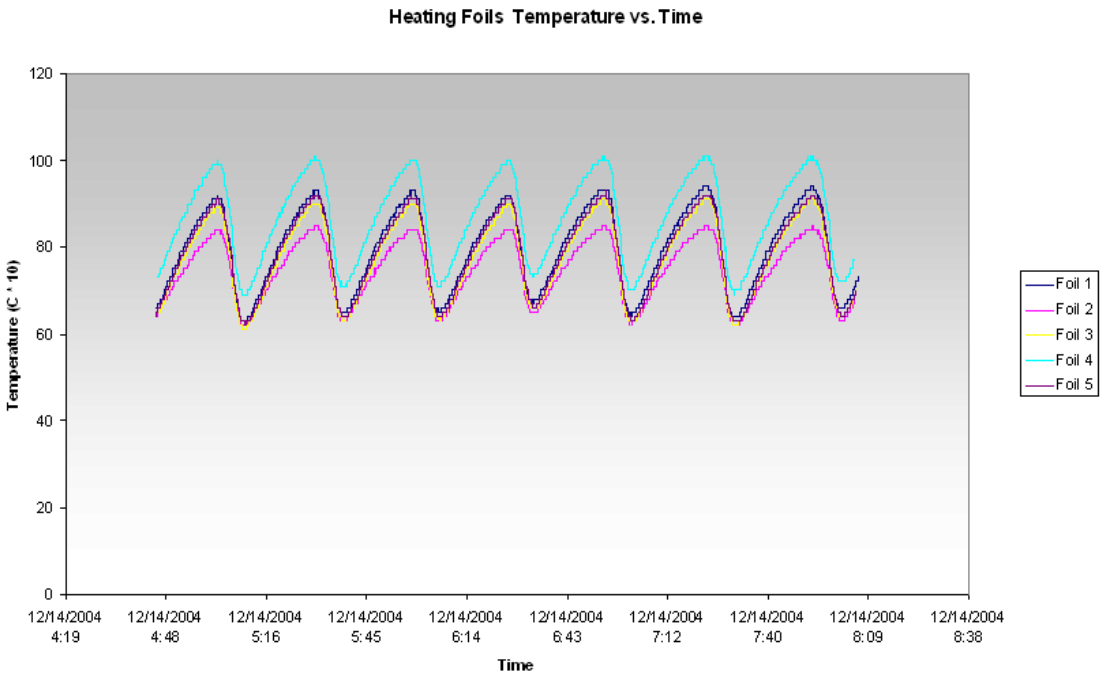


Figure 5-18 Heatig Foils behaviour

From these curves one can see that heating foils number two and four exhibit a behaviour slightly different from the others: foil 4 spans from 7 to 10 °C, while foil 2 ranges from 6 to 9 °C. In other words, the maximum temperature of foil four and two is, respectively, 1 °C higher and 1 °C lower than the others. Only further tests on similar panels will let us fully

understand the reason for this small discrepancy that does not affect, anyway, the overall performances.

We have also simulated the malfunctioning of an individual heating foil by disconnecting it from the power supply and trying to keep the system at the nominal temperature, with good results: even though the time constant was longer, the panel could keep up to the controller.

The operational conditions become critical if we disconnect two foils: in this case, the available heating power is not sufficient to track the set point; thus, with two malfunctioning heating foils , the system is in danger, the correct functioning is not assured any longer, and an interlock (or a warning/alarm) must be introduced to prevent the system from being damaged.

The curve in figure 5-20 shows the behaviour of the support tube temperature when three heating foils are turned off: the support tube temperature falls rapidly, and in about seven hours the temperature drops from 13.5 ° C to 5.5 ° C.

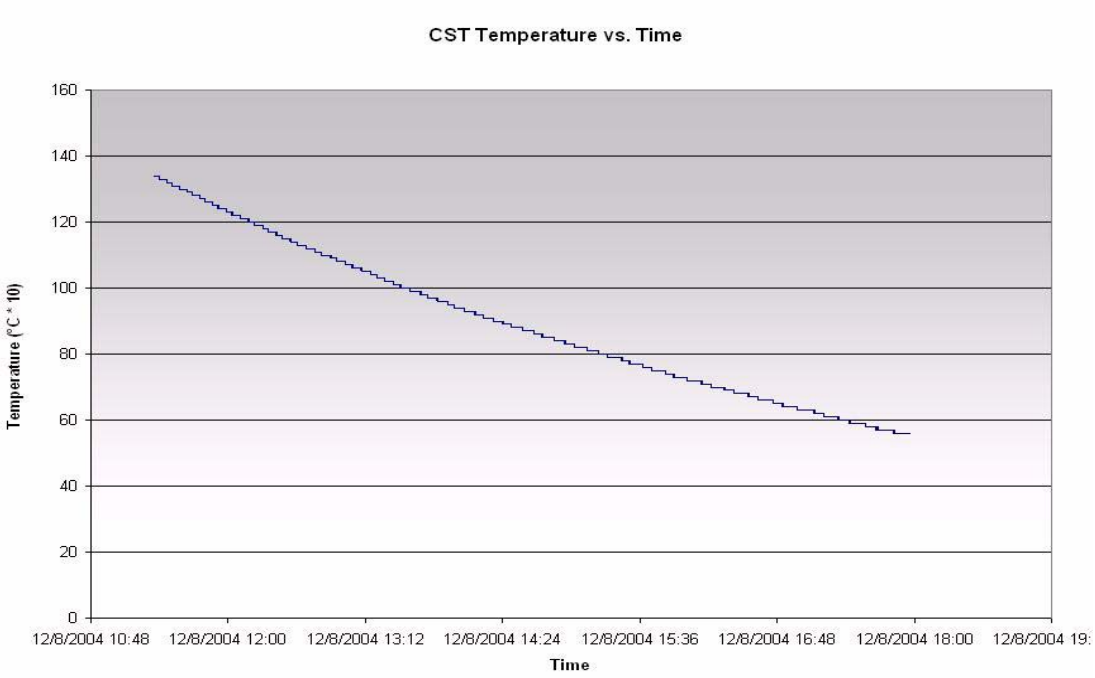


Figure 5-19 CST Cooling Down

5.5. Interlock tests

During the Tracker beam tests (May and September 2004), we had the opportunity to test a realistic copy of the interlock system needed for the Tracker, thus verifying the solidity of the choices made. In what follows we take a closer view on an interlock prototype built using the chosen hardware and software. On this prototype, the interlocks implemented were the ones dealing with power cuts and temperature excesses.

There were three detector groups participating to the test beam, reflecting the composition of the future Tracker. The detector groups equipped with samples of the TIB (Tracker Inner Barrel), TEC (Tracker End Caps) and TOB (Tracker Outer Barrel) detectors.

There were two complete subsystems mounted on final-like structures present, those of the TIB, participating with three, not completely full, ladders and the TEC, participating with an incomplete petal. while The TOB participated with final design detectors, but not yet mounted onto the "ROD" structure which will be the way it will be done in the final Tracker. All three detector systems had their own envelopes.

This detector segmentation in three basic groups will be the one to be addressed by the interlock system in the final Tracker (further segmentation in sub-groups will also be foreseen to be implemented in the system).

All three detector systems had their own separate power supplies both for biasing the detectors (referred to as High Voltage power supplies) and the Front End and control systems (referred to as Low Voltage power supplies). They were also equipped with different cooling and gas (N_2) circulation systems. The TOB and the TIB detectors were simply stabilized at temperatures below 20 ° C, while the TEC detectors were cooled to subZero temperatures.

All three systems were equipped with well-calibratedhumidity and temperature sensors that were appropriately conditioned and whose analog values, as well as the status of the conditioning electronic boards, were continuously scanned by the PLC. If any of the analog signals exceeded the set limits, or if there was a power failure, the PLC would respond by an interlock or a warning signal, depending on the user selection. The PLC was ramping down and then preventing automatic ramp-up of the TIB detectors in case of power failure and temperature excess, and issuing warnings for the TEC and the TOB detectors in case of humidity or temperature excesses.

The temperature sensors used were the ones chosen for the Tracker, thermistors for the TOB and TEC and Pt1000s for the TIB. The only significant difference between the beam test prototype and the Tracker interlock system was the fact that the logic of avoiding an abrupt OFF state for the power supplies was implemented through the PLC and not built in the power supply, as will be the case in the future and, obviously, the size of the system.

It nevertheless helped to find out that the test power supply would not ramp down properly when an interlock signal was present but rather set itself abruptly to OFF. This was modified on the spot by the system expert.

We used a subset of the thermal screen PLC configuration (see fig. 5-20): a medium range PLC of the SIMATIC S7 family, equipped with a CPU 15 (6ES7 315-1AF03-0AB0), an Ethernet module (6GK7 343-1EX00-0XE0), two twelve bit analogue input modules (6ES7 331-7KF02-0AB0 and 6ES7 331-7KF012- 0AB0), one digital input module (6ES7 321-1BH01-0AA0), and one digital output module (6ES7 322-1BH01-0AA0). The programming of the PLC was done in STL, and the PLC itself was powered through an Uninterruptible Power Supply (UPS) unit; the program was run cyclically, reading the values of four calibrated humidity and six calibrated temperature sensors as well as the output of a DC 24V power supply connected on the same power line as the Power Supply system.

The interlock was generated during the cyclic execution and was not event driven. The reason for that was the short duration of the PLC cycle and the fact that another Uninterruptible Power Supply was feeding to the power supply system.

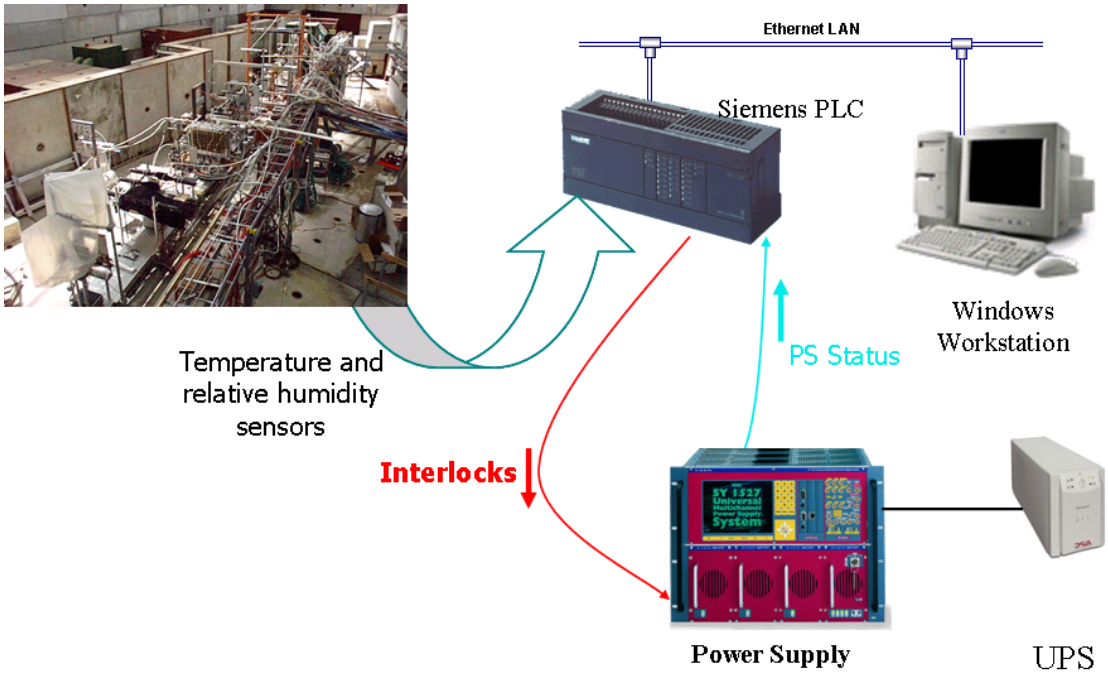


Figure 5-20 Beam Test Experimental Setup

The outputs, driven by the PLC logic shall be driving the interlock lines of the power supplies of the Tracker. Each power supply crate hosts 9 modules, each of which can receive any subset of the four different interlock signal inputs to the crate. It means that all the high and low voltage channels in one crate can be addressed by four interlocks that can be independent. The agreed-upon logic with the company producing the power supplies is *normally closed*. In Figure 5-22 one can see a power supply with modules and the grey cable transmitting the interlocks to the crate (grey cable).

Figure 5-21 CAN power supply with the interlock line

As for the interface, we used the Siemens native interface for small systems (ProTool/Pro environment). Data were transferred to the interface via the Siemens OPC server with the transfer of the complete variable name from the Step 7 symbol editor. In this way, integration of any Step 7 application is flawless.

In the next figures one may see characteristic snapshots of the interface. The interface starts with a display dedicated to the state of the whole three detector system (Fig 5-22)

Every area is "active": it means that when the user notices the pointer changing shape, it means that by clicking on that area another window will open, taking the operator to a more detailed view of the corresponding sub detector. The main screen also has room for an alarm window on top.

Every sub-detector page contains data by means of trends and gauges related to the parameters of interest. An "archive" button causes a Microsoft Excel window to automatically open and let the operator study the past performances up to 10 days back

.

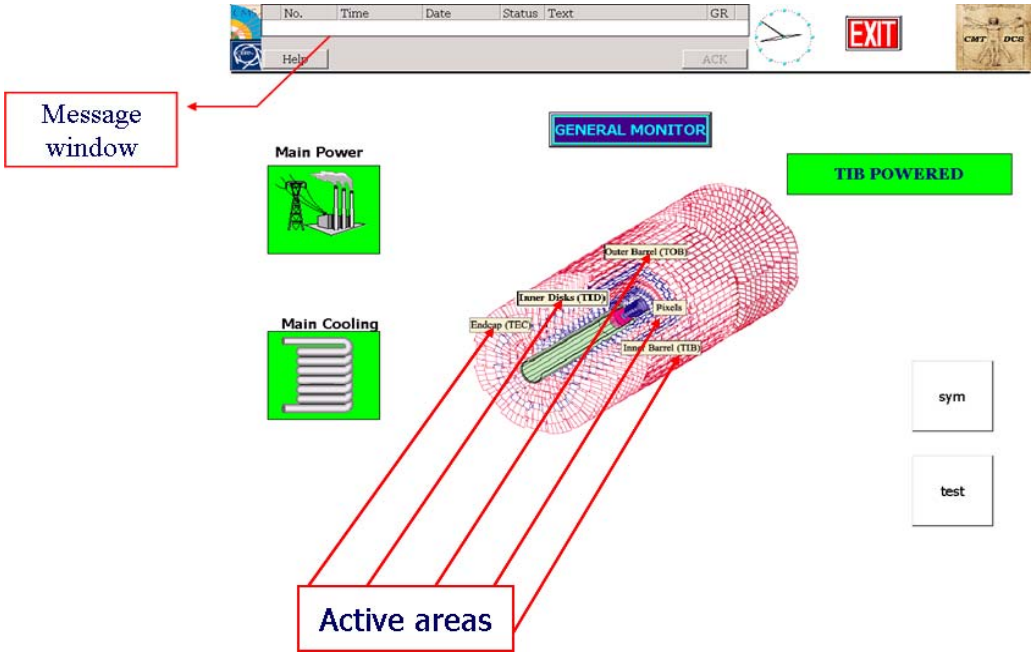


Figure 5-22 Beam Test HMI Main

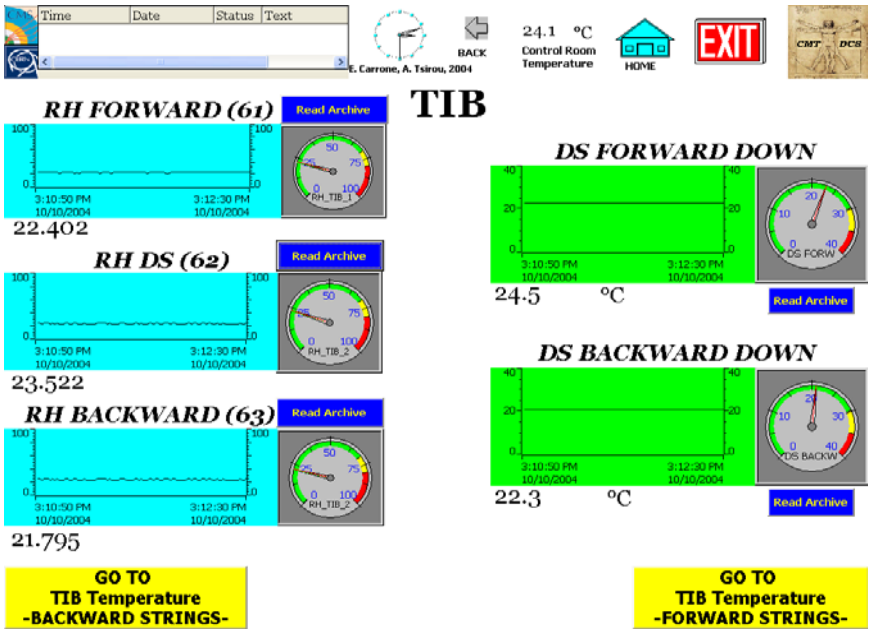


Figure 5-23 TIB HMI

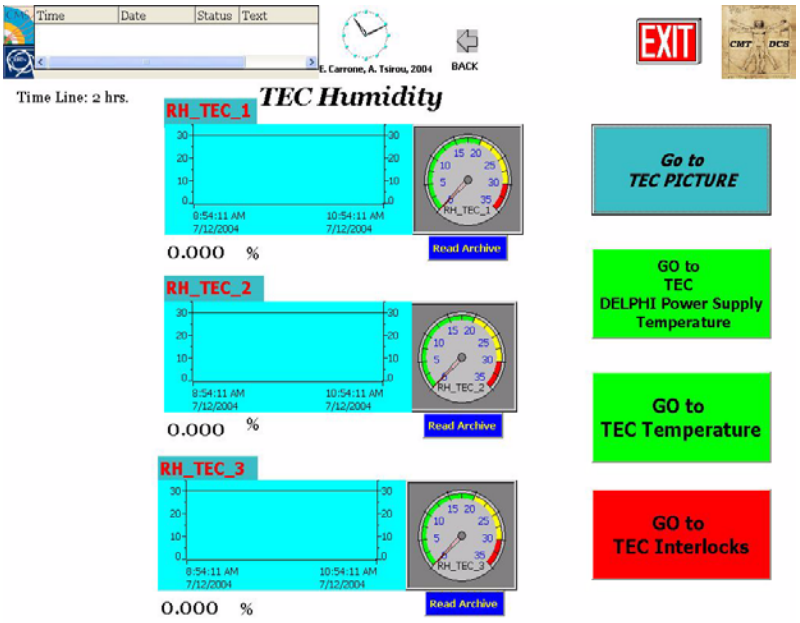


Figure 5-24 TEC HMI

The results of these tests show that the Tracker can very well be protected by means of an effective interlock system based on the very same hardware components used for the thermal screen and following the same programming logic and algorithms.

This test serves as a general tracker interlocks validation and, at the same time, as a thermal screen interlock test. In fact, the thermal screen control system will be the final resource in case of a cooling system failure; therefore, the interlocks that will be issued by the thermal screen PLCs will be of the same kind of the Tracker general interlocks.

Such an homogenuity in design and components is also successful for an easier maintenance, documentation and programming itself, and works as a solution to a problem that could be overly complex because of the large number of variables to be controlled and of the different needs and specifications of the various sub-detectors.

Conclusions

This thesis aims to prove that the thermal screen, a device which thermally isolates the CMS Tracker from the ECAL, is fully controllable by means of a PLC.

The goal was to model the system, design a PID controller and tune it, program the PLC and the human machine interface, install and commission hardware and software (included temperature and humidity sensors, power supplies and human-machine interface) and, eventually, plan and perform acceptance and quality assurance tests using a climatic chamber. Using the test setup, a system identification was performed, and the controller tuning checked.

All the above activities have been successfully performed: the system was studied using a model-based technique, where the process model has been derived with a thermal-electrical analogy representation; then, a controller has been designed and different tuning methods have been used.

While the PLC rack was fully assembled and all the cables routed and connected, a climatic chamber was built to simulate the real operating condition of one thermal screen panel at time.

System identification has been performed by exciting the process with a random signal and studying the output using the MatLab SysId toolbox. A Siemens PLC has been programmed, and an FB51 library has been configured to parametrize the PID. The same setup has been used to check and compare the controller performances, and to decide that the IMC tuning is the more suitable. In this way, we were able to prove that the design constraints were fully satisfied and that the control apparatus could exit the test phase and be ready for the final installation in the CMS underground control room .

We also successfully checked the control hardware and software on a set of tracker test modules exposed to a pion beam,connecting the interlock lines to the power supplies, while monitoring temperatures and humidity. There, we also provided a Human Machine Interface similar to the final one, and we proved that The tracker can actually be kept in a safe state via a PLC station that will control the thermal screen and the interlocks.

The whole project has been managed and documented using a systems engineering approach, following international standards (IEC, IEEE): the documentation includes a Systems Engineering Management Plan, Design Plan, Test and Acceptance Plan, User Requirements Document, Risk Analysis.

Future Trends

Control systems for temperature control are widespread; yet, there is still room for improvement, and the proof is the constant interest of the process control community to more effective controllers tuning. The efforts are in the direction of simplifying the tuning procedures, since they always prove to be quite time consuming and not always easy to be performed (since they could take the system to instability, which is not always feasible). This is why techniques like that of Skogestad, which rely upon the variation of a single parameter, are of big interest and triggering other attempts to find "simple tuning rules".

At the same time, techniques that make a heavy use of the large computational power exhibit good performances, like the IMC. The drawback is that this technique can not yet be used with all the PLCs commecially available: most of the times these controllers run on industrial workstations.

While the process control practitioners are looking for simple and reliable solutions, at the same time, the high energy physics community is witnessing a new trend in controlling particle physics detectors which strongly reliy upon PLCs. In this respect, needs of such kind of installations are often in between a classic PLC control and a totally PC driven one. The choice between the two approaches will be dictated by technical but, also, economic issues: it will be designers' responsibility to find an effective trade-off. The PLC can be used for many tasks, from interlocks to regulation. Therefore, it will be probably be the standard equipment as long as it will manage to provide satisfactory regulation capabilities (which is already happening, with IMC dedicated modules already on the market).

More on the project management side, the current (and future) size of HEP experiments will force to manage and document projects that are more and more complex in a more rigorous way; that is, in a Systems Engineering fashion. This will also help collaborations that show more and more an interdisciplinary approach, where different technical languages are spoken, and where the homogeneity must be given by design criteria and standards that find a common ground on coherent and well documented methodologies.

Appendix

A.1. Polynomial Representation of Transfer Functions

Rather than specifying the functions *G* and *H* in (Equation A-1) in terms of functions of the frequency variable ω, one can describe them as rational functions of q^{-1} and specify the numerator and denominator coefficients in some way.

A commonly used parametric model is the ARX model that corresponds to

$$G(q) = q^{-nk} \cdot \frac{B(q)}{A(q)}; \qquad H(q) = \frac{1}{A(q)} \tag{A-1}$$

where *B* and *A* are polynomials in the delay operator q^{-1}.

$$A(q) = 1 + a_1 q^{-1} + \ldots\ldots + a_{na} q^{-na}$$
$$B(q) = b_1 + b_2 q^{-1} + \ldots\ldots + b_{nb} q^{-nb+1} \tag{A-2}$$

Here, the numbers *na* and *nb* are the orders of the respective polynomials. The number *nk* is the number of delays from input to output. The model is usually written

$$A(q)y(t) = B(q)u(t-nk) + e(t) \tag{A-3}$$

or explicitly

$$y(t) + a_1 y(t-1) + \ldots\ldots + a_{na} y(t-na) =$$
$$b_1 u(t-nk) + b_2 u(t-nk-1) + \ldots\ldots + b_{nb} u(t-nk-nb+1) + e(t) \tag{A-4}$$

Note that (Equation A-3) and (Equation A-4) apply also to the multivariable case, with *ny* output channels and *nu* input channels. Then $A(q)$ and the coefficients a_i become *ny*-by-*ny* matrices,

and $B(q)$ and the coefficients b_i become *ny*-by-*nu* matrices.

Another very common, and more general, model structure is the ARMAX structure

$$A(q)y(t) = B(q)u(t-nk) + C(q)e(t) \quad \textbf{(A-5)}$$

Here, $A(q)$ and $B(q)$ are as in (Equation A-5), while

- $$C(q) = 1 + c_1 q^{-1} + \ldots + c_{nc} q^{-nc}$$

An *output-error* (OE) structure is obtained as

$$y(t) = \frac{B(q)}{F(q)} u(t-nk) + e(t) \quad \textbf{(A-6)}$$

with

- $$F(q) = 1 + f_1 q^{-1} + \ldots + f_{nf} q^{-nf}$$

The so-called *Box-Jenkins* (BJ) model structure is given by

$$y(t) = \frac{B(q)}{F(q)} u(t-nk) + \frac{C(q)}{D(q)} e(t) \quad \textbf{(A-7)}$$

with

- $$D(q) = 1 + d_1 q^{-1} + \ldots + d_{nd} q^{-nd}$$

All these models are special cases of the general parametric model structure.

$$A(q)y(t) = \frac{B(q)}{F(q)} u(t-nk) + \frac{C(q)}{D(q)} e(t) \quad \textbf{(A-8)}$$

The variance of the white noise $\{e(t)\}$ is assumed to be λ.

Within the structure of (Equation A-8), virtually all of the usual linear black-box model structures are obtained as special cases. The ARX structure is obviously obtained for $nc = nd = nf = 0$. The ARMAX structure corresponds to $nf = nd = 0$. The ARARX structure (or the *generalized least squares model*) is obtained for $nc = nf = 0$, while the ARARMAX structure (or *extended matrix model*) corresponds to $nf = 0$. The output-error

model is obtained with $na = nc = nd = 0$, while the Box-Jenkins model corresponds to $na = 0$.

The same type of models can be defined for systems with an arbitrary number of inputs. They have the form

$$A(q)y(t) = \frac{B_1(q)}{F_1(q)}u_1(t-nk_1) + \ldots + \frac{B_{nu}(q)}{F_{nu}(q)}u_{nu}(t-nk_{nu}) + \frac{C(q)}{D(q)}e(t)$$

A.2. CMS Cooling Plant

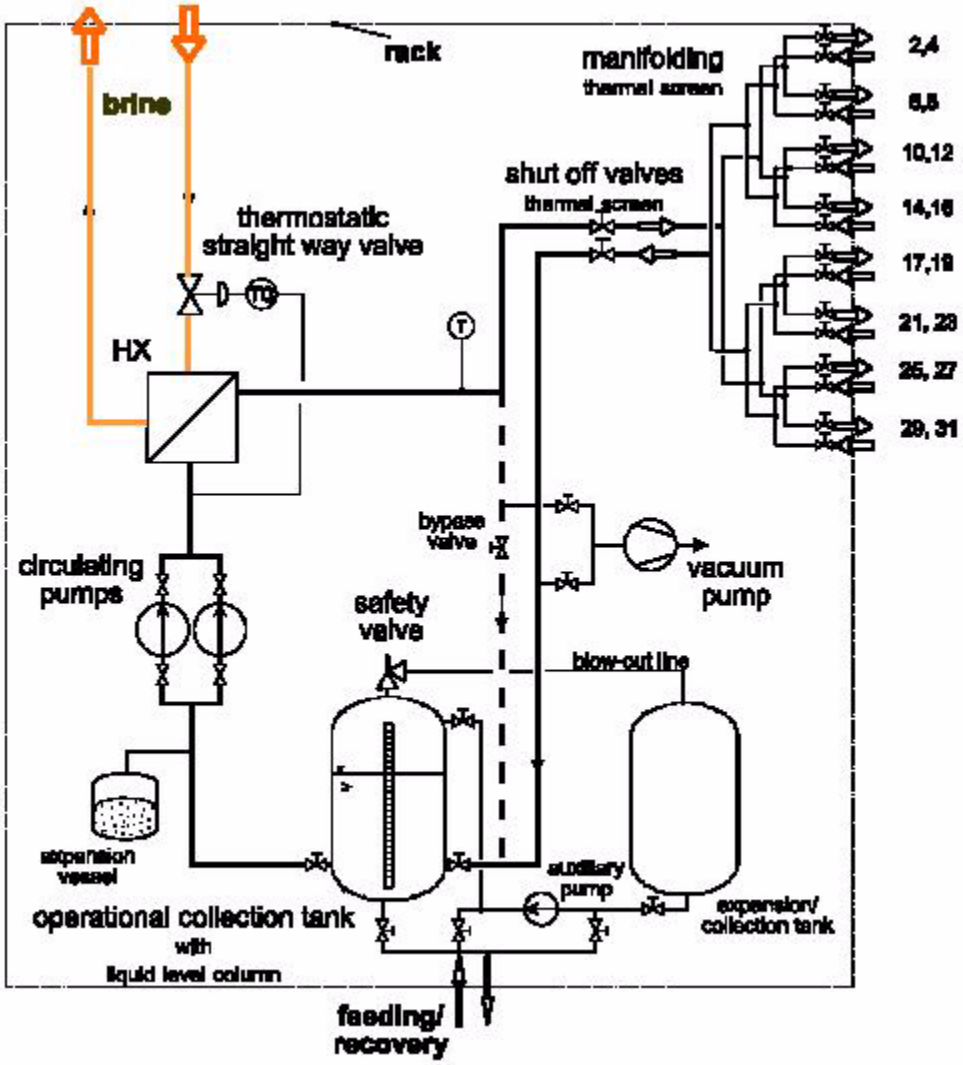


Figure A-1. Thermal Screen Pumping Station

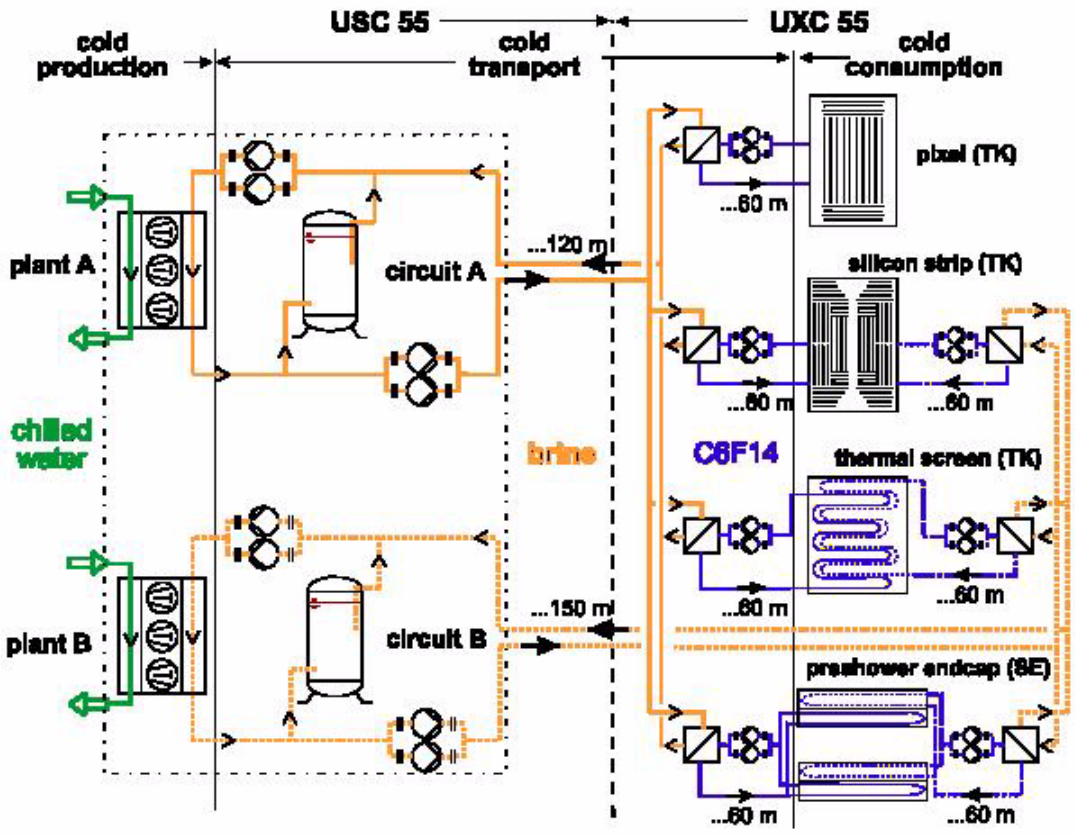


Figure A-2. CMS Cooling Plant

cooling task	refrigeration load $\dot{Q}_{0,max}$ kW	operation temperature $t_{operation}$ °C	inlet and outlet temperatures of cooling liquid t_{inlet} / t_{outlet} °C	possible part load conditions and load fluctuations
silicon detectors (silicon strip tracker and pixel detector)	…100	- 10 [1)]	- 20 / - 17	• winter shut-down (Nov. – April): no experimentation, read out electronics are switched off 30…50 % refrigeration load
preshower	…20	- 6 [1)]	- 14 / - 11	• experimentation (April – Oct.) procedures of experimental work and possible interruptions e.g.: beam dump (few min.) acceleration (few min.) adjustment of collision (…1h) injection and filling (…10h) physics data taking (…7h) 30…100 % refrigeration load
thermal screen (cold plate)	…10	- 10 [2)]	- 11,5 / - 8,5	in operation throughout entire year …100 % refrigeration load

[1)]…temperature of "Si-wafers" during operation
[2)]…average temperature on external surface

Figure A-3. CMS Tracker Cooling Requirements

A.3. System Identification with MatLab Toolbox

The MatLab System Identification toolbox is used to analyze the data gathered from a process excited with, for instance, a random input.

Using the GUI, one can generate discrete or continuous linear models that accurately describe the input-output behavior of the system using experimental data. These types of models are widely used in control applications where a simple but accurate plant model is required.

There are two main steps to the system identification process. The first one is to import our experimental data into the tool. The data to be used for identification can be in time- or frequency-domain. The second step is to create different types of models until a satisfactory one is found. This step is usually an iterative process.

In figure A-1 is shown the GUI.

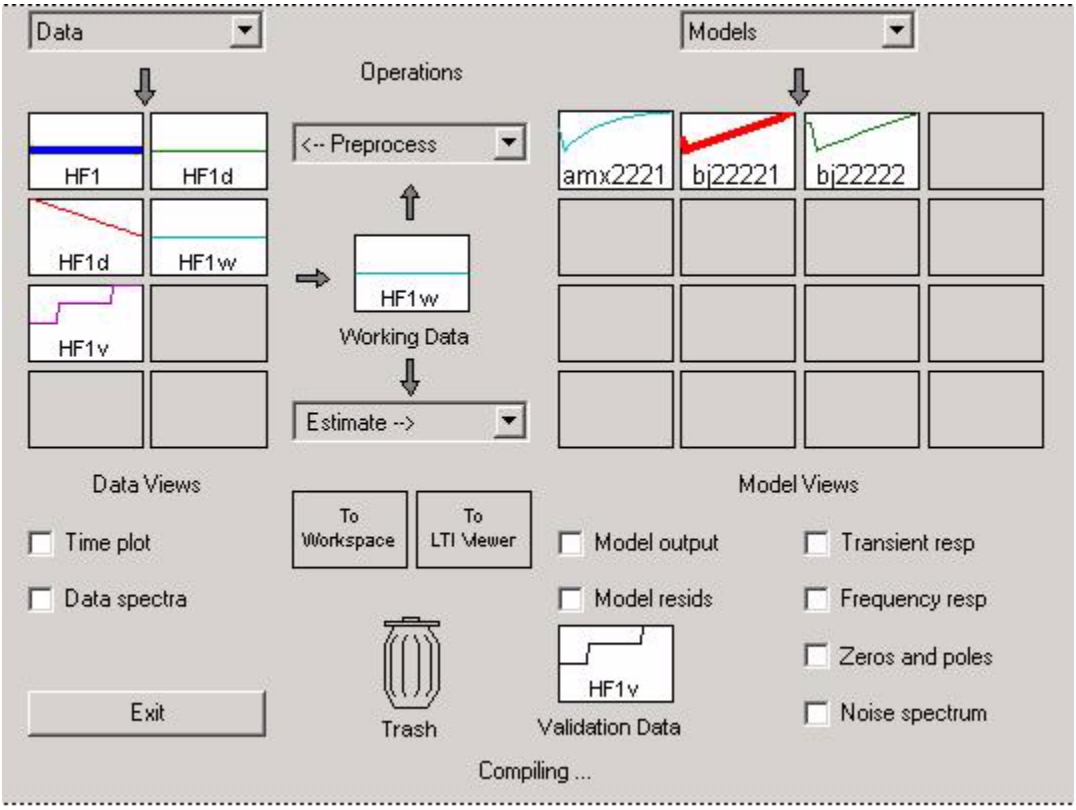


Figure A-4. Ident GUI

After that, the user is to define the so-called Data Sets. This is done by clicking on the "Import data" menu and selecting "Time domain data..." item. This will open a new dialog called "Import Data" dialog, where we can specify our data. The resulting data set would then be added to "Data Views" panel.

We can plot these data sets by clicking on the "Time plot" checkbox and then selecting the desired data icons in the GUI. The top plot shows the temperature from our experiments, and the bottom one shows the voltage command signals.

Now we are ready to create linear models using our data. The System Identification Toolbox is capable of creating both continuous and discrete models of our system.

In order to create a new model, we first need to select a "Working Data" set. For our demo, we will use the third data set (data3) which resulted from driving the DC motor with a short pulse. We will also use this data set for validation.

We will first create a number of discrete transfer function models. To do this we go to the "Estimate -->" menu and select "Parametric models..." item. This opens a new dialog called "Parametric Models", where we can specify various model properties.

If the desired model order is not known, you can press the "Order Selection" button that will let you specify a range of model orders. We will specify up to fifth order models with a time delay of five to ten time steps. Various other options can also be specified in this dialog:

Focus: Simulation

Initial state: Zero

Pressing "Estimate" button will generate these models and will open a "Model Structure Selection" dialog, where we can select the models that provide the best fits.

Now we can compare the response of these models with our experimental data by selecting the "Model output" checkbox. As you can see, all of these models provide a good fit with our data. In particular, the second order model with a delay of ten time steps has the best fit (arx2210). By double-clicking on this model, we can also see its transfer function.

Using the System Identification Toolbox, one can also generate continuous-time models. To do this we again go to the "Estimate -->" menu and select "Process models..." item. This opens a new dialog called "Process Models", where we can specify various model properties.

We will create a transfer function model with two underdamped poles and a time delay. Various other options can also be specified in this dialog:

Name: P2DU; Focus: Simulation; Initial state: Zero.

Pressing "Estimate" button will generate this model and place it in the main GUI. We can plot the response of the resulting model by re-opening the "Model Output" plot.

Once we are satisfied with a model, we can export it to the MATLAB Workspace by dragging it onto the "To Workspace" region in the GUI. This model can then be used in MATLAB or imported into our Simulink model.

To import this model into our Simulink diagram, we specify the model name in the model block. Then we can simulate this system to make sure that its response matches with the experimental data.

A.4. Initial Risk Evaluation

A.5. HAZOP Analysis

A.6. Event Tree

A.7. Test plan

ANSI/IEEE Standard 829-1983 describes a test plan as: "A document describing the scope, approach, resources, and schedule of intended testing activities. It identifies test items, the features to be tested, the testing tasks, who will do each task, and any risks requiring contingency planning."

The Test Plan describes the approach to all development, unit integration, system qualification and acceptance testing needed to complete a project properly. Establishing a test plan based on business requirements and design specifications is essential for the successful acceptance of a project's deliverables.

It is important to note that the higher risk a project has, the greater the need for a commensurate amount of testing. The project Schedule & Task Plan and the project Staffing Plan need to account for testing requirements during the planning and execution phases of the project.

Testing validates the requirements defined for the projects objectives and deliverables.

The first test is a static one, on the thermal screen panel: test of the temperature sensors.

Then the power supply will be tested.

Problems will be reported in the "Problems report" sheet.

Test results will be recorded in Microsoft Excel and/or Word format.

Problems will be tracked to resolution with Controls Team meetings.

Resource Requirements

❖ Hardware requirements:

1. Xantrex power supplies
2. Digital multimeter
3. RTD measuring device (Fluke)
4. 2-wire cables for power transmission
5. 4-wires cables for RTD and voltage readout
6. Climatic chamber for a thermal screen panel
7. Thermal screen panel fully assembled
8. Cooling plant
9. Windows workstation for data recording
10. PLC for PID testing
11. Cable labels
12. Panel enclosure (wooden box)

❖ Software requirements:

1. Microsoft Windows
2. Siemens Step 7
3. Siemens PID module
4. Siemens ProTool/Pro

❖ Documentation requirements :

1. Power supplies specifications
2. Heating foils specifications
3. PLC manuals.
4. Power supply manuals.

❖ Test environment :

A test bench is has been set in Building 186. The climatic chamber is a wooden box, thermally insulated, which will accommodate one thermal screen panel at time. The final PLC rack is installed. The cold plate must be connected to a heat exchanger.

The following schematic shows the electrical components of the heating foils:

❖ Identification of tests (One test case each):

1. Power supply. Behaviour when connected to one heating foil via 30 meters long cable.
2. Power supply. Behaviour when connected to more than one heating foil (namely two and three heating foils) via 30 meters long cable.
3. Heating foil. Measured resistance of the heating foil.
4. PID-Tuning and performances. Analysis of the behaviour with one, two three, four, five panels connected.
5. RTD.
6. Influence of temperature on RTD readings.
7. Mutual influence of adjacent panels in terms of temperature.
8. Temperature changes vs. time.
9. PID tuning. Analysis of the PLC response curves.

❖ Acceptance Test Report :

The Acceptance Test Report is actually the Test Scenario Specification Form [1]. The last version of this report closes out the acceptance process.

❖ Corrective action :

Failing cases will be discussed in group meetings, and corrected or re-tested where necessary. This process is iterative, until each test case has successfully executed.

❖ Test Cases (IEEE Std. 829:A set of test inputs, executions, and expected results developed for a particular objective):

The following Test Scenario Specification Form provides a description of individual test cases.

1.

[Thermal Screen Control System] [Panel] [Item]
Test Scenario Specification Form

Version No.:	Panel #:	Requirement #:	Page:
Release No.:	Test Scenario #:	Environment:	of
	Retry #:	Object tested:	

Written By	Date	Reviewed By	Date	Executed By	Date	Retry #

Instructions:

Test Scenario Objective:

Assumptions/Dependencies:

Test Files/Test Data:

STEP#	DESCRIPTION	EXPECTED RESULT	ACTUAL RESULT	

Figure A-5. Test Scenario Specification Form

The following is a description of the content of key fields.

Version No.: The version number of the item being tested.

Release No.: The release number of the item being tested.

Panel #: The tracking number associated with the panel.

Test Scenario#: A number (sequential) assigned to this test case for tracking purposes.

Requirement #: The number of the requirement that will be proven by this test case.

Environment: e.g., panel inside/outside the climatic chamber.

Object tested: e.g., PLC, RTD sensor, heating foil, power supply.

Retry #: A sequential number representing the number of times the test case has been executed.

STEP#: A number (sequential) associated with the step to be performed in the test scenario.

Description: What happens in this step

A.8. Acceptance Test

The scope of the acceptance test is to test the power supplies, test the heating foils and the pt1000s which constitute all the items that take part, directly or indirectly in the control process.

In order to perform these tests, every thermal screen panel will be cooled through a Freon circulation circuit, the heating foils will be powered by their own power supply, and the environment will be thermally insulated.

The following items have to be tested:

1) Temperature Sensors

Verify the existence and correct reading of pt1000.

There should be one pt1000 per heating foil (a total of five per panel). The pt1000s should be measured and their recorded temperature should be cross checked with the temperature recorded by another pt1000 in the same environment. This is why this test should take place before the panels enter the climatic chamber. When exposed in the open of the storage room, any difference larger than 1^{0}C from the ambient should be considered as an indication of a problem. The expected duration of the test should be one working day (20 sensors/hour).

2) Power supplies.

Verify the performance of the power supplies.

Verify the capability of the power supply of being driven through the CURRENT output of the PLC. That means performing the following tests, for which the PLC-PC system is needed. No insertion in the environmental chamber is needed.

Verify the curve **Icon vs. Vout**, where Vout is PLC monitored. Expected needed time for that (if the PLC program exists): 1 working day (4 power supplies/hour). Very few points are needed in the curve. The importance is to verify the same endpoints of the plot.

Verify the voltage drop along the ~150m distance between the thermal screen and the power supply rack, using an appropriate 150meters long cable.

3) Heating foils.

Verify the thermal foil integrity and heating circuit distribution

Verify the continuity of the thermal foils. This should be done before the insertion in the environmental chamber, using a voltmeter. This is a simple test that should take about 30 minutes before the insertion of the panel in the environmental chamber.

Check the integrity of the heating captons up to the PLC distribution bar ("to the "weidmuller" connector) to prevent any problem with connectors and cables.

Test the heating foils interaction in different configurations. This should be tested after the insertion of the foil in the environmental chamber and while the cooling is on.

The procedure is the following:

We have to attach at pt100 or a pt1000 on the carbon fiber cover of the cooling panel-insulator-heating foil system. This is the carbon fiber panel outside which the ECAL will start, therefore any temperature sensor attached to it shall measure the same way it will when the thermal screen is inside the detector. If the carbon fiber to be provided by the engineering group covers the whole Thermal screen panel, we should use a temperature sensor per heating foil, situated on the carbon fiber and approximately in the middle of each heating circuit. We should monitor the temperature of the 5 sensors attached to the carbon fiber as well as the temperature recorded by the 4 pt1000s attached to the heating circuits (the 5^{th} is the process variable) while attempting to have the system reach 18^{0}C (provided it is possible). We should attempt this individually using circuits one, two,…,five, the using circuits one and five, two and four, one, three, five and all of them.

After this test is completed, we should study the temperature distribution using two power supplies and two PIDS and studying the temperature distribution for two set point temperatures: circuits one, two and three should heat the system to 18^{0}C, while circuits four and five would heat the system to 20 ^{0}C. This is a test to be performed once and we expect to spend two weeks on it. From this we expect to have enough data for the temperature profile outside the carbon fiber and what would be the effect of using a second, independent power supply per thermal screen panel.

4) PID

Verify the PID parameters and tuning

Verify the tuning of the PID controller. Identify the variables, in the Siemens PID control software and specify the ones that are of any relevance

5. Test the whole PLC configuration.

6. Optimize the PC algorithm

A. The level at which testing is performed is:
System level
Component level
Integration level.

B.Test methods, tools, harnesses and procedures to be used:
Voltage ramp will be applied to the heating foils, one by one.
Voltage output will be read at the power supply output and at the heating foils input, along with the controlling voltage.
Temperature sensors RTD will be checked by means of a dedicated instrument.

C. Source(s) of test data
Power supplies manuals and specifications
Thermal screen panels specification files.
Instruments reading
PLC reading (Human Machine Interface display, PLC operator panel, Microsoft Excel type archives)

Bibliography

[1] P. Burkimsher, H. Milcent. "Applying industrial solutions to the control of HEP experiments." International Conference on Accelerator and Large Experimental Physics Control Systems, Trieste, Italy, 1999.

[2] D. Blanc, "Distributed control software for high performance control loop algorithm." International Conference on Accelerator and Large Experimental Physics Control Systems, Trieste, Italy, 1999.

[3] The Tracker Project Technical design report, CERN/LHCC 98-6, Geneva, Switzerland, 1998.

[4] Operator screen (HMI) design guidelines. Hexatec, Ochrelands Hexham Northumberland, England, 2002.

[5]] A. Canning, G.T. Moran, S.J. Clarke, D. Maisey, S.Pegler, D.Hedley, "Sharing idea - the SEMSPLC project", Software for engineers, supplement, March 1994, IEEE.

[6] G.F. Knoll, "Radiation detection and measurement", Third edition, J. Wiley and sons, 1999

[7] D. Groom, "Energy loss in matter by heavy particles", Particle Data Group Notes, PBG-93-06, 1993.

[8] CERN EDMS CMS-TK-ED-0162, Oct. 2002.

[9] CERN EDMS CMS-TK-CI-003, June 2003.

[10] BaBar Collaboration, "Bound on the Ratio of Decay Amplitudes for $B \to J/\psi K^{*0}$ and $B^0 \to J/\psi K^{*0}$", Phys. Rev. Lett. 93, 081801 (2004).

[11] A. Ghandakly, M.E. Shields, A.M. Farhoud, "Enhancement of existing PLC's with an adaptive control technique", IEEE Industrial Application Society annual meeting, October 1995.

[12] M. Tzhorn, "The role of knowledge based systems in fault diagnosis in applications using human operators", Human Interaction with Complex System, Kluwer, 1996.

[13] N. Weston, L.F. Baxter, J.E.L. Simmons, "Software design: concept generation as part of formal and informal process", IEEE Transactions on Engineering Management, Vol. 45, N. 4, November 1998.

[14] M. Klein, D. Norton, "Increasing the productivity and profitability of machine builders using PC-based control", IEEE?...

[15] L. Zhi, J. Shi Quin, T. Bao Yu, Z. Jun Hu, Z. Hao, "The study and realization of SCADA system in manufacturing enterprises", Proceedings of the 3rd WOrld COngress on Intelligent COntrol and Automation, Hefei, P.R. China, June 28 - July 2, 2000.

[16] A. Ramirez-Serrano, S.C. Zhu, S.K.H. Chan, S.S.W. Chan, B. Benhabib, "A hybrid PC/PLC architecture for manufacturing system control-implementataion", IEEE International Conference on Systems, Mand and Cybernetics, Nashville, USA, October 2000.

[17] "ISA S88.01-1995 Batch Control -Part 1: Models and terminology", International Society for Measurement and Control.

[18] "IEC 61512-1-1998 Batch Control - Part 1: Models and terminology", International Electrotechnical Commission.

[19] "IEC 61511-1-2000: Functional safety: Safety Instrumented Systems for the process industry sector", International Electrotechnical Commission.

[20] Seok Kim H., Young Lee J., Hyun Kown W., "A compiler design for IEC 1131-3 standard languages of programmable logic controllers", Proceeding of SICE 99, Morioka, 1999.

[21] "Safety Integrity Level Selection, Systematic Methods Including Layer of Protection Analysis, Editor P.E. Marszal and E. Scharpf, ISA Instrumentation Systems andsAutomation Society, 2002.

[22] V.M. Becerra, "Lectures on automatic tuning", University of Reading, United Kingdom, February 2003.

[23] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flanery, "Numerical recipes in C", Cambridge University Press, Cambridge, 1996

[24] S. Skogestad, "Simple analytic rules for model reduction and PID controller", Journal of Process Control, June 2002.

[25] D.J. Lartz, H. H. Cudney, T. E. Diller, "Heat Flux Measurement Used for Feedforward Temperature Control," Proceedings of the 10th International Heat Transfer Conference, vol. 2, Brighton, UK, pp. 261-266, 1994.

[26] Kok Kiong T., Qing-Guo W., Chang Chieh H., “Advances in PID control”, Springer-Verlag, London, 1999.

[27] G.K McMillan, “Tuning and control loop performance”, Research Triangle Park, NC: ISA - The Instrument Society of America, 1994.

[28] M. Katayama, T. Yamamoto, Y. Mada, “A practical design scheme of multiloop robust PID control systems”,

[29] K.E. Haggblom, “Control structure selection via relative gain analyis of partially controlled systems”, AICHE Annual Meeting, Los Angeles, 1997.

[30] P.J. Campo, M.Morari, “Achievable closed-loop properties of systems under decentralized control: Conditions involving the steady-state gain", IEEE Trans. Autom. Control, 39, 932-943 (1994).

[31] M. Hovd, S. Skogestad, “Simple frequency-dependent tools for control system analysis, structure selection and design", Automatica, 28, 989-996 (1992).

[32] M. Tham, “Multiloop systems (relative gains)”, Lectures at the Department of chemical and process engineering, University of Newcastle upon Tyne, 2002.

[33] “Dynamic Behaviour of Buildings”, Environmental Engineering Science 2 Course at the University of Strathclyde, Glasgow, Scotland, 2002.

[34] “Comparison of a Single-Loop Digital Controller to a DeltaV Multi-Loop Controller”, Emerson Process Management Ltd., White Paper, April 2001.

[35] O.H. Bosgra, H. Kwakernaak, G. Meinsma, “Design methods for control systems”, Notes for a course of the Dutch Institute of Systems and Control Winter term 2002–2003

[36] A. Mader, “A classification of PLC models and applications”, WODES Workshop on Discrete Event Systems, Gent, Belgium, August 2000.

[37] S. Bornot, R. Huuck, Y. Lakhnech, B. Lukoschus, “Utilizing static analysis for Programmable Logic Controllers”, ADPM 2000, International Conference on Automation and Mixed Processes: Hybrid Dynamic Systems, Dotrmun, Germany, September 2000.

[38] J. La Fauci, “PLC or DCS: selection and trends”, ISA Transactions, VOl. 36, No.1, pp. 21-28, 1997

[39] K.H. Johansson, “Relay feedback and multivariable control”, Doctoral dissertation, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, 1997.

[40] J. Mikles, M. Fikar, “Process modelling, identification and control I.Models and dynamic characteristics of continuous processes”, STU Press, Bratislava, 2000.

[41] C.D. Johnson, “Process control instrumentation technology”, Seventh Edition, Prentice Hall, Upper saddel RIver, New Jersey, 2003

[42] B.J. Kuo, F. Golnaraghi, “Automatic control systems”, Eighth edition, J. Wiley and sons inc., USA.

[43] E Carrone, “Design Specifications and test of the HMPID’s control system in the ALICE Experiment”, LEB2001,Stockholm, 2001.

[44] K. Dutton, S. Thompson, B. Barraclough, “The art of control engineering”, Prentice Hall, 1997

[45] E.O. Doebelin, “Engineering experimentation”, McGraw-Hill international editions, 1995

[46] T.E. Marlin, “Process control”, Second Edition, Mc Graw-Hill Higher Education, 2002

[47] C.S. Howat, “Proceaa HAzard Identification”, Lecture Notes, Kurata Thermodynamics Laboratory, University of Kansas, 2002.

[48] M.J. Baker, “Risk Modelling and Quantification”. Lecture N. 2 in :An engineer’s responsibility for safety”, Hazard Forum, London, 1996

[49] K. Ogata, “Modern control engineering”, Third edition, Prentice Hall, 1997

[50] A.M. Law, W.D. Kelton, “Simulation modeling and analysis”, Third edition, Mc Graw-Hill International Series, 2000

[51] “Process control engineering”, Edited by M. Polke, VCH, 1994

[52] Siemens Simatic S7-300 Programmable Controller Hardware and Installation, Order N. 6ES7 398-8AA03-8BA0, Siemens AG, Numberg, 1998.

[53] Siemens Simatic Standad PID Control - First steps in commissioning, Siemens AG, Numberg, 2003.

[54] G. Baribaud, Private communications, 2004

[55] G.R. Mosard, “A generalized framework and methodology for systems analysis”, IEEE tran. Eng. Manag., Vol. EM-29, 1982.

[56] M.L. Shoman, Software engineering: design, reliability, management", London, McGraw-Hill, 1983.

[57] J. Lewis, "Project Planning, Scheduling and Control", McGraw-Hill, 2000

[58] CMS, Technical Proposal, CERN/LHCC 94-38 LHCC/P1, 15 December 1994.

[59] LHC: The Large Hadron Collider Conceptual Design, CERN/AC/95-05 (LHC), October 1995.

[60] CMS, The Tracker Project, Technical Design Report, CERN/LHCC 98-6 CMS TDR 5, 15 April 1998.

[61] CMS Collaboration, Addendum to the CMS Tracker TDR, CERN/LHCC 2000-016, CMS TDR 5 Addendum 1, 21 February 2000.

[62] H. Berger, "Automating with STEP7 in STL", Erlangen and Munich, MCD Verlag, 1998.

[63] S. M. Shinners, "Modern control system theory and design", New York, John Wiley & sons, inc., 1988.

[64] "High Risk Safety Technology", Edited by A.E. Green, New York, John Wiley & sons, inc., 1982.