Tampere University Dissertations 48

ARTO NIEMI

Modeling Future Hadron Colliders' Availability for Physics

CERN-THESIS-2018-389
17/04/2019

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty Council of the Faculty of the Engineering and Natural Sciences
of Tampere University,
for public discussion in the auditorium FA032 Pieni sali 1
of the Festia building, Korkeakoulunkatu 10, 33720, Tampere,
on 17 April 2019, at 12 o'clock.

ACADEMIC DISSERTATION
Tampere University, Faculty of Engineering and Natural Sciences
Finland

Responsible supervisor and Custos	Professor Eric Coatanéa Tampere University Finland	
Pre-examiners	Doctor Annika Nordt The European Spallation Source Sweden	Professor Rebecca Seviour University of Huddersfield United Kingdom
Opponents	Doctor Annika Nordt The European Spallation Source Sweden	Professor Samuel Jean Bassetto Polytechnique Montréal Canada

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

Copyright ©2019 Arto Niemi

Cover design: Roihu Inc.

ISBN 978-952-03-1056-1 (print)
ISBN 978-952-03-1057-8 (pdf)
ISSN 2489-9860 (print)
ISSN 2490-0028 (pdf)
http://urn.fi/URN:ISBN:978-952-03-1057-8

PunaMusta Oy – Yliopistopaino
Tampere 2019

Abstract

This thesis presents an availability model for the Future Circular hadron Collider. The model is based on the current operations of the Large Hadron Collider. The thesis shows that hadron collider's availability for physics is a complex problem as it depends both on the system availability and the operational cycle. The availability itself is not the key performance indicator but it is closely linked to collision production and so-called integrated luminosity that are essential for physics research. The thesis shows that taking into account the operational cycle is critical for modeling luminosity production.

The thesis validates the model against LHC operations and shows preliminary results on the FCC availability and luminosity production. Ramentor Oy' ELMAS software was chosen as the platform for the model. ELMAS is designed as a fault tree software. However, the developed model combines fault trees with Markov models. This feature was implemented by adding custom Java code and libraries to models.

This reliance on a custom code was not the ideal solution and lead to the development of the OpenMARS approach. This approach allows combining the most common risk assessment and operation modeling techniques and connecting the models made with these techniques. This thesis presents the basic concept of this approach and shows how the collider operations model can be implemented with it. The discussion section also provides ideas on how the study would have proceeded if the collaboration with Ramentor Oy had not been an option and further applications of the OpenMARS approach.

Another issue encountered during the study was the lack of reliability data. This issue resulted in a study of reliability data sharing practices in the industry where the author was a key contributor. The thesis presents the idea to use a similar concept in the accelerator field and comments on this.

Preface

This study was carried out at the European Organization for Nuclear Research (CERN) as a part of the Future Circular Collider study. My work at CERN has taught me that the big science is a team effort. The work is carried out in collaborations and working groups. This was the case also in my project, and I would like to express my sincere gratitude to the people who contributed to my work.

The majority of the work was made in collaboration between CERN, Ramentor Oy and Tampere University of Technology (TUT). At CERN, key people were Andrea Apollonio, Johannes Gutleber, and Peter Sollander. Andrea introduced me to accelerator failures and fault tracking. His prior work was the starting point for the collider model and he gave essential help in writing the journal publication to the Physical Review of Accelerators and Beams with Johannes Gutleber. Johannes Gutleber was supervising the project from the FCC study and a key contributor in developing the OpenMARS approach that is one main of the outcomes of my work. Peter Sollander was my official supervisor at CERN. Through him, I had a chance to attend CERN Accelerator School at Budapest and Joint Universities Accelerator School that comprised the majority of my formal Ph.D. studies. I want to also express gratitude to numerous experts on the accelerator field that I had a chance to work with. + An extra special thanks for Estrella Vergara Fernandez who provided helpful comments to my thesis while I was writing it.

Ramentor Oy provided the ELMAS tool for the FCC availability study. Support and additional software features were provided by Jussi-Pekka Penttinen who was also partly working at TUT during the project. Later in the project, he was the main developer of the OpenMARS approach. At TUT, Seppo Virtanen was my first Ph.D. supervisor, and I am still amazed that I had a chance to work at CERN through Seppo. After his retirement Eric Coatanéa and Kari Koskinen took over the CERN collaboration at TUT and Eric was my Ph.D. supervisor. Before the CERN project, I was working at TUT for my master thesis project for Finnish aviation and later in other projects for Finnish industry. For this period, I would like to especially thank Jouko Laitinen, Per-Erik Hagmark, Henry Paajanen, Olli Salmela and Timo Galkin.

This work would not have been possible without funding. This thesis and its findings are part of the global Future Circular Collider Study hosted by CERN. The FCC Study has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 654305 (EuroCirCol), and this project has also received funding from the under grant agreement No 730871 (ARIES). In Finland, the work for the FCC availability study has received funding from Ramentor Oy and Tampere University of Technology.

Author's contributions

This thesis includes and extends material from several other publications to which the author has contributed. The following list briefly presents the author's contributions and how this thesis extends these findings. More details are provided in the chapter 10.

A. Niemi et al. Phys. Rev. Accel. Beams, 19:121003, 2016 [1]: This paper was written by the author and revised together with the co-authors. The paper presents the first version of the collider availability model; its validation with LHC operations data; and studies of different operation scenarios.

J.-P. Penttinen, A. Niemi et al. Technical Report CERN-ACC-2018-0006 [2]: This report was written in collaboration together with the other authors. It introduces and gives the specification for OpenMARS, which is an open modeling approach for availability and reliability of systems. The initial concept was developed in collaboration with all authors, and the final report is the result of combined work by the authors A. Niemi and J.-P. Penttinen.

A. Niemi et al. Technical Report CERN-ACC-2018-0013 [3]: This report was written by the author and revised together with the co-authors. It presents the requirements for the FCC-hh operation schedules and what these requirements mean for the cryogenic system maintenance.

A. Preinerstorfer et al. Technical report [4]: The author did not write the report, but it is a result of collaboration where the author is a key contributor. The report relates to a study on the accelerator system reliability database, and it presents a literature review on maintenance data collection.

J.-P. Penttinen, A. Niemi et al. submitted [5]: This paper was written in collaboration together with the main author J.-P. Penttinen. It gives a scientific introduction to the OpenMARS approach. Key contributions of the author were in the introduction, case study, discussion and conclusion chapters.

Future Circular Collider Study Volume 2, chapter 2.5 [6]: The author wrote the chapter in collaboration with A. Apollonio. It presents a brief summary of operations schedule, operations cycle and availability issues of Future Circular hadron Collider.

This thesis presents and extends these findings. For the collider availability model [1]: Chapter 3 provides extended background on what options would have been available for complex system availability modeling. Chapter 4 provides a background on how colliders operate to assess what features the model should have. The model itself is presented in the chapter 5. The publication [1] does not go into the details that are presented in this chapter, and the injection modeling has not been previously published.

The Future Circular hadron Collider operations schedule and shown in the report [3] is presented again in the chapter 7. This chapter also presents findings for the Future Circular hadron Collider operations cycle and availability. These findings have partially been presented in FCC week

presentations, in the technical report [7] and the summary of them is given in the conceptual design report [6].

The work for the FCC availability resulted in two other studies: OpenMARS approach and on accelerator reliability database. These are presented in the chapter 8. Section 8.1 gives an introduction to the open modeling approach for availability and reliability of systems (OpenMARS) with few original comments. This chapter also shows how the collider operations model could look in the OpenMARS. This work has not been previously published. For reliability database, the section 8.2 introduces the idea and provides few original comments.

This thesis also benefited from author's prior work in the field of reliability engineering. The author's master thesis [8] described software for aircraft system component reliability assessment based on maintenance data and the publication K. Mahlamäki, A. Niemi et al. [9] presented challenges in maintenance data collection for a model to assess extended warranty contract costs. These findings are partially reflected in the chapter 3.

Contents

1 Introduction

Currently, the world's most powerful hadron collider is the Large Hadron Collider (LHC) at CERN. The collider is 27 km long and it is designed to produce collisions at 14 TeV energy [10]. The LHC's most famous achievement is the discovery of the Higgs boson during 2012 [11, 12]. The discovery or exclusion of the Higgs boson was one of the primary goals of the LHC and its physics program is planned to continue to mid-2030's [13].

Large science infrastructure projects have long timescales. The LHC was initially proposed in 1983 [14] and currently, multiple studies are ongoing to determine what should be the post-LHC collider [15]. One of these projects is the Future Circular Collider (FCC) study [16]. Its main emphasis is on the conceptual design study hadron collider (FCC-hh) with a centre-of-mass energy of in order of 100 TeV. The plan is to locate the collider in a new 80-100 km circumference tunnel at Geneva area. More details are provided in the chapter 2.

FCC-hh has ambitious goals for physics production [17]. Reaching them requires maintaining the operational availability achieved in the LHC. This is a challenge, as the system complexity and stored energies are foreseen to be significantly higher than in the LHC. As a consequence, the FCC availability study was launched to ensure feasible operations and this thesis is a part of the availability study. During a conceptual design phase of the FCC, the main tasks of availability study were deemed to be: (i) understanding the key challenges in achieving the required availability and (ii) deriving availability goals for the accelerator complex. The effort to accomplish these tasks leads to questions:

- How to estimate and model collider's availability while taking into account the whole collider complex and the collider operations?
- Can the modern simulation approach be applied to modeling reliability and availability of research infrastructures?
- Is this model able to produce useful insights for the FCC-hh?

These questions lead to the main original contributions of this thesis that are:

- A model to estimate hadron collider's availability for physics production;
- model validation against LHC operations;
- first results and thoughts about Future Circular hadron collider availability.

This thesis also contains contributions that were not originally planned to be a part of the study. They are a response to the first findings of the study that the chosen model platform did not fully support the required model features and the lack of reliability data hindered the

model applicability [1]. The desire to have software with native support to special features in my collider model lead to a development of the new OpenMARS system reliability and availability approach [2, 5]. Thus, this thesis also presents a contribution that describes how the key requirements for OpenMARS were derived from the collider model. The lack of reliability data from research infrastructures led to a study on reliability data sharing practices in the industry and assessment if this is applicable in the accelerator community. This thesis introduces this study.

CERN has had a good experience on model-based approach for assessing collider availability in the LHC high luminosity upgrade project (HL-LHC) [18, 19, 20]. In that project, the model was done using Matlab software. The FCC study wanted to test if commercial reliability modeling tools were better suited for this purpose and chose Ramentor oy's ELMAS software [21] as the platform for the model. The plan was also to substantially increase the model details to improve modeling collider operations and failures.

The primary goal of this thesis is to describe the resulting model to estimate hadron collider's availability for physics production. The focus is on how the model was developed to resemble the collider operations and validating its appropriateness. Parts of my model were published earlier in [1], where it was validated against the operational data from the LHC. This thesis provides more details on the model implementation and how it was extended to model operations at collider complex level. The model is further used for deriving availability goals for FCC-hh systems.

The key argument of this thesis is that collider's availability for physics can only be estimated if the whole collider complex and the collider operations are taken into account. Regarding practical methods, the model uses fault trees [22] for system availability modeling and semi-Markov chains [23] for operations modeling. The fault tree and Markov models are interconnected. Failure events can affect the state transitions in the Markov chains, and the active Markov states affect the failure rates in the fault trees. Special attention is given to model operational schedules, physics production and the process required to inject the particle beam into the collider. The ELMAS software uses discrete event simulation to obtain results [24].

This thesis provides additional considerations on the model applicability. The section 9.2.3 shows that there is clear potential in applying similar techniques in modeling industrial process or manufacturing applications. However, the approach has also limitations. The lack of good sources of reliability data proved to be a challenge in model utilization. This lack of data is now addressed by a new study that looks into the possibility to develop an accelerator reliability database [4]. Modelling in this thesis is only applied to manage risks related to loss of production. Additional considerations are needed when dealing with risks involving catastrophic consequences or personnel safety.

The work for FCC availability study was performed in collaboration between CERN, Ramentor Oy and Tampere University of Technology. The author's contribution for this project was the design and the implementation of the collider operations model in the ELMAS tool. The design of the model involves the decisions on what aspects of collider operations should be modeled and what methods should be used to model specific aspects. The model implementation in the ELMAS tool, available during the project, was not straightforward. ELMAS was designed mainly for fault trees, but it allowed adding user code to a model for implementing special features. The collider model required extensive amounts of additional code to make model functional.

Having a high amount of user-defined code in a model is problematic. Programming is a particular skill, and expertise is required to understand the code and to write understandable code. The use of custom code creates a high skill floor for using and developing a model. This can lead to a dependency where practically only the model creator is able to develop the model further. Due to this reason, CERN launched a collaboration with Ramentor Oy to develop a new version of

the modeling software that would provide native support for these special features. This lead to the development of an Open Modeling approach for Availability and Reliability of Systems (OpenMARS) [2, 5]. The collider model provided the key requirements for OpenMARS approach and the author was heavily involved in developing and documenting it. The OpenMARS approach is potentially the most significant outcome of my research. Details of how my work lead to requirements and features in OpenMARS are provided in the section 8.1.

This thesis is organized as follows. Chapter 2 offers a brief overview of the FCC study. Chapter 3 summarizes state-of-the-art in complex system availability modeling and applications in accelerator facilities. Chapter 4 describes the availability and operational aspects of hadron colliders and the chapter 5, techniques that the model uses to depicts them. Chapter 6 shows the model validation against LHC operations and the first results for the FCC-hh availability. Chapter 8 provides details on the OpenMARS approach and its application to the collider complex model as well as introduces the concept of the accelerator reliability information system in the ARIES study. Chapter 9 offers thoughts on availability of circular electron-positron colliders and applicability of the developed model before conclusions in the chapter 10.

2 Future circular collider study and motivation on availability study

The 2013 update of the European strategy of particle physics [25] recommended that: *"CERN should undertake design studies for accelerator projects in a global context, with emphasis on proton-proton and electron-positron high-energy frontier machines"*. This lead to the launch of the FCC study [16]. The scope of the FCC study is multi-pronged. The main emphasis is the FCC-hh, but the study also analyses: (i) corresponding injector chain for the collider taking into account the existing CERN accelerator infrastructure; (ii) options for electron-proton collider (FCC-he); (iii) a lepton collider with a centre-of-mass energy in the order of 90 to 350 GeV (FCC-ee), as a potential intermediate step before hadron collider; and (iv) potential to build a new hadron collider in the LHC with more powerful magnets to reach higher energies (HE-LHC). FCC-hh and FCC-ee would be located in a new collider tunnel in the Lake Geneva basin as shown by the figure 2.1.

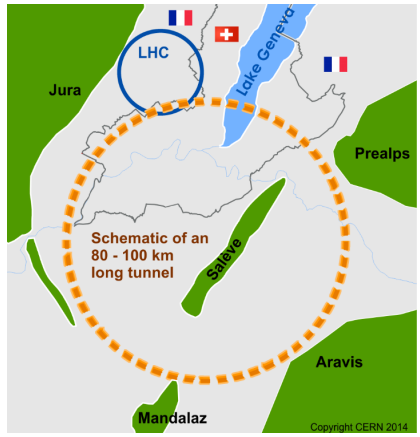

Figure 2.1: Schematic of a 100 km collider tunnel in the Lake Geneva basin [26]

In circular colliders [27], the energy of circulating protons E depends on bending radius r and the field of the bending magnets B:

$$E[GeV] = 0.3 * B[T]r[m]. \tag{2.1}$$

The centre-of-mass energy in a head-on collision is:

$$E_{cm} = 2E. \tag{2.2}$$

The design centre-of-mass energy for the LHC is 14 TeV. It achieves this energy with 27 km ring and magnets that can reach up 9 T. FCC-hh aims to achieve 100 TeV centre-of-mass energy by increasing the ring circumference to 100 km and the bending magnet fields to 16 T. The LHC magnets are made from NbTi. Increasing the field strength will require changing the magnet material to Nb_3Sn. Developing these new magnets is one of the key technological challenges of the FCC-hh [28].

As with the LHC, the purpose of the FCC is to study high energy physics. The particle physics strategy of the United States describes that within 10-20 year time scale a very high-energy proton-proton collider is the most powerful future tool for direct discovery of new particles and interactions [29]. Experimental measurements at a 100 TeV collider would cover previously unexplored territory at energies never reached before in a laboratory environment [30]. The high energies specifically create an opportunity to study multi-Higgs physics [31]. At the same time, the high rate of Standard Model (SM) events allows to pushing to new limits in the exploration of rare SM phenomena (e.g., rare decays of top quarks or Higgs bosons) [32]. In these studies the new physics can be found indirectly as the structure of the standard model could be tested at unprecedented energies and with unparalleled precision [30]. If the observations do not agree with theoretical predictions, this would mark a breakdown of the standard model of particle physics and the rise of new physical processes.

Gaining a credible statistical significance to confirm a new physics discovery relies on the production of a sufficient number of particle collisions over an extended period of time. Over the years scientists have become more and more careful before calling the evidence of high energy physics phenomenon as a discovery or observation of this phenomenon [33]. A practical reason for this is to avoid canceling the physics discoveries. A recent example exists where a potential hint of a new particle was detected at LHC in 2015 [34], but it was not confirmed by data collected in 2016. The event gained much attention and demonstrated the need for a conservative approach towards the evidence of new physics.

Physics production in colliders is measured using luminosity

$$L(t)\sigma_p = \frac{dR}{dt}, \tag{2.3}$$

which gives the rate of events $\frac{dR}{dt}$ for a given time t for interactions with production cross-section σ_p [35]. Rare events have a small production cross section, but the size of it increases with the collision energy. Thus, the higher collision energy improves the chances of discovering new events and observing rare events. For example for Higgs boson: $\sigma_p \approx 50$ pb at 13 - 14 TeV and $\sigma_p \approx 800$ pb at 100 TeV [36] leading to a 16 fold increase in production of these particles in the FCC-hh.

The luminosity produces over time is called integrated luminosity

$$L_{int} = \int_0^t L(t)dt. \tag{2.4}$$

Integrated luminosity is a measure for the number of events potentially observed by the particle detectors. It is one of the Key Performance Indicators (KPI) for particle colliders designed for fundamental physics research. The FCC-hh plans to produce 17.5 ab^{-1} integrated luminosity: during 25 years of operations [17]. To put this into perspective, the LHC aims to deliver close to 3-4 ab^{-1} of integrated luminosity by 2039 [37].

Achieving the production goal will require stricter availability goals for accelerator systems and the injector chain. Understanding these goals early in a project is highly beneficial as the ability

to affect the outcome of the project decreases as the project matures [38]. The reason for this is that the cost of change increases as the project matures. Making a change in a design phase is relatively inexpensive compared to construction and operations costs and decisions made in this phase can affect the whole project life cycle. To give an example of a project cost structure figure 2.2 shows the annual cost of the LHC project. The cost peak is associated with building the LHC.

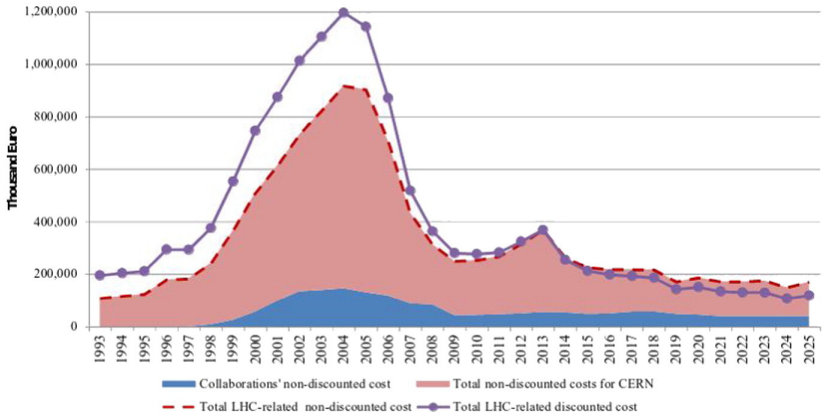

Figure 2.2: The annual cost of the LHC project [39]. The cost prediction assumes steady operations and does not include future upgrades like HL-LHC or the decommissioning costs.

3 Complex system availability modeling and applications in particle accelerators

3.1 Availability as a key performance indicator

3.1.1 Terminology and definitions

The vocabulary of International Electrotechnical Commission (IEC) [40] defines availability as an item's ability to be in a state to perform as required. The instantaneous availability $A(t)$ is the probability that an item is in a state to perform as required at a given instant. The mean availability can be calculated over a period of time (t_1, t_2) by integrating over the studied period to sum the times when the system is available:

$$A = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} A(t) dt. \tag{3.1}$$

Achieved availability depends on the system's combined characteristics of reliability, recoverability, maintainability and the maintenance support performance.

The reliability is the probability that a system performs as required for the time interval (t_1, t_2), under given conditions. The notation $R(T)$ is used for reliability function that gives the likelihood that an item is intact at time t. The likelihood that an item is failed is given by failure distribution function:

$$F(t) = 1 - R(t). \tag{3.2}$$

The probability that a failure occurs at a certain instant of time is:

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \frac{F(t + \Delta t) - F(t)}{R(t)} = \frac{f(t)}{R(t)}. \tag{3.3}$$

If the failure rate is assumed to be constant and then the mean operating time to failure (MTTF) can be calculated as:

$$\text{MTTF} = 1/\lambda, \tag{3.4}$$

and the exponential distribution is used to depict failure distribution:

$$F(t) = 1 - e^{-\lambda t}. \tag{3.5}$$

If a failure occurs, corrective maintenance is required to regain the upstate where the system is available. The IEC defines the term maintainability as system's ability to be retained in or restored to a state to perform as required, under given conditions of use and maintenance. The IEC also defines the term recoverability as system's ability to recover from a failure, without corrective maintenance. The actual failure related downtime depends on maintenance support's performance.

The downtime associated with a failure can be summarized as the time to restoration. With the mean time to restoration (MTTR) and MTTF, the mean availability can be calculated as:

$$A = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}. \tag{3.6}$$

The exponential distribution is easy to use as the only parameter is the occurrence rate. For example, if a system consists of 100 individual units and if during a one year 15 units fail:

$$\text{MTTF} = 100 \text{ units} * 1\ a/15 \text{ failures} \approx 6.67\ a, \tag{3.7}$$

when instantaneous repairs are assumed. A major setback in the exponential distribution is that it is memoryless, it assumes that the mean time to an event is constant over the time. This can lead to unrealistic assumptions. For example, let's assume that MTTF of an item is ten years. If this item were intact after ten years the MTTF at that point would be still ten years. Assuming memoryless distribution leads to an assumption that the age of an asset does not affect its failure behavior. For the MTTR, this assumptions can lead to a situation where if the MTTR is one week and if the repair were still ongoing at this point the MTTR would be still one week.

There are many alternative options for the exponential distribution, as the failure and restoration behavior can be modeled with any distribution. There also exist non-parametric methods to express reliability function such as Kaplan-Meier estimator [41]. However, the Weibull-distribution is commonly used to model age-dependent failure behavior. It can model a failure rate that increases, decreases or stays constant over time. The distribution is given by the equation:

$$F(t) = 1 - e^{-\lambda t^{\beta}}. \tag{3.8}$$

The shape factor β defines if the failure rate is decreasing ($\beta < 1$), increasing ($\beta > 1$) or constant ($\beta = 1$).

Figure 3.1 shows a practical example on how a distribution models failure data. The failure event sample is given on the top of the figure. The lower part of the figure shows a failure distribution based on a cumulative histogram, an exponential distribution, a normal distribution, and a Weibull distribution. The exponential distribution is the farthest away from the histogram as it fit was based only on the mean value. The normal distribution is fitted based on mean value and sample variance, and the Weibull is fitted with so-called Weibull-paper [42]. In this case, the Weibull distribution provides the best fitting. However, using it will require information on the age of the unit at the time of failure which is not always available.

The availability and reliability are not the only performance indicators of interest. Industry often uses the indicator Overall Equipment Efficiency (OEE) [43] to measure production performance

$$\text{OEE} = \text{Availability} * \text{Performance} * \text{Quality}. \tag{3.9}$$

In the OEE, the availability depends on time reserved for planned maintenance and the amount of corrective maintenance. The performance depends on the production rate and the quality can be understood as the number of rejections or the number of produced goods that do not meet the quality requirements. Alternatively, OEE can be calculated simply as:

$$\text{OEE} = \frac{\text{Achieved production during scheduled hours}}{\text{Theoretical maximum production within the time frame}}. \tag{3.10}$$

The OEE does not consider costs. For this purpose, EN 15341 [44] defines metrics such as:

$$\frac{\text{Total maintenance cost}}{\text{Quantity of output}}, \tag{3.11}$$

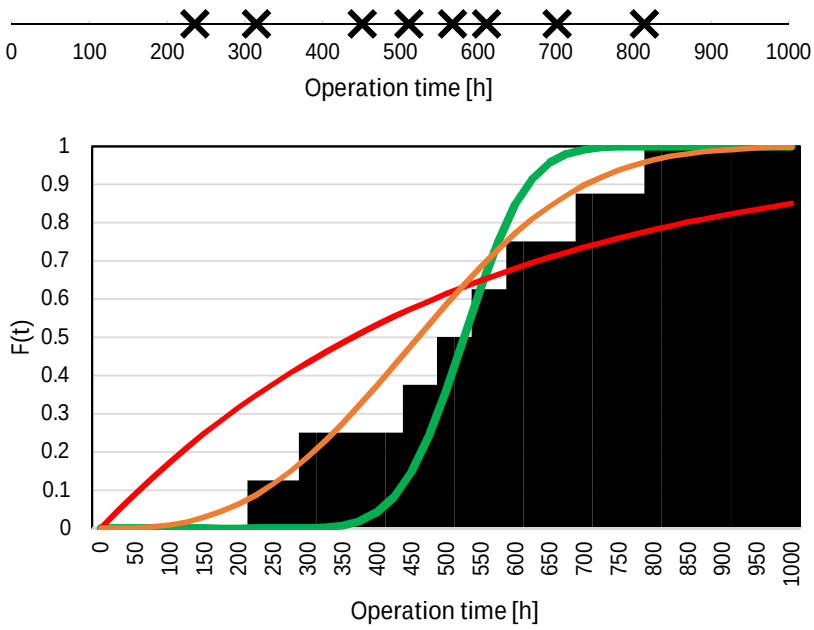

Figure 3.1: The top part shows a failure event sampleto which a cumulative histogram, an exponential distribution (red curve), a normal distribution (green curve) and a Weibull distribution (orange curve) are fitted.

$$\frac{\text{Total maintenance cost+ Unavailability costs related to maintenance}}{\text{Quantity of output}}, \tag{3.12}$$

and

$$\frac{\text{Availability related to maintenance}}{\text{Total maintenance cost}}. \tag{3.13}$$

Although modern standards define coherent availability terminology, this is not always reflected in practice [45, 46]. Source [47] says that one problem is that a failure simply is an unpleasant word. Another problem is that different definitions exist. For example, source [48] defines failure as: *"Failure denotes an element's inability to perform its designed function because of errors in the element or its environment, which in turn are caused by various faults"*. Source [48] further defines a fault as: A fault is an anomalous physical condition, which can be caused by a design error, manufacturing problem, damage or fatigue.

These definitions are different than the ones in the IEC vocabulary [40]. There the failure is: *"A loss of ability to perform as required, a failure of an item is an event that results in a fault of that item."* and a fault is: *"An inability to perform as required, due to an internal state. A fault of an item results from a failure, either of the item itself, or from a deficiency in an earlier stage of the life cycle, such as specification, design, manufacture or maintenance."*. So, in the IEC a fault is a state of an item and a failure an event that starts the fault state.

Alternatively, the definition in [48] says that failures are caused by faults. Another definition can be found in [22] that states that: *"All failures are faults but not all faults are failures. Failures are basic abnormal occurrences, whereas faults are "higher order" events."* To explain this, the reference [22] uses an example where a fault is caused by an event outside the system boundary.

For such an event IEC vocabulary [40] would use a term secondary failure as a failure caused by a failure or a fault of another item.

Different and sometimes conflicting definitions can cause problems. To avoid this issue, CERN started a training course to ensure that everyone uses a common terminology [49].

3.1.2 Applications in accelerators

In the accelerator field, the availability is nowadays a standard requirement for machine design. It is clear that machine availability and production goals are key ingredients to achieve sustainable operations. This kind of target can be found in most of the on-going accelerator projects this list includes: linear proton accelerators: SNS [50], Myrrha [51], Linac4 [52], ESS [53] and IFMIF that is fusion material test facility consisting of two accelerators [54]; linear electron-positron colliders: ILC [55] and CLIC [56]; and circular colliders: SSC [57], HL-LHC [58] and CepC [59]. Following text presents some of these examples in more detail.

Superconducting Super Collider (SSC) project was a hadron collider project in the early 90's. In this project, a goal was set that the SSC should achieve 80% availability during the scheduled operations time [57]. It was argued that if the availability falls below this level efficiency of operations falls rapidly. However, the LHC achieved operations with 80% availability first time in 2017 [60] and the reference [57] even mentions that the availability of Tevatron was at that time below 80%. This example shows that giving an availability target only as a percentage of the time machine needs to be operational can be somewhat vague. Due to this, in the on-going HL-LHC project the target is given as an integrated luminosity production goal, and the availability targets are derived from this goal. This makes the target more meaningful as it is directly linked to the physics goals. Thus, this same approach is also the chosen approach in the FCC project.

Availability targets are a rather new addition to large accelerator projects. To the best of the author's knowledge, only one published case exist where the requirement and the achieved availability observations are available. The Spallation Neutron Source (SNS) is a high power linear proton accelerator used for neutron production. During a design phase, a goal of 95% availability was set for the accelerator [50]. This goal can be compared to achieved availability. In an article from 2013 [61], SNS is reported to achieve between 70 to 90% availability. However, these values are not comparable, as they are counted differently. The reported values include failures of the beam target that were not taken into account in the requirement. Also, the requirement assumes 2 MW beam energy, but only 1.2 MW was reached at the end of the observed period.

Improving the reliability of high power linear accelerators opens a new accelerator application. Availability of an accelerator driven sub-critical nuclear reactor depends heavily on the accelerator reliability. The accelerator is used for providing neutrons to the reactor to sustain the chain reaction. If the accelerator fails, the nuclear reactor starts to rapidly cool-down. Myrrha project aims to build the world's first prototype of a sub-critical lead-bismuth cooled reactor [51]. In this project, the current assumption is that if a beam interruption lasts more than 3 seconds, the reactor needs to be shut down. As the recovery will take a significant amount of time, the current reliability target is that: the accelerator can have only ten failures lasting more than 3 seconds per 3 months. This requirement has lead to a design with redundant particle sources and accelerator cavities.

Comparing requirements and achieved availability between laboratories is difficult as different laboratories use different metrics. This challenge is present even if the machines were built for similar purpose and were operated similarly. This was proven in a survey studying operation statistics of storage ring light sources [62]. Based on this finding, common operations metrics have been proposed for these facilities [63]. One of the considered factors is the beam current in

the light source because it has a direct effect on the intensity of the provided synchrotron light. Lüdeke et al. propose limits I_{tol} and I_{min} that would mark low beam current and no beam failure modes.

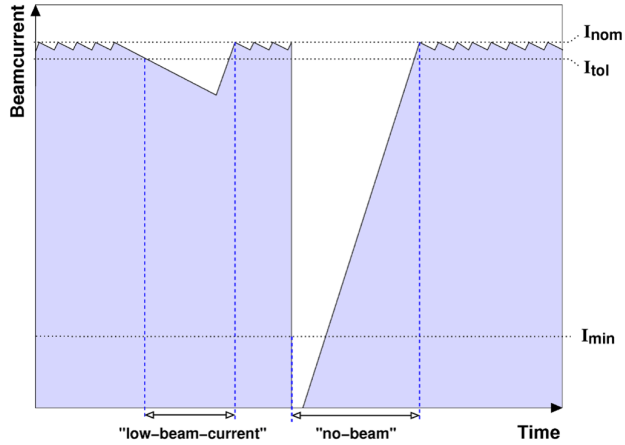

Figure 3.2: Beam current in a synchrotron light source and limits I_{tol} and I_{min} that are related to ow beam current and no beam failure modes [63].

For a collider, as a part of his thesis [18] A. Apollonio introduced definitions for availability:

Run time (RT): scheduled operational time of the accelerator.

Beam delivery time (DT): effective time for beam delivery to experiments or other accelerators.

Availability for physics (PA): the probability that an accelerator provides beam to the experiments or to the next accelerator in the accelerator chain. (PA = DT/RT)

However, measuring particle accelerator availability for physics only as a percentage of time is problematic as the physics production rate is not constant. So-called Hübner factor H has been invented to measure production efficiency [64]. The idea is quite similar to the OEE. The integrated luminosity can be roughly estimated as a product of peak luminosity L, run time RT, and H:

$$L_{int} \approx H * RT * L. \tag{3.14}$$

The Hübner factor contains every element that restricts the collider from operating continuously with the design peak luminosity. These element range from system availability to operational considerations. However, this approach does not consider the fact that most colliders cannot physically operate continuously with maximum peak luminosity. For example, in the FCC and the LHC, the reasons for this are the luminosity lifetime and the operational cycle that are presented in the chapter 4. So, the Hübner factor cannot be directly compared to the OEE [43] that takes this limitation into account.

3.2 Complex system availability modeling

3.2.1 Methods for availability modeling

Unavailability of an asset can be seen as a risk as it has a certain likelihood and usually monetary consequence. Several methods have been developed for risk assessment. For example, the standard IEC 31010 [65] lists 31 different methods for this purpose. This thesis will not list all of them but gives examples of the most important ones.

Fault tree handbook NUREG-0492 [22] divides reliability modeling methods to inductive and deductive methods. Here an inductive approach stands for a bottom-up technique that tries to analyze on what components a system consists of and what consequences a failure of a component has on the system level. A deductive approach, on the other hand, is a top-down technique. The first step is to determine what are the top level failures of the system and the analyses determine possible causes for these top events. Another important distinction is between qualitative and quantitative methods. Qualitative methods aim to identify risks and quantitative methods aim to predict risk probability or failure frequency.

Failure Mode and Effect Analysis (FMEA) is a classical method in reliability engineering has been presented in many different ways. NUREG-0492 [22] presents the method as being deductive and qualitative. Table 3.1 shows the example presented in the book. With this approach, an analyst would list every component in the system, determine their failure modes, assign probabilities and determine the criticality of the failure. Such an approach is today supported by electronics reliability libraries that can be used by commercial software such as Isograph [66].

Table 3.1: FMEA example showing the list of components within a system with failure probabilities and determination if a failure is critical for the system.

Component	Failure probability	Failure Mode	% Failures by mode	Critical	Non-Critical
A	$1*10^{-3}$	Open	90		x
		Short	5	x	
		Other	5	x	
B	$1*10^{-3}$	Open	90		x
		Short	5	x	
		Other	5	x	

Alternatively, FMEA has also been presented as an inductive and qualitative method. Table 3.2 shows first columns in VDA 86 FMEA form and an example of risk priority number [42]. Risk Priority Number (RPN) is an estimate for ranking risks. In VDA 86, it is based on three factors: Occurrence, Severity and Detection which are ranked on a scale from 1 to 10 where 10 is the most adverse value. The RPN is calculated with the equation:

$$\text{RPN} = \text{Occurrence} * \text{Severity} * \text{Detection}. \tag{3.15}$$

Occurrence depicts the likelihood of occurrence and Severity the severity of the consequences. VDA 86 is designed for production quality assurance. In reference [42], the detection determines the likelihood that the quality control fails to detect a faulty product. The consequence of this would be that a faulty product is delivered to a customer. However, alternative standard IEC 60812 [67] gives a more general definition for the detection as the chance to eliminate the failure before system or customer is affected.

Table 3.2: Columns related to risk assessment in VDA 86 FMEA sheet the full sheet contains columns for information on how the situation could be improved

Column	Description
Function	List of system elements or process steps and their functions
Potential failure modes	List of possible failure modes
Potential effects	Potential effect of failure modes
Potential causes	Potential causes of failures
Current controls	Current measures used for failure detection
Occurrence	1-10
Severity	1-10
Detection	1-10
Risk priority number	1-1000

A shortcoming of the FMEA is that it cannot depict failure combinations such as redundancy. These combinations can be modeled in methods like the fault tree or the Reliability Block Diagram (RBD). The fault tree is a deductive method [22]. An analysis starts by determining top level faults. For each top event, a fault tree is constructed to find all credible ways in which the top event can occur. The tree forms a graphical representation of the faults that can lead to the top event occurring. Fault tree can be used both qualitatively and quantitatively. The fault tree model is always qualitative, but it can become also quantitative if likelihoods are attached to events.

In contrast to a fault tree, Reliability Block Diagram studies what system functions need to be present in order for a system to work properly. RBD model forms a success path that is required for the preferred event or state to exist. Fault tree can always be converted to RBD and vice versa if the model contains only AND and OR gates. Figures 3.3 and 3.4 show how similar logic can be presented in both modeling methods.

Chapter 5 in book by M. Rausand [68] compares fault tree and RBD approaches. As the methods are largely comparable, the interesting question is which one is better? The book claims that using fault tree tend to lead more complete models and thus it always recommends to form a fault tree for a system model. The given reason for this finding is that the analysts think differently when they use a fault tree compared to when they use a RBD. While using a fault tree, an analyst thinks: How this system can fail? When using a RBD, an analyst thinks: How does this system function? The book argues that this focus on functions can lead to forgetting safety functions out of the analysis if they are not essential for a system to work.

Modeling availability requires a concept of states to show if an asset is available or not. Figure 3.5 shows a simple Markov model with two states available and unavailable and transitions between the states. A failure logic can also be presented with a Markov model. Figure 3.6 shows the logic present in figures 3.3 and 3.4 in a Markov format. This figure also shows why this is not usually done as the complexity of the model starts to rapidly increase with additional states.

Markov model can be solved analytically if all transitions are exponentially distributed. For example, the availability of the model in the figure 3.5 can be calculated with an equation 3.6. This assumption is problematic if the failure is age dependent. For example, in the figure 3.6 if transition probabilities to "A fault" states depends on components A age, the model needs to carry this information in it. Achieving an analytical solution becomes difficult as the age dependency affects the transfer rates to all states where component A is failed. A Markov model that is

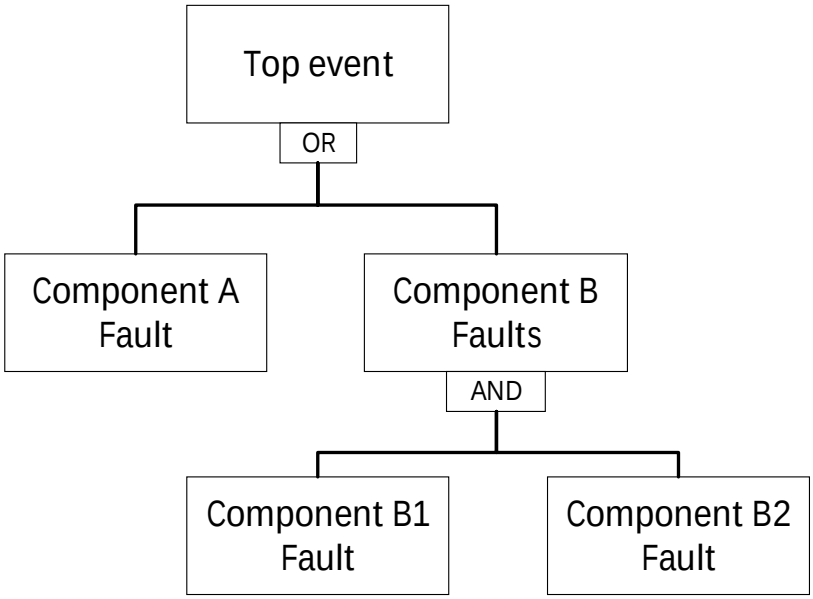

Figure 3.3: A fault tree for a system consisting of two main components A and B where component B has internal redundancy. Failure of component A or B causes a top event that results in system failure. Component B failure can only occur if the internal redundancy is lost and requires the failure of both component B1 and B2.

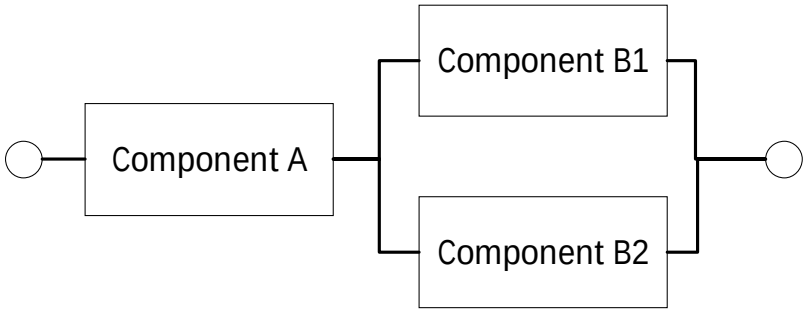

Figure 3.4: RBD of the failure logic presented in the figure 3.3. RBD forms a success path. If this path is free, the system works. In this example, the path can be blocked if component A fails or both components B1 and B2 fail.

generalized to contain any transition distributions is called a semi-Markov model [42]. However, with this generalization, the results cannot be obtained analytically anymore. Another variation of Markov model is so-called Hidden Markov Model (HMM) where certain states are not directly observable. In reliability engineering HMM is used for example for modeling condition based maintenance where certain observations are used to assess condition of a system that is not directly observable. For instance, such model has been used for predicting failures in telecommunication system based on error sequences [69]. In accelerator field it has been suggested that HMM could be used for predicting beam trips [70].

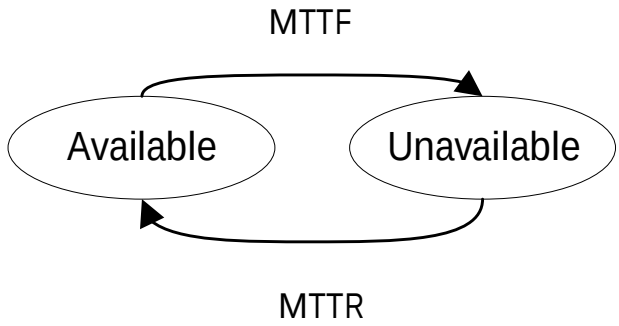

Figure 3.5: Simple Markov model with two states and transitions to model component availability.

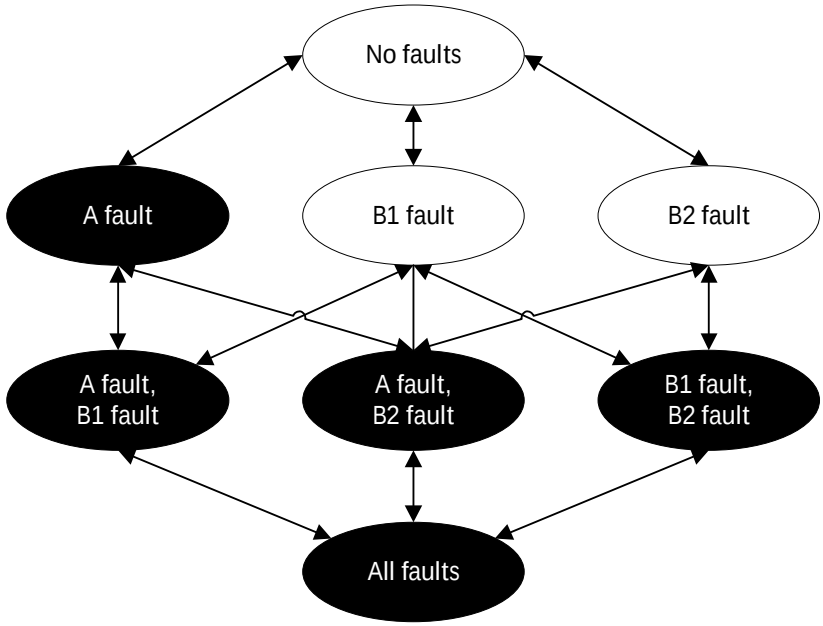

Figure 3.6: Markov model of the failure logic presented in the figures 3.3 and 3.4. The black color distinguishes the states, where the system is unavailable. Transition rates between states are MTTF and MTTR distributions. For example, transitions between "No faults" and "A fault" are the MTTF and MTTR times for failure A. Transition from "A fault state" to a state with component B1 or B2 failure is the failure distribution of B component.

In a Markov model, only one state is active at a certain time. This is a limitation if one wants to model the availability of manufacturing process that depends on the state of several processes. Examples of such models can be found in the manufacturing industry [71]. Petri net solves this problem with tokens that can move between places in the net where possible connections are defined with transitions and arcs. In practice, a transition can present a process. Figure 3.7 shows

an example of an assembly process where two parts are connected to a base assembly to create a final product.

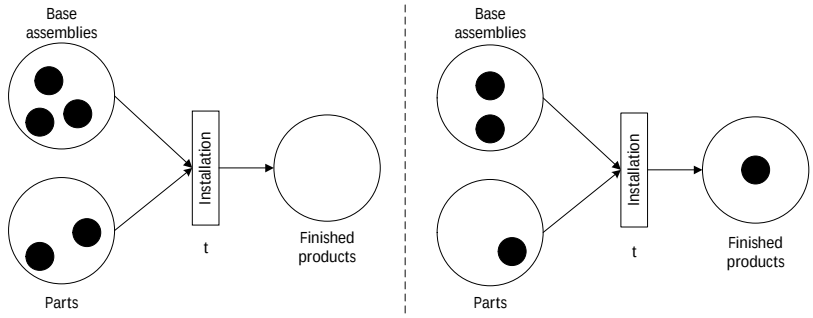

Figure 3.7: Petri net of an assembly process from reference [71]. Part is installed to a base assembly to finish a product. This figure shows how, in the model, a part and a base assembly tokens are combined by an installation transition to create a product token. In this example, if new parts do not arrive the last assembly token cannot be finalized as a product.

Colored Petri net is a further generalization of the Petri net where tokens can have different colors which enables distinction between tokens. Example cases can be found in modeling reliability of hybrid car drive system [72] and system with multiple state and units [73, 74]. In these specific cases colored Petri net was needed to model: in [72], development of the component's age as it affects the transition probabilities and in [73, 74], distinguish between different unit types that had different transition rates.

Alternatively, manufacturing operations can also be modeled with queuing theory models. For example in a Jackson network production facility is modeled as a network of interconnected queues where the service time in each queue is exponentially distributed [75]. However, compared to Petri net, queuing theory are more limited in terms of state-phase definitions and process time distributions [71]. Other modeling methods used in reliability engineering include Graph-based Functional-Failure Identification and Propagation framework (FFIP) that is based on functional failure logic (FFL) [76]. In FFIP, functional relations are mapped to see how failure in a function propagates within a system. Also, Bayesian networks are used in reliability engineering.

Article [77] says that the advantages of Bayesian networks over fault trees are treating the uncertainty and "multi-state" variables. On the treatment of uncertainty, the article says that in Bayes a variable can be a probability distribution whereas in fault tree these type of uncertainties are handled by sensitivity analyses where parameter value is changed within a certain range. The article describes multi-state variables with an example where a power supply may induce a CPU failure when failing, while other failure modes do not cause a CPU failure. Article [77] claims that it is not possible to model this in a fault tree, while in a Bayes network this can be modeled naturally by connecting each power supply to the CPU. However, modern fault tree software can model common cause failures where a single failure can cause multiple consequences, which seemingly solves the issue presented by the article [77].

The present paradigm in availability and operations modeling, especially for complex systems, prefers method free discrete event simulations for this purpose. The state-of-art in complex system reliability modeling is discussed in sources [78, 79]. Source [78] states that modern engineering systems with complex maintenance policies, with aging or otherwise non-exponential failure distributions, and standby redundancies are too complex to model with analytical approaches.

These approaches might require unrealistic assumptions when the case situation is presented with analytical modeling technique. Due to this, simulation analysis has become a necessity to model these features. Source [79] agrees by the first account by stating that modeling modern systems goes beyond the classical approaches such as fault trees and they fail to capture the distinct features of system behavior like production, maintenance, and emergency management.

An introduction on how Discrete Event Simulation (DES) works is given in source [80]. In practice, a simulation consists of event timeline, simulation time and states defined by a model. In reliability engineering, interesting events are generally failures and repairs, and the interesting states determine whether a system is in fault or not. So, generally at the start of a reliability simulation each element, that can fail, creates a failure event to the event timeline. Figure 3.8 shows an example of a timeline with events.

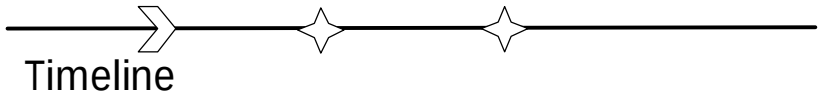

Figure 3.8: A simulation timeline with two events.

Usually, the event time is a random value based on associated distribution. Figure 3.9 show how a random value in the range between 0 to 1 can be turned into random value from a distribution by using the cumulative distribution function. In DES the simulation states do not change between the events. So, practically the simulation jumps from event to event, where these events update the simulation time and can change the modeled states. For example, a failure event can be implemented as an event that changes the state of the failed item to the fault and creates a new repair event to the event timeline. Figure 3.10 shows an approach that implements a DES calculator for OpenMARS.

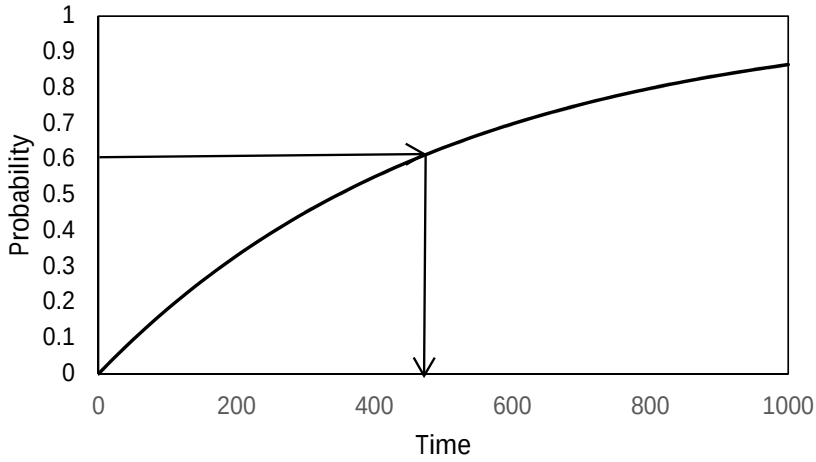

Figure 3.9: Cumulative distribution can be used to generate random values from the distribution. Here a value from random number generator (0.6) is transformed to random value from the distribution (450). Depending on the studied phenomenon the timescale can be from seconds to thousands of years.

Simulations can be classified to be deterministic or stochastic. In a deterministic simulation, events are generated with predetermined times and the result does not vary between simulations. However,

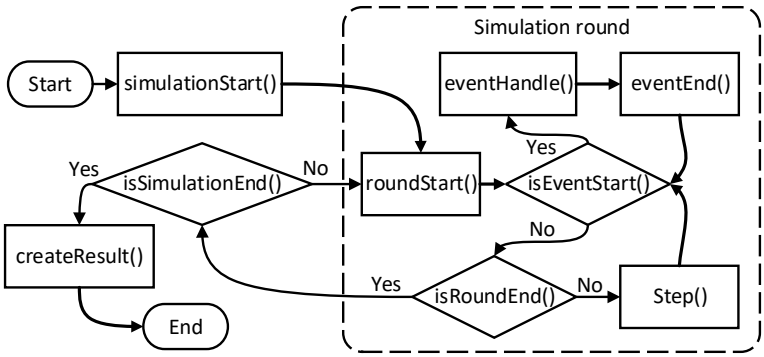

Figure 3.10: Concept representation of a DES calculator [5]. Simulation job is defined by the number of rounds, i.e., how many times simulation job is run and the length of a round. This figure shows how after each round simulator checks if the completed round was the final one. During a round simulation proceeds from an event to another. For each event, the simulator runs the eventHandle method that can change the simulation parameters and create or alter the events in the timeline.

as the failures are probabilistic events, they are generated with a random number generator and the results vary between simulations. These kinds of simulations are called stochastic. Due to the randomness, a result of a single simulation round is rather meaningless. In literature, this problem is called aleatoric uncertainty which arises from the intrinsic randomness of a phenomenon [81]. The word *alea* means rolling the dice in Latin. This is different from the epistemic uncertainty that is the result of a lack of knowledge of exact input parameters. Again, in Latin *episte* means knowledge. Controlling aleatoric uncertainty requires running the simulation multiple times to create a result sample that is not affected by statistic randomness.

There are several formalisms for model-based safety assessment [82] and commercial products for complex system modeling [83]. A detailed comparison of these approaches requires a deep understanding of the individual formalisms (e.g. [84, 82]) and is out of the scope of this study. This study uses Ramentor Oy's ELMAS software [21] as a platform for the model where the calculations are based on discrete event simulation. This thesis does not, however, claim that alternative approaches would not have worked.

3.2.2 Survey of accelerator availability studies

Our hypothesis before the study was that availability modeling is not common in accelerator facilities. However, availability has become consistently more important with tightening user requirements and increasing complexity of infrastructure [85]. This section reviews the work done in different studies. Earlier this kind of review has been done in a paper by O. Rey Orozco et al. [86] and in my publication [1]. This thesis extends the review in [1] by updating it and providing more details in modeling techniques. However, A. Apollonio's modeling work for HL-LHC study [18, 19] is covered in chapter 5 as it was an essential ingredient in developing my model.

At Fermi National Accelerator Laboratory, a C++ based Monte Carlo or discrete event simulation model was developed for Tevatron hadron collider [87, 88]. This tool allowed modeling the luminosity production based on varying operational parameters. In particular, downtime for

the various parts of the complex was included as well as beam-related parameters like transfer efficiencies, emittance growths and changes in the luminosity lifetime. The tool helped to determine and optimize operation strategies and to forecast how improving reliability could improve the luminosity production.

For the International Fusion Materials Irradiation Facility (IFMIF) study, several approaches have been tried for availability modeling. In 2005, I. Balan published his Ph.D. thesis on availability calculation technique that was based on Markov chains and implemented with Fortran 77 programming language [89]. The model was inputted as a fault tree that the code translated to Markov chains for obtaining the results. In 2011, C. Tapia et al. [90] published a similar type of analysis as [89]. However, they were using RiskSpectrum PSA Professional® as a platform for their fault tree. Later in 2014, E. Bargalló et al. published a paper on in-house availability model for IFMIF, arguing that commercial tools are insufficient to accurately model the complexity of the accelerator facility [91].

In their work, E. Bargalló et al. adapted the Availsim software for IFMIF. The AvailSim was originally developed for the International Linear Collider study (ILC) [92, 55]. The first version of the AvailSim was a Monte Carlo simulation platform based on Matlab software. The motivation for the AvailSim was to accurately model unique features in accelerator operations such as degraded operation modes and operational procedures. These type of arguments motivated the adaption also for IFMIF. The latest version of the Availsim was developed in European Spallation Source [93]. The new version is based on the Python programming language. The reason for the change was to avoid requiring a Matlab license to use the software.

One of the most reliability critical applications of accelerators is accelerator-driven sub-critical reactors or accelerator-driven systems (ADS). In this application, a three second long beam trip can force a shutdown of the nuclear reactor [51]. Two papers focus on modeling accelerator in ADS use. L. Burgazzi and P. Pierini used reliability block diagrams in Relex software[1] [94]. Later, A. Pitigoi and P. Fernández Ramos used Risk Spectrum software to model the reliability of Spallation Neutron Source linear accelerator [95]. They plan to use this model to develop an ADS model in the future.

Several studies at CERN have also used Isograph software for reliability and availability modeling. In accelerator facility modeling this software has been used for reliability analysis of Linac4 [52], radio-frequency system upgrade for Proton Synchrotron Booster [96], and in Compact Linear Collider study [97].

P. Gupta published a study assessing annual maintenance budget of Pelletron accelerator in Inter-University Accelerator Centre at New Delhi [98]. Assessment is based on the graph-theoretical approach that takes into account parameters such as availability, spare-parts, personnel aspects, and work planning.

The survey shows that both commercial and in-house tools are used availability studies. The use of the in-house software is commonly motivated by a need to model case specific features that are not supported by commercial software. In-house software is almost always based on Monte Carlo or discrete event simulation techniques. This review did not take into account reliability studies aimed to prevent severe failure. These studies are a well-established activity in the accelerator community. For example, the LHC has a highly reliable machine protection system [99] and studies are also performed to prevent failures that can affect personnel safety [100].

[1] The software is currently known as PTC Windchill Quality Solutions.

3.3 Data for availability models

3.3.1 Methods to obtain data

Quantitative models depend on data or estimates on failure rates and repair times to calculate availability. Here three potential sources are described: operations data; handbooks and expert estimates. Also, concept of accelerated life testing is introduced.

Calculating reliability metrics from operations data requires collecting information on failures during operation. Here a source of this information can be a maintenance database or data from warranty claims. One of the greatest successes of the data-driven approach in reliability engineering comes from the aircraft industry where Weibull-distribution was used to prove that maintenance does not necessarily improve the reliability [101]. This lead to the reliability centred maintenance approach where every maintenance task must be justified by its effect on improving the reliability or reducing the operational risks. The approach helps to find the most cost-efficient maintenance policies that maximize the system reliability [102]. It is nowadays successfully applied in aviation, process and nuclear industries.

The author has experience in performing Weibull-analyses in two projects that tried to use operational data to assess system or component reliability. Author's master thesis [8] describes a project for Finnish Air Force. In this project, the author developed a software to calculate a cumulative failure distribution using Weibull-distribution and item's age dependent failure tendency using so-called Power-law function [103]. Thanks to the Air Force's vigorous data collection practices, the project was highly successful. An analysis that was conducted as a part of the thesis revealed aging in one component of the aircraft hydraulic system. This result justified an overhaul task for these components to resist the aging. Another study, that the author took part in, aimed to repeat the success in the Air Force study by studying warranty data and using the results to predict warranty costs. The publication [9] lists the difficulties in using operations data for reliability prediction that were encountered in this study. The maintenance data was impartial and thus the results calculated with this data were potentially not credible. The publication [9] describes that one of the key reasons for impartial operations data is that the return of investment in collecting this data is not immediate and this return is not obvious to the part of the operations that is tasked to collect the information.

If operation data do not exist or if data are unreliable, there are few possibilities to estimate the reliability. Accelerated life testing tries to give information on failure rates and aging in a time-limited test [104]. The hypothesis is that one can create a test scenario that represents loads that system would encounter during a life cycle of a unit. The key questions in this area are, can the test scenario represent realistically the life cycle and what is the minimum sample that is required to estimate the reliability of a product. In other words, what is the most economical way to produce credible results.

Another option is to compare the product to similar products and use the data collected from them as a basis for a reliability prediction. Such data can be found in handbooks. One famous example is the OREDA project that has collected and published reliability data from oil industry equipment since 1984 [105]. The OREDA approach has been described in the standard ISO 14224 [106]. For electronic components, several standards exist that use mathematical estimates to predict component reliability in different operating conditions. The most well-known standard is the military handbook MIL-HDBK-217 series that started in 1965 and ended in 1995. The publication of these handbooks was ended due to criticism that electronics advance in such a rapid pace that these type of handbooks cannot keep up and predictions based on them are incorrect because they rely on obsolete data and concept of constant failure rate in electronics is flawed [107].

However, recently this type of handbooks have resurfaced: FIDES guide was published 2004 and RIAC-Handbook-217Plus was published 2006 [108]. The main argument that supports this activity is that estimates given by these standards are useful in the early design stage even if the predictions are not fully accurate.

Sometimes no published estimate or data exist on system reliability. Several methods have been created to generate estimates based on expert judgement. The methods have different criteria on how to judge individual systems. AGREE allocation method was introduced in 1957 and it is based on system complexity and criticality [109]. In 1965 the Karmiol method included the factors: complexity, state-of-the-art, operational profile, and criticality [110]. In 1980, T. Saaty described an Analytical Hierarchy Process (AHP) in which deciding factors are compared pairwise against each other [111]. In practice, if systems are judged with absolute scales, the question is: What are system A's and B's complexities in scale 1-10. However, in pairwise comparison question is: Is system A more complex than system B. The latter seems more intuitive and allows allocating reliability by performing a series of pairwise comparisons between systems. S. Virtanen presented an example of this type of analysis in [112]. Studies on reliability allocation methods are on-going. For example, an article by C.-S. Liaw et al. [113] suggested using Decision-making trial and evaluation laboratory (DEMATEL) method for reliability allocation. Earlier this method has been used for evaluating marketing strategies, R&D projects, e-learning evaluations, managers' competencies, control systems, and airline safety problems.

3.3.2 Examples of data collections activities in the accelerator world

Accelerator availability has always been a keen interest of operators. In the first particle accelerator conference (PAC) 1965 several papers presented results on operational availability of accelerators [114]. One of the most the best examples of maintenance data collection in accelerator domain is the CATER-database in Stanford Linear Accelerator Center that has been used since 1988 [115]. At CERN, the availability of the LHC drew interest after the first operational years that established safe LHC operations [116]. This interest in availability led to launching the Accelerator Fault Tracker (AFT) project that has provided tools and workflows that allow capturing the failure statistic of CERN beam operations [117, 118]. Before AFT, the CERN fault tracking depended on numerous, diverse and unrelated systems used by different organizational units within CERN. As a result, achieving a clear and consistent overview of the availability was difficult and burdensome. Figure 3.11 shows the data collection process upgrade between 2012 and 2015. In 2012 statistics were combined from various sources at the end of the year, but since 2015 data collection and combination is done as a part of the operations.

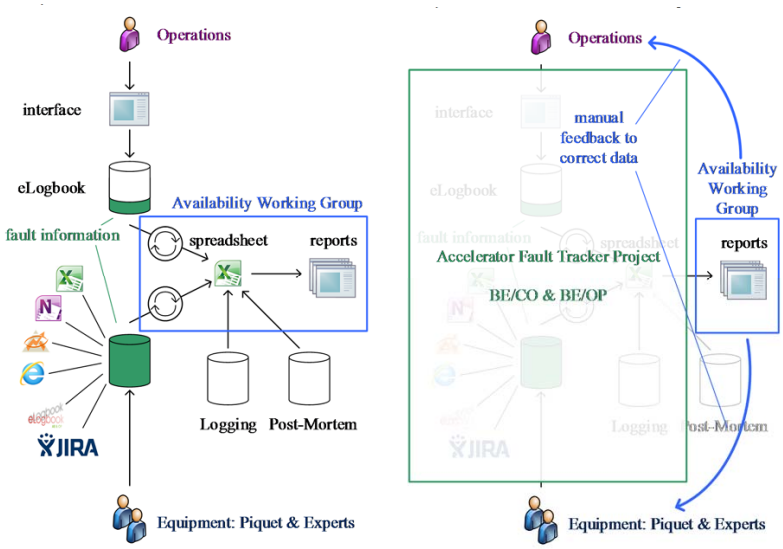

Figure 3.11: Evolution of failure data collection at CERN [117]. On the left old process where individual failure data repositories where studied to form the overall picture. On the right, the new system where failure data is collected continuously to a common platform.

When reliability data have not been available a few published cases exist where expert judgment methods have been used to generate availability estimates. Specific cases can be found on Superconducting Super Collider (SSC) project [119], Spallation Neutron Source (SNS) project [120] and CLIC project [121]. SSC project used, so-called, feasibility of objective technique to allocate reliability requirements for Low Energy Booster for SCC. SNS project used Analytical Hierarchy Process to allocate availability goal among facility systems and subsystems and CLIC project tested DEMATEL procedure for availability allocation.

4 Hadron collider operations and luminosity production

4.1 Basic operational concepts of synchrotrons

This section shows that in a synchrotron [122], such as LHC, a particle beam is accelerated with radio frequency system that is synchronized with the circulating frequency. The beam travels in a circular track with a constant radius. Magnets provide the force to bend the beam to the track and the constant circulating radius is maintained by synchronizing the strength of the magnetic field to the beam momentum.

A stream of accelerated particles is called a particle beam. Usually, a beam is not contineous, and it consists of bunches of particles separated with a specified interval. For example, in the LHC a nominal bunch contains about $1.15 * 10^{11}$ protons and nominal bunch spacing is 25 ns [10]. Figure 4.1 shows trains of bunches populate one LHC ring.

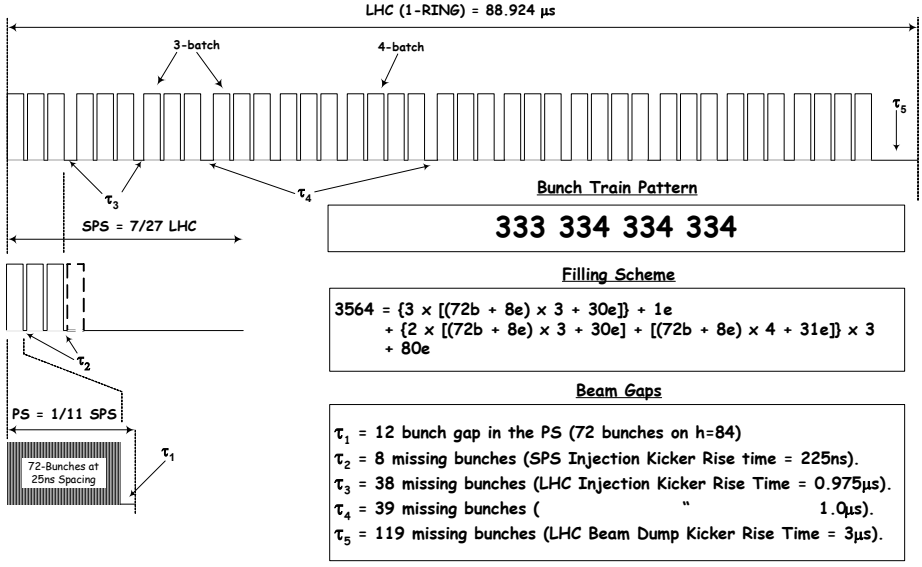

Figure 4.1: Structure of the LHC beam with 25 ns bunch spacing [123]. The length of PS accelerator determines the bunch train lengths, and the gaps between trains are caused by rise times of SPS and LHC kickers. Without these gaps, the beam would see an insufficient kicker magnet field that could lead to a part beam colliding with vacuum pipe or high beam losses. This situation is however taken into account in the design [10].

A common way to accelerate a beam is to use Radio Frequency (RF) cavities. The figure 4.2 shows the operating principle of a RF-cavity: as a bunch travels through the cavity, a resonating electromagnetic field inside the cavity accelerates the bunch. The resonating field is necessary. As if the field in a cavity was static, the cavity would de-accelerate the bunch before it enters and after it exits the cavity. The left side of the figure 4.2 shows how a combination of two electromagnetic waves creates RF buckets that accelerate bunches. The acceleration is caused by an electrical field that forms within a bucket. The field vectors are shown on the right side of the figure 4.2.

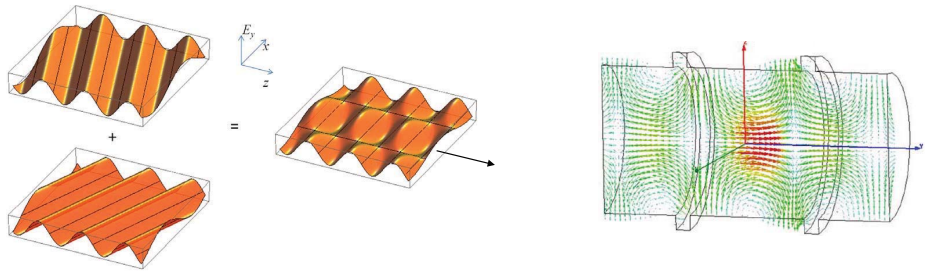

Figure 4.2: Left side of the figure shows how a combination of two waves creates a wave pattern traveling in the z-direction. The right side shows the electrical field vectors of a notched circular waveguide with propagating oscillation mode. [124]

This introduction on beam optics mainly is based on [125, 126]. The beam is steered to the circular orbit with a magnetic field. The basis for this is the Lorenz force:

$$\bar{F} = e(\bar{E} + \bar{v} \times \bar{B}), \tag{4.1}$$

where $\bar{E}$ is the electrical field, $\bar{B}$ is the magnetic field, $\bar{v}$ the velocity of the particle and e is the charge of the particle. At speeds near the speed of light, the effect of 1 T magnetic field is similar to a significantly high electric field of $3 * 10^8$ V/m. Thus, the electrical field can be neglected, and the magnetic field required to bend the beam can be calculated from the equation

$$\bar{B}\rho = \frac{m\bar{v}}{e}, \tag{4.2}$$

where m is the relativistic mass and ρ is the bending radius. Figure 4.3 shows the effect of the magnetic field.

In a synchrotron, dipole magnets are utilized to generate the fields to bend the beam. The strength of the magnetic fields is synchronized to the momentum of the beam

$$\bar{p} = m\bar{v}. \tag{4.3}$$

This keeps the accelerating beam in a stationary orbit. Also, the resonance frequency of the RF-system is synchronized to the circulating frequency of the beam. When the speed of the particle nears the speed of light, the velocity hardly increases and the circulating frequency is practically constant at that point. In these conditions, the acceleration increases particle's momentum mainly by increasing the particle's mass.

The dipole magnets create a design orbit that ideally every particle should travel. However, orbit deviations, momentum deviations and magnet imperfections exist and require additional magnets to focus the beam to the design orbit. As seen in the figure 4.4, the orbit deviation results from particles oscillating around the design orbit. This oscillation is called the betatron tune and it is

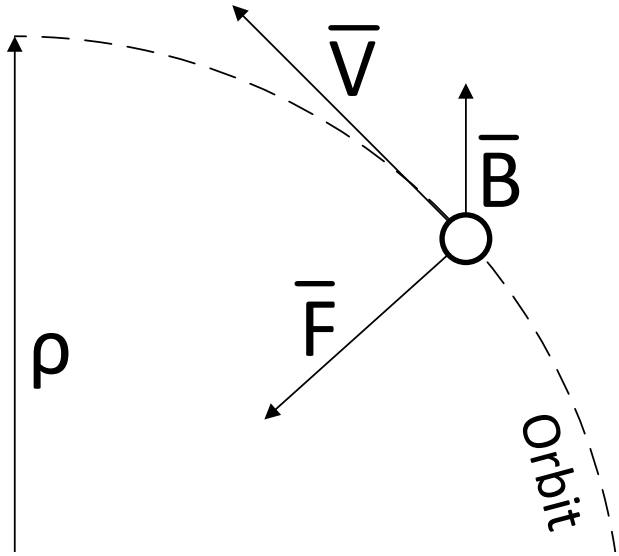

Figure 4.3: Magnetic field $\bar{B}$ steers the charged particle to the orbit

stabilized with quadrupoles. The magnetic field of a quadrupole: (i) is zero in the centre as there no correction is needed; (ii) the field changes linearly along the axis providing focusing force that is strongest at the edges of the magnet; and (iii) a quadrupole focuses beam in one direction and de-focuses it in another.

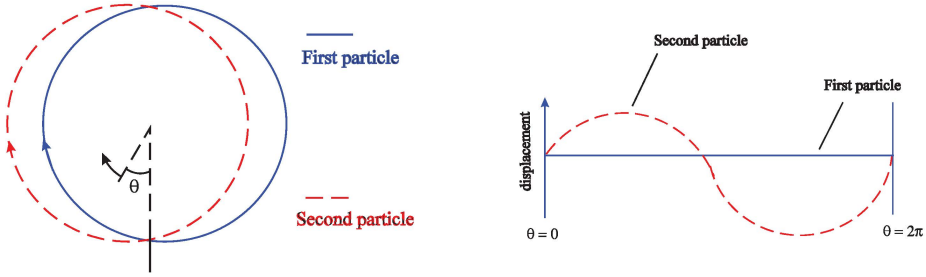

Figure 4.4: Particle orbit deviations result in particles oscillating around the design orbit forming the betatron tunes [126].

An important parameter concerning the individual bunches is so-called emittance. The emittance is the area of which contains a certain percentage of particles. This area forms an ellipse in x, x' coordinates where x is the displacement of a particle and x' an angle with respect to design orbit [126]. This idea can be understood with the figure 4.4, where the second particle has certain displacement and an angle of motion compared to the design orbit. A low emittance means that a large percentage of particles is close to the design orbit. This leads to high instantaneous luminosity in a collider.

Due to this focusing / de-focusing effect, an accelerator has pairs of vertically and horizontally focusing quadrupoles throughout the accelerator ring. The figure 4.5 shows these focusing and defocussing quadrupoles that are commonly called as FODO cells.

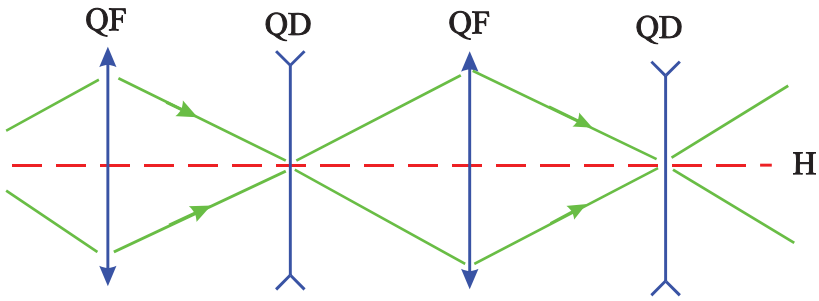

Figure 4.5: Series of focusing and defocusing magnets from [126]. Focusing the beam with quadrupoles show clear similarities to optical lenses.

The momentum deviations affect the betatron tune based on the following function:

$$\frac{\Delta Q}{Q} = \xi \frac{\Delta p}{p}, \tag{4.4}$$

where ΔQ is the change of betatron tune, Δp the change of momentum and the quantity ξ is called as chromaticity. It tells how betatron tune changes with particle momentum. The momentum spread is corrected by increasing the betatron correction for high energy particles and decreasing it for low energy particles. This correction can be achieved with sextupole magnet. The figure 4.6 shows field of the sextupole magnet and quadrupole magnet and their combination.

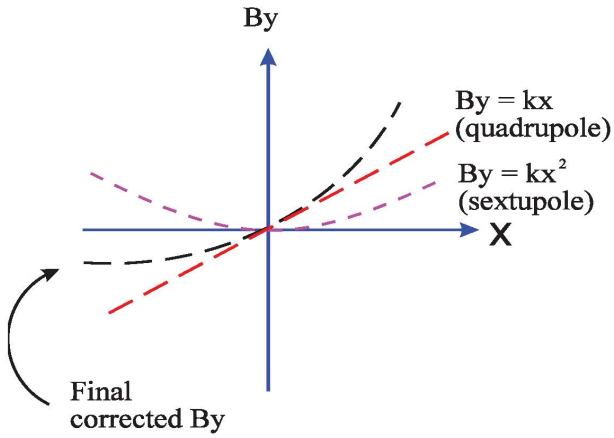

Figure 4.6: Magnetic field of a quadrupole and sextupole magnets and their combination. Sextupole field strengthens the one edge the quadrupole field and weakens the other. [126]

4.2 Operational schedules

Over the life cycle of a particle accelerator complex, periods of luminosity production runs and maintenance shutdowns alternate. Figure 4.7 shows long-term schedule for LHC and HL-LHC [37]. The long shutdowns are used for maintenance, consolidations and to upgrade the machine.

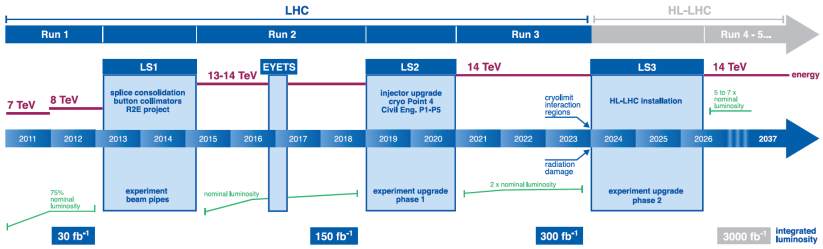

Figure 4.7: Long-term schedule for the LHC operations and HL-LHC upgrade

Within production runs, the performance increases stepwise. To reproduce the operations, this has to be modeled. Figure 4.8 shows the luminosity production evolution in the LHC operations [127]. The first year of LHC operation (2010) focused on machine commissioning and 45 pb^{-1} of integrated luminosity were produced. In 2011 the commissioning was still ongoing and the production was 5.5 fb^{-1}, and in 2012, the production was 23.27 fb^{-1}. For LHC run 2, the bunch spacing was decreased from 50 ns to 25 ns and the collision energy was increased from 4 TeV to 6.5 TeV. These changes, together with improved availability, allowed increasing the luminosity production. For example, In 2016, the LHC reached 40 fb^{-1}, surpassing the production estimates. However, in 2015 the first year of run 2, the production was relatively low (about 4 fb^{-1}), due to machine recommissioning.

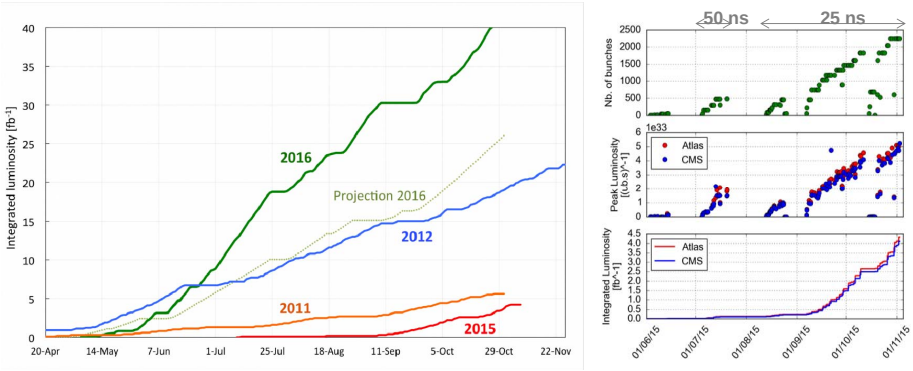

Figure 4.8: Luminosity production in the LHC for different years [127] and intensity ramp-up for 2015 from [128]. 2011 & 2012 were during run 1 and 2015 & 2016 during run 2. The production in 2010 was so limited that it would be barely visible in this scale.

Increased production rate within operational years can be explained by beam intensity ramp-up. In the ramp-up, the beam intensity is increased stepwise to commission the machine and to avoid the prompt changes to the operation conditions to reduce risks. Luminosity production evolution is small in the LHC if it is compared to Tevatron. A paper by V. Shiltsev [129] shows that many accelerator facilities have achieved significant increases in their production performance over their life-cycle. For example, in Tevatron about a hundredfold increase in peak luminosity was achieved gradually over ten years of operations. The increase was due to numerous improvements, both during operations and annual machine shutdowns.

An idealized plan for annual LHC operations is shown in the figure 4.9. During a run, each winter has a technical shutdown. In the LHC, the cryogenics system is warmed to a standby mode

during the stop. The stop is followed by machine commissioning periods that aim to verify the readiness for operations. The commissioning is separated into hardware commissioning and beam commissioning. The beam commissioning is carried out with low-intensity beams to limit damage potential. Beam conditioning also improves vacuum conditions. For example, at the start of the second LHC run problems with macroparticles [130] and so-called electron cloud [131] limited the beam intensity. Then, the proton physics period starts and the beam intensity is progressively increased to reach the nominal parameters as shown in the figure 4.8.

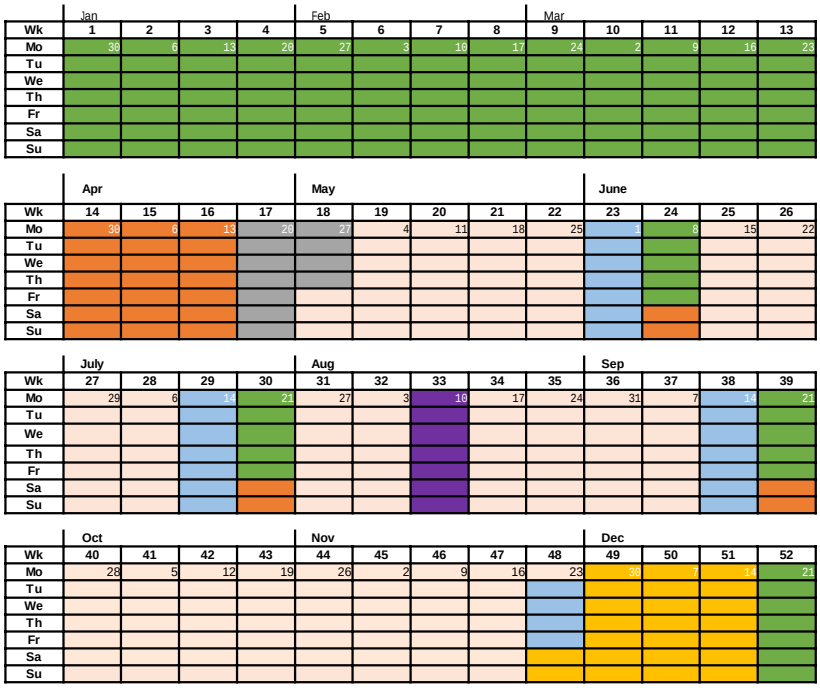

Figure 4.9: An idealized LHC schedule based on [37]. Green color signifies technical stops, orange commissioning, gray scrubbing, beige proton physics, blue machine development, violet special physics, and yellow ion physics.

The proton physics are stopped several times by short technical stops. They are often accompanied with time for machine development that aims to optimize and further study the performance of the machine. Technical stops are planned according to operational necessities. For example, deferred corrective maintenance can be carried out during these. Operations time is also reserved for special physics and ion physics.

4.3 Operational cycle

4.3.1 Overview of the cycle

Particle colliders have distinct operational cycles. Published literature on this topic can be found for example for LHC and FCC-hh [7], and the reference [132] lists turnaround times for HERA, RHIC and, Tevatron. The reason for the operations cycle is that before collisions are produced, the collider needs to be filled with beams that are accelerated to collision energy. During the acceleration, the magnetic fields are ramped up to keep the beam in the orbit. After the collision phase is over the energy in the magnets needs to be ramped down before a new cycle.

A report by R. Alemany Fernandez et al. [7] divides hadron collider operational cycles into nine phases. Phases are shown in the figure 4.10 and they consist of:

Setup: machine systems are prepared for particle beam injection.

Injection probe: A low-intensity particle bunch is injected for measurement and correction purposes to prepare the machine for nominal beam injections.

Injection physics: the machine is filled with nominal intensity bunches injected in a bunch trains.

Prepare ramp: systems are prepared to accelerate beams.

Ramp-Squeeze: beams are accelerated and squeezed to make a beam physically smaller in interaction points. A synchrotron, like the LHC, maintains the stable beam orbit by ramping up the magnetic field strength while the beams are accelerated.

Adjust beams are adjusted to collide in interaction points.

Stable beams once adjustments are finished beams are declared stable and experiments start to collect data from interaction points.

Beam dump beams are dumped out from the machine due to an operational decision or an interlock activation in the machine protection system.

Ramp down energy in the magnetic fields is ramped down to a pre-injection level.

Table 4.1 shows the minimum and average times for turnaround phases in the LHC during 2017.

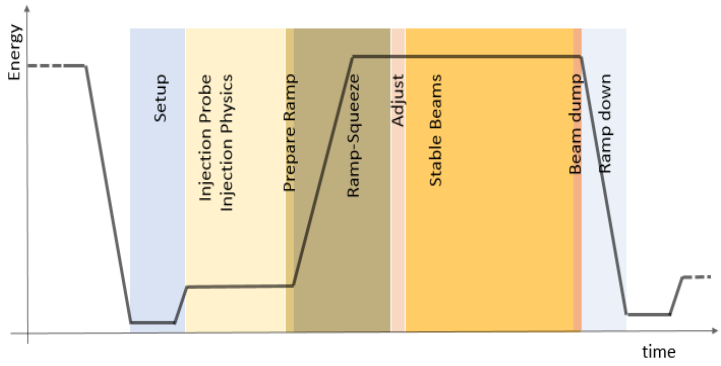

Figure 4.10: Collider cycle phases [7].

Table 4.1: Observed minimum and mean turnaround times in 2017 [133] in minutes.

Phase	LHC min 2017	LHC mean 2017
Injection	28.0	77.1
Prepare ramp	2.3	5.0
Ramp-Squeeze-Flat top	20.2+13.4+2.8	20.5+18.1+4.5
Adjust	3.3	7.9
Ramp down	36	153.2[a]
Total	106.0 (1.8 h)	286.3 (4.8 h)

[a]The ramp down phase includes the recovery time from failures.

4.3.2 Injection phase

There are several factors that limit the minimum injection energy for an accelerator. This is discussed in detail in High Energy LHC design report where different injector options have been analysed [134]. A higher injection energy is favourable in terms of impedance[1], beam stability, aperture and energy (field) swing. For practical example, errors caused by persistent currents in super conducting magnets become more significant in low energies. This motivates the research to decrease the magnet filament size to limit the amount of persistent currents. A beam has to be accelerated to 450 GeV before it can be injected into the LHC. For this purpose, the LHC has an injector chain that is shown in the figure 4.11. The whole chain from Linac2 to SPS must be operational during the injection.

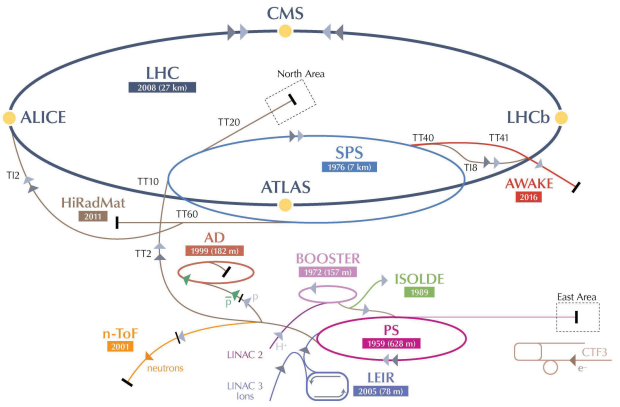

Figure 4.11: Linac2, PSB, PS, SPS, and corresponding transfer lines are required to produce and inject the beam to the LHC [135].

The preparation for the injection phase starts before the LHC requires the beam. Optimally, the LHC operators inform the injector operators well in advance. This gives time to prepare the beam for the LHC and to verify the beam quality. As when LHC does not require beam, the SPS produces beam only for the experiments in the so-called CERN's north area that is shown in the

[1]Beam communicates with its surrounding vacuum chamber through electromagnetic fields generated by the beam itself. This self-interaction mediated by the beam environment is called impedance in accelerator physics.

figure 4.11. On average, in 2017, the LHC had to wait the beam for 10 minutes before injection could start [136].

The injection phase starts with the injection of so-called pilot beams or probes. In the LHC, a pilot beam is a low intensity bunch that cannot damage the machine. It is used for verification of safe operations before physics beams are injected. Certain measurements are also performed using the pilot beams such as orbit and tune corrections. On average, in 2017, the measurements with pilot beams took 15 minutes [136].

Multiple injections from SPS are needed to fill both LHC rings. Also, the SPS needs four injections from the PS to fill it. This is shown in the figure 4.12 that shows the process to prepare the SPS to inject the beam to the LHC. Limiting factors are the length of the SPS and the beam energy. The SPS is 7 km long and the LHC has two 27 km long rings. Amount of stored energy in the beam delivered by the SPS has a potential to damage components. Due to this the LHC has systems to protect the machine against this damage [10]. In an injection to the LHC, injection kicker magnet system (MKI) provides the final steering. An error in the MKI functioning can lead to a serious damage. The primary protection against this is the by injection beam stopper (TDI). TDI and additional shielding elements absorbs sufficient amount of energy to prevent the damage if the MKI fails or behaves erroneously during an injection [10]. TDI's ability to safely absorb the energy was found to be lower than the design performance [137]. This lead to a limit in the injection energy and subsequently increased the amount of injections needed to fill the LHC.

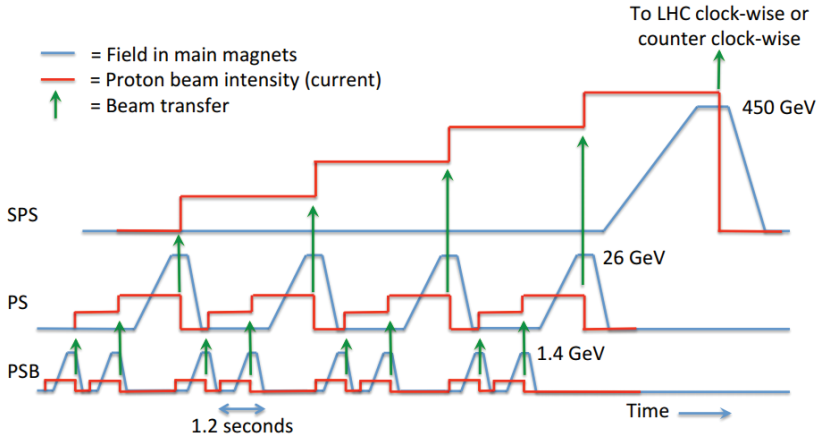

Figure 4.12: Preparing the beam in the SPS to be injected into the LHC [138].

With nominal parameters filling of the LHC requires 24 injections and theoretically takes only 8-16 minutes [10]. This time does not include pilot beam phase. Theoretical turnaround also assumes SPS injection interval of 21.6 sec. Achieving this interval would require SPS to be devoted to LHC during the injection phase. Currently, the SPS provides beam also to fixed target experiment during the LHC injection. Every other beam goes to the North area and the injection interval to the LHC is 59 sec.

As seen from the table 4.1 the average length of the injection phase is furthest away from minimum operational time. There are several possibilities why an injection phase can last longer than expected. The most common issues are that the beam is not ready or the quality of the beam is not within the set parameters.

The quality of injected beams is controlled to protect the machines [139, 138]. A SPS injection can be rejected by multiple systems including SPS Beam Quality Monitor BQM, interlocks in the LHC, and in the beam transfer lines. In this context, An interlock monitors signal produced by a system and stops an injection if the monitored signal is outside the tolerable limits. The SPS Beam Quality Monitor is specified to: i) verify the synchronization between the SPS and LHC frequencies to inject the beam in the correct position; ii) measure individual bunch qualities; iii) beam losses; and vi) check for the presence of satellite bunches in the LHC and in the SPS [139]. Satellite bunches are bunches that are located between regular bunches (e.g., in the LHC within the 25 ns bunch spacing).

Additionally, the Injection Quality Control (IQC) system is designed to stop the injection process if an error occurred during an injection [138]. Through several measurements, the system decides if an injection was successful. In addition to BQM, the IQC analyses data from beam transfer lines, LHC injection kickers and beam instrumentation. Figure 4.13 shows possible IQC analyses results and operational consequences and the figure 4.14 shows IQC result percentages during 2011 and 2012. Also during 2015, about 20% of injections had to be repeated and the IQC warnings caused operational delays [140]. Few analyses have been conducted on this topic [138, 140, 136], however, further studies should be conducted to fully apprehend the causes for missed injections.

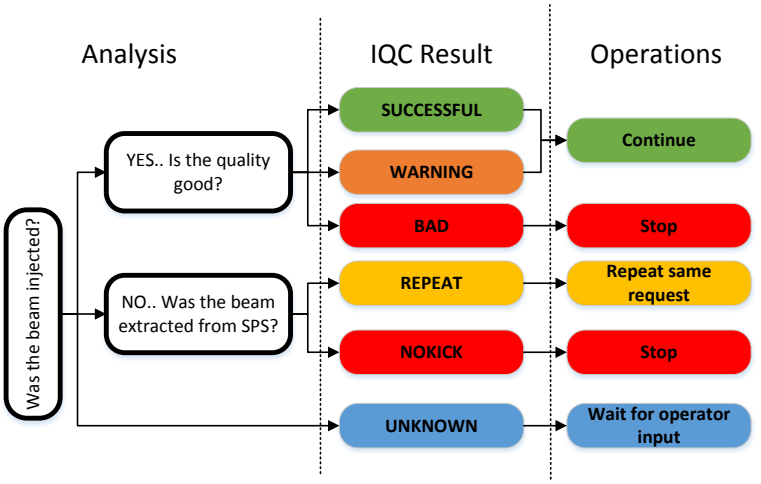

Figure 4.13: IQC analysis logic [138]

Although repeated or bad injections slow down the operations, one of the most critical failures is a complete or partial failure of MKI. A direct impact of the beam produces a highly localised temperature rise of around 600 ºC [141]. This halts the operations as the TDI requires time to cool-down.

4.3.3 Luminosity production

The luminosity is produced when beams are set to collide. The instantaneous production rate can be calculated with an equation [35]:

$$L = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y}, \tag{4.5}$$

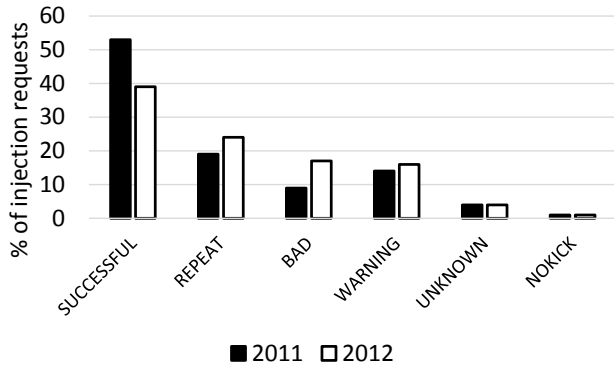

Figure 4.14: IQC analysis results for LHC run 1 [138]

where N_1 and N_2 are the numbers of particles in a bunch in a specified beam, f the revolution frequency, N_b is the number of bunches and σ_x & σ_y define the bunch size in horizontal and vertical directions.

Particles are destroyed in collisions. This combined with the beam losses leads to a decay of luminosity production over time. From the modeling point of view, both in the LHC and in the FCC the luminosity production is time-dependent. So, the time spent in stable beams cannot alone measure the achieved amount of integrated luminosity.

For the LHC and HL-LHC, the integrated luminosity can be modeled with an equation [142]:

$$L_{int}(t) = \begin{cases} L_p t, & : t < t_{lv} \\ L_p t_l v + L_p \tau_L (\exp(-t_{lv}/\tau_L) - \exp(-t/\tau_L)), & : t \geq t_{lv}, \end{cases} \tag{4.6}$$

where L_p is the peak luminosity achieved during a fill, t_{lv} is the luminosity leveling time when the luminosity production rate is kept constant by operational means, and τ_L is luminosity lifetime. Equation 4.6 assumes exponential decay of luminosity. Figure 4.15 shows the evolution in luminosity production rate (instantaneous luminosity) during a LHC fill and with planned parameters for HL-LHC fill. Several factors contribute to the increased production rate: i) number of particles within bunches is increased; ii) magnets squeezing the beams before collisions in interaction point are replaced with stronger ones; and iii) new so-called crab-cavities rotate the bunches to maximize bunch overlap at the interaction point [13].

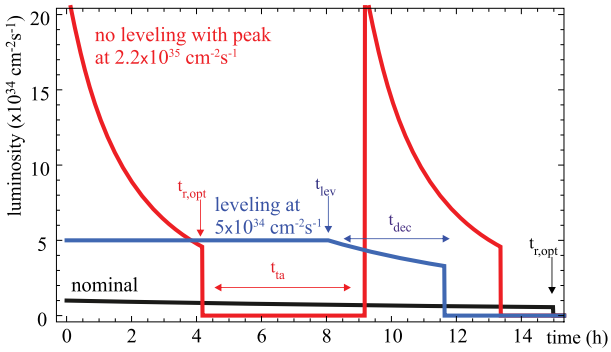

Figure 4.15: Black curve shows the instantaneous luminosity evolution with the LHC design parameters, blue with the HL-LHC design parameters and red with the HL-LHC design parameters, but without luminosity leveling [143].

For FCC-hh two phases of operations are planned [143]. Figure 4.16 shows the luminosity production rates in the two phases. The limiting factor for the FCC luminosity production is assumed to be forces caused by circulating beams interacting on each other. This effect is measured by the so-called beam beam tune shift parameter. The difference in the production rates in the phases results on the fact that phase 2 assumes less strict limitation on the tune shift. The main difference to the LHC beams is that with the higher energy the synchrotron radiation emitted by protons dampens the beam and causes bunches to shrink. This leads to beams be consumed quickly by the collisions and short 3-4 hours optimal fill time.

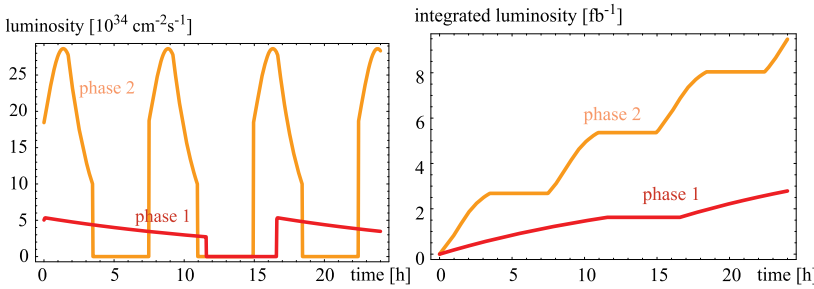

Figure 4.16: Luminosity production in the FCC in phase 1 and 2 [143].

4.4 LHC availability and key systems

Today's particle accelerators are among the most complicated and expensive scientific instruments ever built, and they exploit almost every aspect of cutting-edge engineering technologies [144]. This means that accelerator operations depend on the availability of a large number of systems and the purpose of this section is to give a brief introduction to them. Table 4.2 shows system

Table 4.2: Number of failures and unavailability times in the LHC by individual systems during 2016 and 2017. Fault duration takes into accounts the full time a system was unavailable for the LHC. Root cause time only takes into account the time the root cause of the failure was within the system.

Root cause class	System	Faults		Fault duration		Root cause duration	
		2016	2017	2016	2017	2016	2017
Equipment	Injector complex	138	96	360.4	145.2	313.2	140.2
	Cryogenics	42	31	361.1	207.0	90.3	107.7
	Power converters	66	84	106.6	113.7	75.1	98.9
	Quench protection	45	55	37.0	61.8	25.9	63.8
	Beam dumping system	30	22	70.2	63.7	64.2	60.6
	Cooling & ventilation	(67)	7	(221.7)	14.9	(278.4)	53.4
	Radio frequency	40	32	40.2	47.3	33.4	45.5
	Electrical network	(67)	15	(221.7)	36.3	(278.4)	37.8
	Beam instrumentation	40	22	47.1	32.6	37.6	32.0
	Other	2	22	0.1	20.7	9.2	29.2
	Injection systems	30	18	46.3	20.6	40.7	19.4
	Access system	12	10	24.5	16.5	13.9	18.5
	Experiments	52	28	59.9	16.1	50.3	15.4
	Accelerator controls	19	24	9.6	15.4	12.8	15.0
	Machine interlocks	8	5	6.0	4.3	4.4	5.7
	Vacuum	2	2	1.4	3.4	1.2	3.0
	Transverse damper	10	4	13.8	3.8	10.6	2.1
	Ventilation door	10	2	21.1	0.7	9.2	2.0
	Collimation	23	10	73.4	2.5	10.7	1.7
	Magnet circuits	27	17	74.5	13.7	68.8	1.1
	Beam exciters	1	1	0.0	0.5	0.0	0.5
	Orbit control	0	1	0	0	0	0
	Software interlocks	-	1	-	0	-	0
	IT services	2	1	2.5	0	2.4	0
Beam	Losses	27	66	0.8	3.3	45.2	12.1
	Injection	29	8	9.2	2.3	9.1	2.3
	Induced quench	5	0	0.3	0	0	0
Operations	Access management	8	24	21.5	42.5	15.1	31.0
	Error, settings	44	23	11.4	3.9	11.1	1.9
	Sum	779	631	1620.5	892.7	1232.4	800.7

unavailabilities in the LHC during 2016 and 2017 operations and figure 4.17 visualizes part of the data[2]. Data are originally from references [145, 146].

The lead individual contributor for the LHC unavailability is the hardware availability of the injector complex. This is not surprising as it consists of five individual accelerators and the transfer lines. In 2016 although the lead causes for injector unavailability varied between machines power converters and/or RF system were among the top contributors in all machines [147].

The main contributors for the LHC unavailability are the superconducting magnets and the systems linked to it. The lead contributor of these systems is the cryogenics system. Cryogenics are required to maintain a 1.9 K operating temperature for the magnets. The system itself is quite reliable, but vastness and long recovery time make it a top contributor for the downtime. The system also suffers from secondary failures where the system environment does not allow

[2]In the data, access management is a generic description that is used for example to note deferred corrective maintenance and instances where handling of the access caused downtime.

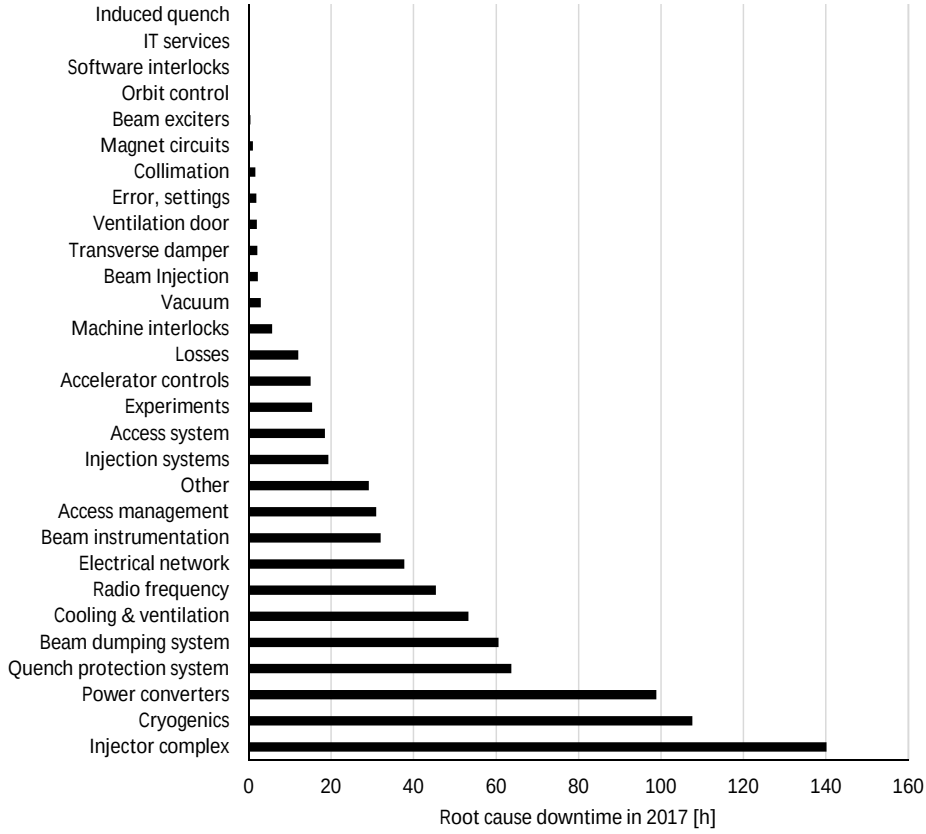

Figure 4.17: Root cause downtime duration for the LHC systems during 2017 [146].

maintaining the correct temperature. For example, power cut or a high heat load lead to an increase in the system temperature.

The LHC has the largest cryogenic system in the world consisting of 8 cryogenics plants [10]. Figure 4.18 shows the system layout and a layout of an individual plant where helium is cooled to 4 K on surface and to 1.9 K underground before it is distributed to magnet sections. Failure can lead to warming up parts of the cryogenics system. This warming leads to a long recovery time as the system needs to cool-down before operations can continue. One hour unavailability of a cooling plant leads to about 6-7 hours recovery time, where the plant cools the system to a temperature where the beam can be injected. Availability of the cryogenic system is further studied in the section 5.2.1.

Another system directly connected to the magnets is the power converters. Converters are needed as magnets require high amounts of Directional Current (DC) for generating the magnetic fields. The provided power affects directly to the magnetic field strength and thus high accuracy, reproducibility, and stability are required from the converters. A special requirement is that, as superconducting magnets have no resistance, power converters have to provide power with a low voltage and high current. Figure 4.19 shows the powering chain from a power plant to accelerator magnets and an illustration of the LHC power converters in an underground area. The main LHC dipoles are powered in eight circuits connecting the dipoles in series. In the LHC there are 1612

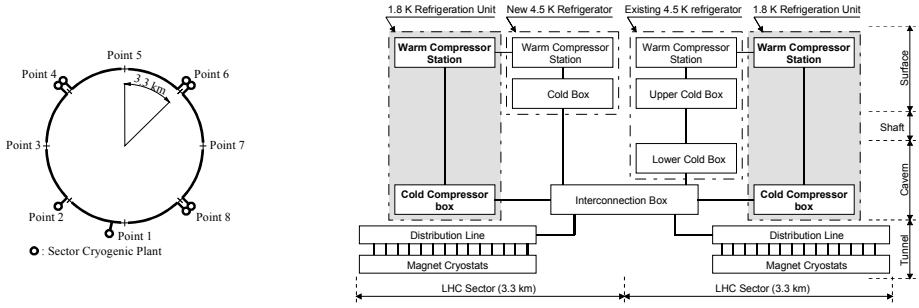

Figure 4.18: Layout of the CERN cryogenics system and the system architecture at points 4, 6 and 8 [148].

electrical circuits that have 1720 power converters. The converters can provide 1850 kA of current. The steady state power is 63 MW and the peak power is 86 MW. Common failures in the power converters are failures in electronics and embedded software.

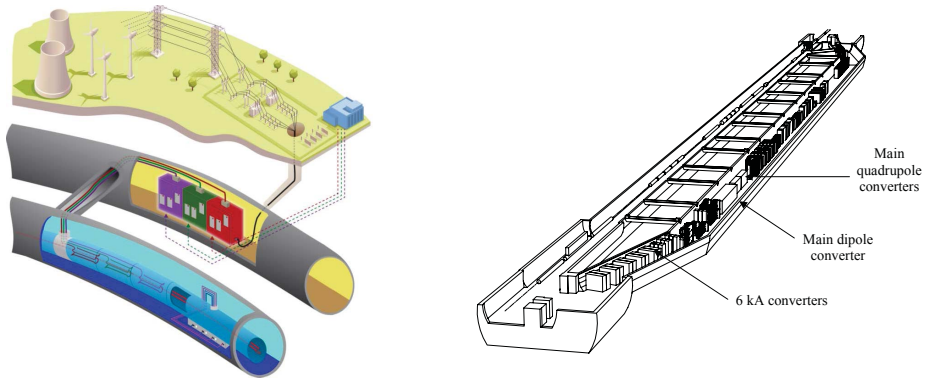

Figure 4.19: Powering chain of an accelerator magnet [149] and an example of LHC power converters from [10]

A special failure mode of the superconducting magnets is quenching. Quench is a sudden loss of superconductivity. When the magnet becomes resistive, it starts to heat up. There are two primary sources for quenches. A superconducting magnet has to undergo multiple training quenches to reach its full performance [150]. The training quenches are believed to be caused by small movements of magnets or crack propagation in epoxy resin that causes heat and leads to a quench [151]. These quenches have been also observed to occur during operations and they constitute to a significant part of the magnet circuit downtime in the table 4.2. The second primary source for quenches are the beam losses. Particles lost from the beam can deposit a sufficient amount of heat to the magnets to cause a quench.

A quench protection system (QPS) is required to heat the magnet evenly in order to avoid stress peaks that could damage the magnets. Without the protection, after a quench, the temperature in the resistive zone would increase within less than 1 s to 1000 K which would destroy the magnet. QPS is a part of a larger group of machine protection systems (MPS). The energy stored in the superconducting magnet system and beams can cause significant damage if this energy is not safely discharged [99]. The MPS's primary task is to prevent the energy to damage accelerator equipment. MPS consists of active and passive protection systems. Active systems detect faults

that can lead to damage and pass this information to start the energy deposition. In such situation, the beam is dumped using a specific beam dump system that is shown in the Figure 4.20.

Reference [152] classifies machine protection system downtime into two categories: downtime due to repairs caused by insufficient machine protection system and downtime due to complex protection systems. All the downtime marked for specific protection system goes to the second category. Despite this, it is worth to remember that a failure to protect the machine could lead to far greater consequences than a downtime due to complex protection system and reference [152] notes that there is a theoretical optimum between these two failure modes.

The QPS system failures in the table 4.2 relate to failures within the system itself. In this case, the most of these failures were within electronics boards that the system needs for functioning. Failures of the beam dumping system mainly relate to triggering mechanism that can activate spuriously. The system also has several kicker magnets that extract and dilute the beam to the bump blocks. These magnets require power converters that provide a quick power pulse. This system requires several parallel switches to hold the required voltage. The system also has in build redundancy that is maintained to ensure high operational reliability.

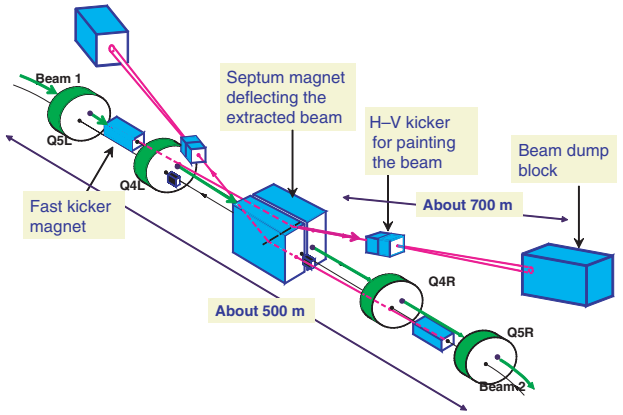

Figure 4.20: Layout of the LHC beam dumping system [99]

Passive protection systems are required as the active systems are not sufficiently quick to dump the beam in all situations. One example of passive protection system is called collimators. Particles are lost from the beam and collimators provide a controlled way to dispose of such particles. This helps to protect the LHC from beam losses, provide a cleaner beam for users and reduce the amount of radiation in equipment located in the tunnel. The figure 4.21 shows how staged collimators clean particles from the edges of the beam in an area called beam halo. The collimation system causes relatively low downtime compared to other systems as seen in the figure 4.17.

Other leading contributors for the LHC primary system failures are the injection system and the Radio Frequency (RF) system. In the LHC, the system consist of 8 RF-cavities per beam with an oscillation frequency is about 400 MHz [10]. In a cavity, the frequency is created by power coupler that transmits the power generated by the source to the cavity. In the LHC, each cavity is powered by its own 300 kW klystron [153]. A klystron is a linear beam vacuum tube that acts as an amplifier for RF power. Figure 4.22 shows the layout where waveguides connect klystrons, via a circulator and load, to their cavity coupler. A set of four klystrons share a high-voltage power supply installed on the surface. The primary failure sources of the RF-system are the klystrons, powering and control systems.

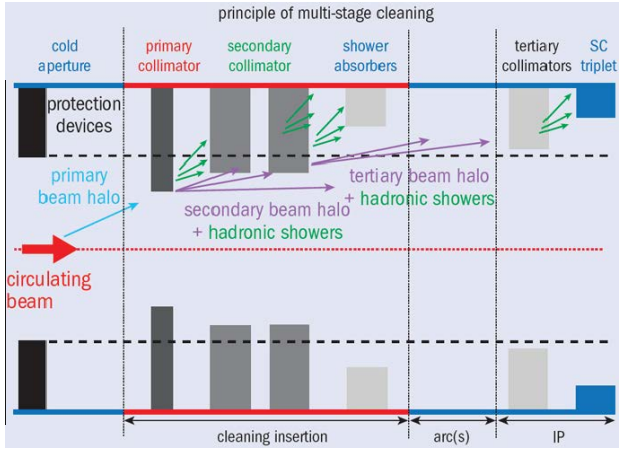

Figure 4.21: Schematic illustration of multi-stage collimation cleaning at the LHC. [58]

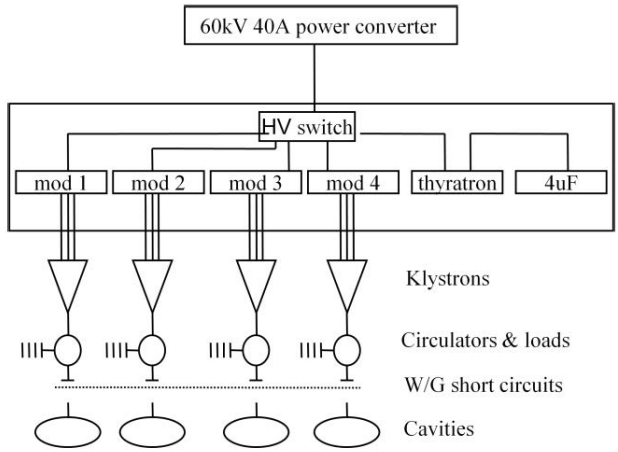

Figure 4.22: Layout of LHC RF-cavities powering [153]

The injection system consists of a set of septum magnets and a kicker magnet to inject the beam. A septum magnet has a magnetic field only one side of the magnet. They are used to steer the injected beam near the beamline so that the magnet field does not affect already circulating beam. A kicker magnet provides the final steering to inject the beam. A kicker magnet pulses quickly to provide a magnetic field only when a new batch is injected. This way an injection does not interfere with the beam before and after the slot where the batch is injected. The injection setup for single turn injection shown by figure 4.23.

In the table 4.2 beam instrumentation failures are also related to machine protection as they mainly consist of beam loss monitor failures [155]. The Beam Loss Monitoring (BLM) is designed primarily to protect the machine equipment against unintended energy deposition by losses. In the LHC the amount of particles lost from the beam is measured with ionization chambers. A beam generates a shower of secondary particles. An ionization chamber contains gas that the particle shower can interact with. The interaction produces ions and electrons that are collected

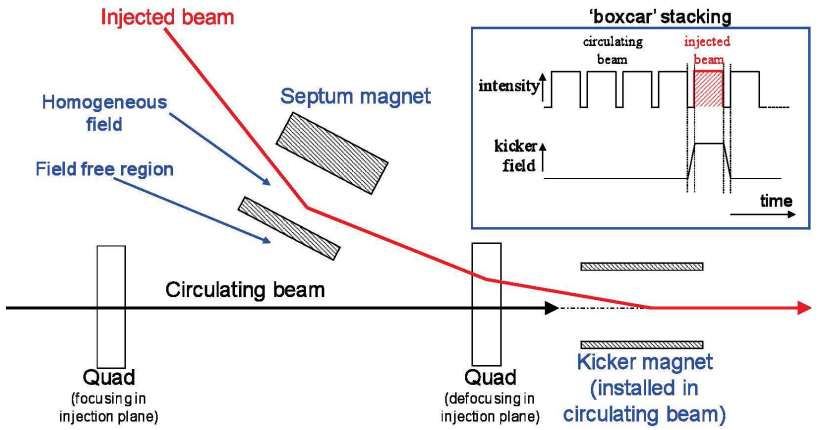

Figure 4.23: Pair of septum and kicker magnet inject a beam into a free slot [154].

with electrode plates (cathodes and anodes) inside the chamber. The resulting current between the cathode and the anode allows monitoring the beam losses. When a partial loss of the proton beam around the ring exceeds a predetermined threshold, the system sends a beam abort request. The figure 4.24 shows the ionization chamber for the LHC beam loss monitoring system [156]. The main source of unavailability for this system are the failures in built in tests that are intended to test system functionality before a fill. Other key contributors are optical links that connect BLM equipment in the tunnel to the equipment on surface and surface electronics [155]. The beam losses that have caused a beam dump are marked in the table 4.2 as beam losses or as beam injection if the losses occurred during an injection.

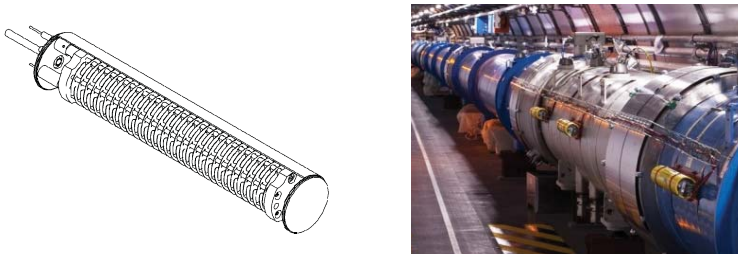

Figure 4.24: On the left, the ionization chamber for the LHC beam loss monitoring system with 60 electrode plates [156]. On the right, three yellow chambers installed to the LHC [157].

A particle beam can be thought as being electrical current traveling in the beam pipe. A particle beam can be though as being electrical current traveling in the beam pipe. An electric current produces a magnetic field and a changing magnetic field within a loaded coil of wire induces a current in the coil. The figure 4.25 shows operation principle of a beam current transformer that transforms the beam current to a current I_0 in an external circuit that can be measured [157].

The horizontal and vertical position of a beam in a beam pipe can be measured by measuring the image current associated to the beam. The figure 4.26 shows the operating principle of position measurement and a button sensor from LHC for position measurement [157]. A set of four sensors gives the beam position in a two-dimensional plane.

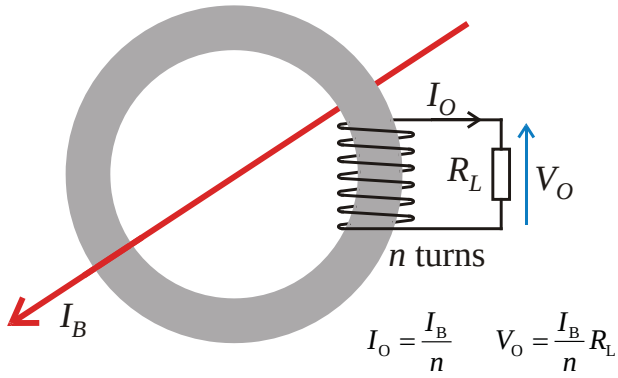

Figure 4.25: Operating principles of a beam current transformer: the current of the beam I_B creates a current in the coil I_0 and the resulting voltage in the circuit V_0 is used for measuring the beam current [157].

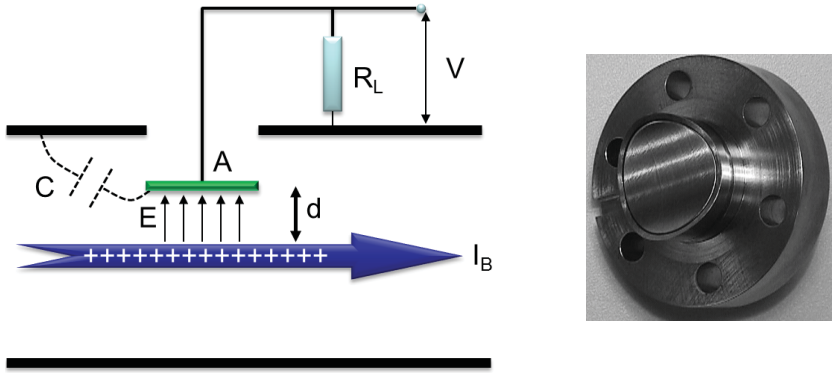

Figure 4.26: The image currents in the walls of a beam pipe are proportional to the beam intensity and are inversely proportional to the position of the beam. This current can be measured with specific sensors. The one used in the LHC is shown on the right. [157]

High beam energy and circulating beam limit the available options the measure the beam profile, as the measuring device must withstand the beam energy and must not destabilize the beam. The two options used in the LHC are wire scanner and synchrotron light monitor. A wire scanner moves a wire quickly through the beam. The interactions between the beam and the wire create secondary particles that can be measured with a scintilator[3] and a photon-multiplier. The quick speed is essential to avoid damage to the wire. Even then the wire scanner can only be used when there is a limited number of bunches in the machine. The scanner is mainly used to calibrate other sensors [58].

The LHC has a synchrotron light monitor for continues profile monitoring. At relativistic speeds, charged particles emit synchrotron radiation when they are accelerated orthogonally to their velocity. The figure 4.27 shows how synchrotron light is separated from particle beam at bending dipole [158]. Both the dipole magnet and the undulator are used to generate the synchrotron light. The undulator is a set of dipole magnets with alternating fields. It causes the particles to oscillate and to generate the light. Undulators are especially used in accelerators build to produce synchrotron light.

[3] A scintilator absorbs the energy from particles and re-emit the energy as photons.

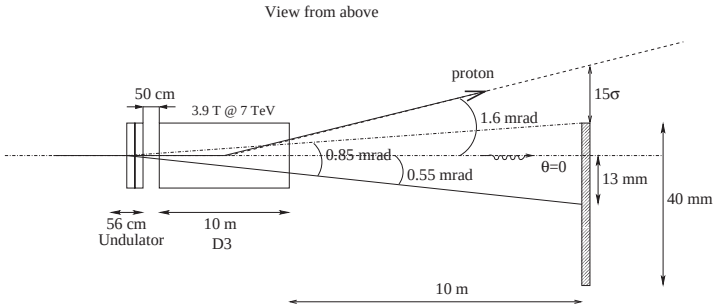

Figure 4.27: An undulator and dipole magnets generate and separate synchrotron light from the beam. The light goes to a mirror that guides the light to a monitoring system. [158]

The beam travels in vacuum. High quality vacuum is crucial to minimize beam interactions with other particles that would lower the beam lifetime. The vacuum system consists of pumps to create and maintain the vacuum conditions and sensors for monitoring it. The pumps and gauges work in pressure levels, so the system needs several sets of them to cover the whole region from low vacuum to ultra-high vacuum (10^{-9} Pa in the LHC [10]). The system is relatively reliable so it not further covered in this document.

Accelerator systems require electricity, cooling and pressurized air. The figure 4.28 shows the CERN power network and general concept of the cooling water distribution. Similarly to the power network, there are de-mineralized water, pressurized air, ventilation, control ssystems and communications networks covering the systems and tunnel areas [159]. Key failure modes of the electricity are the power cuts and small dips that can be caused for example by thunderstorms. In the LHC, these dips can cause a beam dump. In the LHC, electrical perturbations are noticed by the fast current change monitors [160] that protect the machine from inadvertent changes in magnet powering that would lead to a change in magnetic field strength.

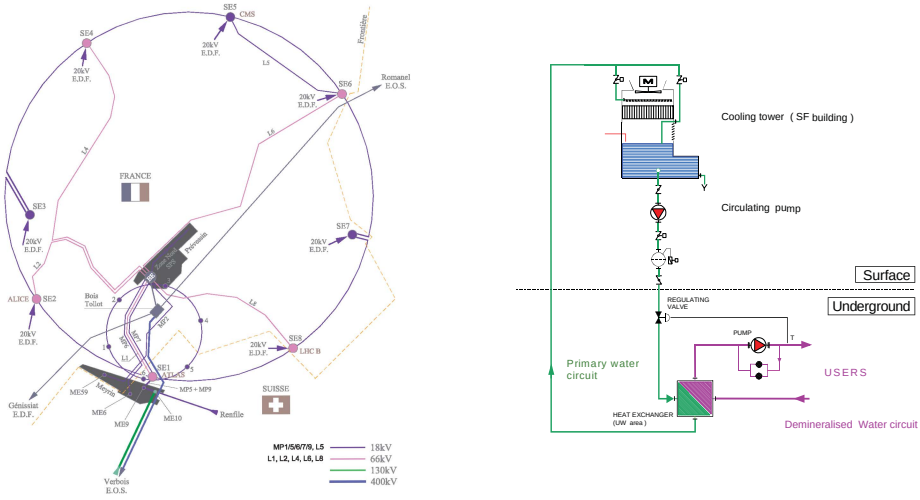

Figure 4.28: General layout of the CERN power network and a cooling circuit with [159].

Beam interactions in collisions are monitored by so-called experiments. The LHC has four beam interaction points that house experiments ATLAS, ALICE, CMS, and LHCb. Experiments contain sets of monitors that collect the data from collisions. This data are recorded in the CERN data center. Figure 4.29 shows the CMS detector and CERN computer center. A failure in experiments or in data processing can lead to a situation where the collisions cannot be detected or the data from them cannot be stored.

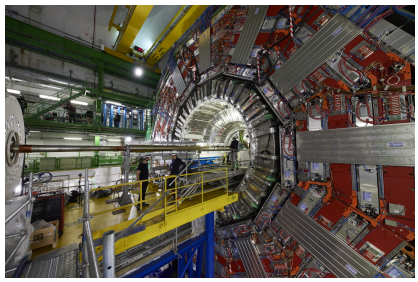

Figure 4.29: CMS detector [161] and CERN data center [162].

There are several hazardous elements that need to be taken into account when accelerators are accessed for maintenance. This requires special systems to limit and mitigate the risks. Operational accelerator produces high amounts of radiation due to stray radiation fields that are created at high-energy particle accelerators by beam interaction and beam losses on structural components of the machine [163]. Operational accelerator creates also so-called synchrotron radiation that occurs when a charged particle is accelerated. In a non-operational accelerator exposure to radiation is mainly caused by so-called residual radioactivity that is the result of radionuclides that were created by radiation activation during operations. To avoid radiation exposure, at CERN access to the machines is prohibited and monitored by an access system [100]. There also radiation level monitors in the tunnel.

Second major hazard is caused by the fact that the accelerators are often located in underground tunnels. This naturally restricts access to the tunnel. There are also risks related to asphyxiating gases that can be a result of fire or helium leaks[4]. Some of the mitigation measures for these risks are smoke and asphyxiating gas detectors in the tunnels. All in all, risk complicate the system maintenance and some additional downtime is caused by the non-critical failures of the access and safety systems, such as false alarms.

[4]Helium is used for cooling the superconductors.

5 Model for hadron collider availability for physics

The objective for the model is to resemble the operations as they are presented in sections 4.2-4.4. The model takes into account operational schedules, operational cycle, luminosity production, and failures. This is chapter presents model background, concept and special features that required developing custom code to add them into ELMAS model.

5.1 Background and modeling platform

The study chose A. Apollonio's model [18, 19] as the starting point. It was validated against the LHC operations in 2012 and has been used for availability assessment of the HL-LHC upgrade [18, 19]. The most recent published application is an analysis of crab-cavities effect on the luminosity production in HL-LHC [20].

The model simulates luminosity production based on beam parameters; turnaround and stable beams lengths distributions; MTTF and recovery time distribution. The model concept is best described by the figure 5.1 that shows the schematic sequence of modes in the simulation. A turnaround leads to stable beams during which the luminosity is produced. A failure can occur in two points of time: during stable beams phase or during the turnaround. A failure during a stable beams shortens the phase length and has a recovery time when the LHC is not available. Failure during the turnaround phase has a recovery time that prolongs the start of the next stable beams phase. Failure during a turnaround can originate from the LHC or from injectors.

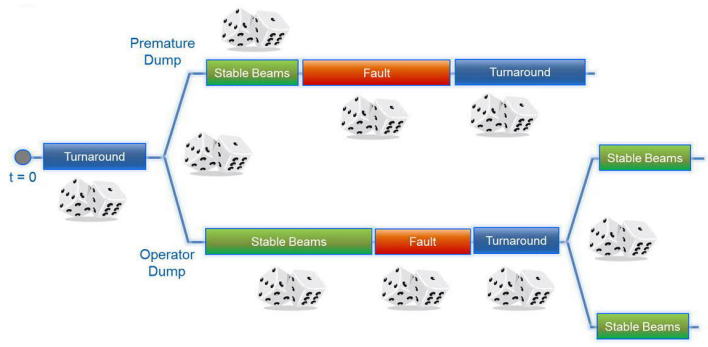

Figure 5.1: The schematic sequence of modes in A. Apollonio's model to simulate LHC luminosity production, as presented in [20]. Dice figures present stochastic randomization of the mode sequence and lengths.

The plan for the work was to increase detail in this model. The finished model should contain: (i) operation schedule modeling, (ii) turnaround time formalized as operations cycle with injection process model, (iii) failure modeling with system fault trees with operation cycle phase-dependent failure rates. The luminosity production was already handled well in A. Apollonio's model, so the plan was to link the production parameters to the operations schedule to formalize the luminosity production evolution during operations. The plan was also to study operations costs.

Ramentor's ELMAS software version 4.8 [21] was chosen as the model platform, thanks to connections that company then had to Tampere University of Technology. ELMAS uses Monte Carlo based discrete event simulations to model system reliability and operation expenses. As a built-in feature ELMAS support fault tree analyses, cause-consequence analyses and reliability block diagrams. ELMAS has about decade-long development history with several published use cases in various fields of industry: i) reliability analysis performed for a production-critical molding crane [164]; ii) maintenance robots for radioactive environment [165]; and power supply grid for a spent nuclear fuel encapsulation plant [166].

As a distinct feature, ELMAS allows users to add Java code to model special features [24]. This feature was used extensively to model the features described in section 5.3. Although this feature allowed detailed modeling of the operations, having large amounts of custom code in the model was clear shortcoming as it made model development somewhat dependent on the individual developer who created the code. Also, custom code in the model creates a high skill floor as the modeler needs to understand both the modeling approach and programming. Collaboration between CERN, Ramentor and Tampere University of Technology was launched to address this shortcoming. The result of this collaboration is OpenMARS which is an open modeling approach for availability and reliability of systems [2]. This approach is further explained in the chapter 8.1.

5.2 System failure modeling

5.2.1 Fault tree for the cryogenic system

ELMAS is a fault tree software and creating fault trees in this software is easy. The problems in this task is related to a lack of data on the failure rates and repair times. This made developing detailed system fault trees seen as an unnecessary task as without this data the detailed fault tree cannot be utilized. Also, evolving FCC design and uncertainties in this design made this task seem impractical.

However, as a case study, a high-level fault tree of the CERN cryogenic system was developed[1]. The basic concept is from NUREG fault tree handbook that distinguishes failures to primary and secondary failures based on system boundary. Figure 5.2 shows this concept. Primary failures have a root cause within the system boundary and they occur in an environment for which the system is designed. Secondary failures have a root cause outside the system environment. Secondary failure occurs when the system environment cannot support the system operations due to a failure outside the system boundary. The failures outside system boundary were further distinguished to failures related to utilities that system functions depend on and failures related to users of system function. Within the system boundary, the system is divided to subsystems with different resolutions based importance of the subsystem.

The first level of the fault tree was formed by categorizing the sources of unavailability in three categories: Primary failures, secondary failures and unavailability due to planned maintenance.

[1] The author worked for this task with B. Bradu and P. Gayet from CERN and E. Rogova from TU Delft. E. Rogova published her results on this study in [167].

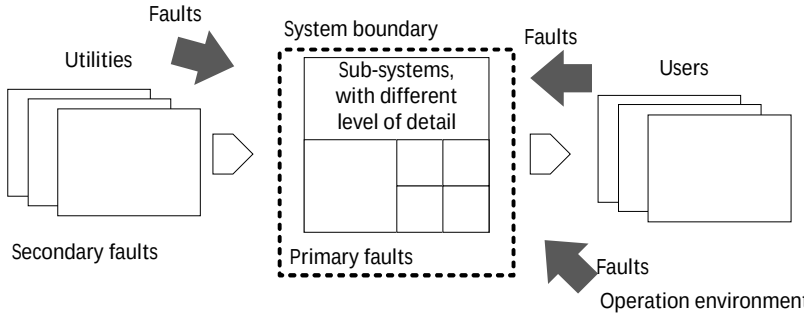

Figure 5.2: System boundary defines the primary and secondary faults based on the origin of the root cause. Faults are further classified based on the system outside of the boundary or subsystem inside the boundary.

The first level tree is shown in the figure 5.3. Unavailability due to planned maintenance is not a failure. However, the categories in the tree were used as a basis to form categories in a maintenance database, so this was included for the sake of completeness.

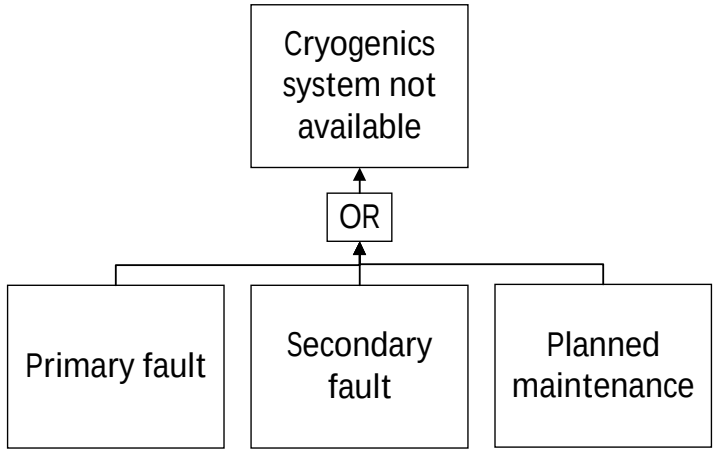

Figure 5.3: Top level classification for cryogenics fault tree.

Secondary failures branch contains utility, users and environment categories, as shown in the figure 5.4. Utilities are systems that cryogenics operations rely on: such as cooling and ventilation, electricity, controls network and a specific vacuum system to insulate the distribution pipes. Users need cryogenic cooling for their operations, but can also cause cryogenic system unavailability. At CERN key user are the superconducting magnets. Their main failure mode that affects cryogenics is a quench as it heats the cryogenic system. The only identified and relevant failure category related directly to the environment in this tree are single event upsets that are electronics failures caused by radiation [168]. This decision was controversial as on the other hand side a SEE is a failure of the electronics and can be made less likely by improving radiation hardness.

The fault tree of primary failures of the cryogenic systems contains categories for equipment, local control systems, human factors [169] in operations and maintenance. This tree is shown in the figure 5.5. Equipment category is the most elaborate. The equipment is first classified by

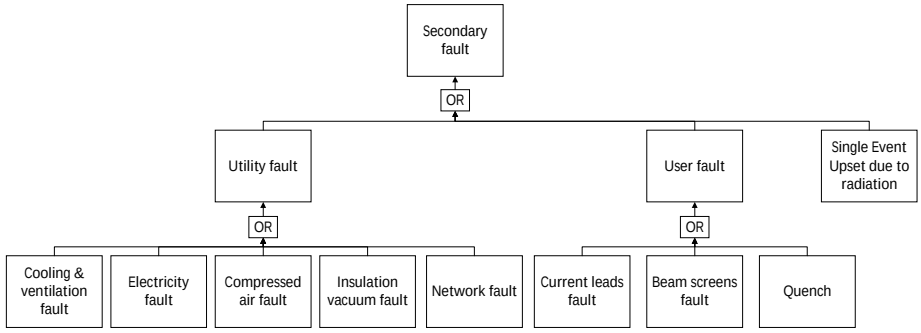

Figure 5.4: Categories for secondary faults.

location: 4.5 K refrigerator, 1.8 K refrigerator, and distribution. The distinguishing difference between the locations is the access time 4.5 K refrigerators are on the surface, 1.8 K refrigerator are underground, and distribution can be anywhere in the LHC tunnel. A specific category was formed for process optimization failures. An example of such issue is presented in [170] that describes how higher than expected heat load to the beam screens was compensated with improved control schema. Before this schema, the operations were frequently halted when parts of the system were overheated.

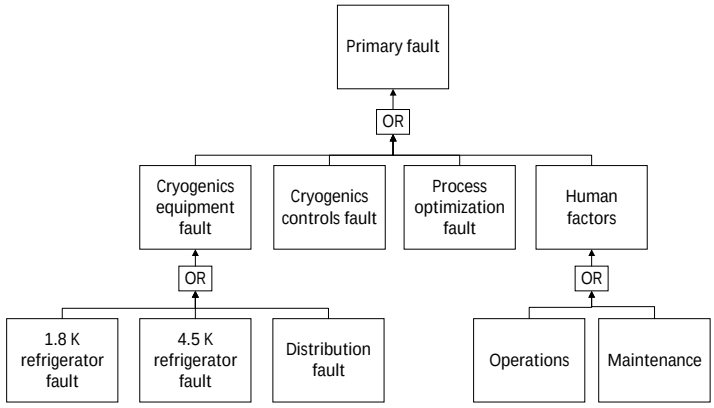

Figure 5.5: Top level categories for primary faults.

In the equipment fault classification, shown in the figure 5.6, special attention was given to rotating machines in the cryogenic system [171]. These categories include compressors, turbines and oil and vacuum pumps. Other categories tried to list the active equipment in the system such as instrumentation, valves, and heaters. The passive equipment that does rely on external inputs [172] was categorized separately. In the LHC main failures in this category relate to equipment that maintains the quality of the cryogenic fluid, such as filters and absorbers.

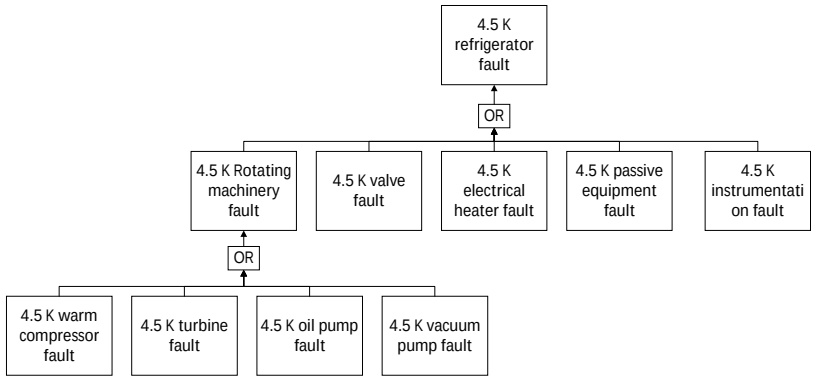

Figure 5.6: Categories for 4.5 K refrigerator faults.

The aim was to use these categories first to understand the current unavailabilities and then to see how the complexity of each category would scale for the FCC. The problem was that the statistics were not collected to this level of detail from multiple years and the plan for the FCC cryogenics system design was quite open at that point. For example, the number of the cryogenics plants was not yet decided [173].

5.2.2 Parameters depicting failure and recovery behavior

Original aim for model parameters was highly ambitious. The model would have different failure rates for systems for each operation phase and failure rates could depend on beam parameters or system age. However, the lack of the data stopped most of these ideas.

Phase-dependent failure rates require information on the active phase at the time of the failure. This information is difficult to attain as the failures are recorded manually and do not contain information on the active mode. Linking a fault to a mode afterward is hard as the failure time is recorded manually. In some cases, the operation mode when the failure occurred was not the same mode that was present at the recorded failure time. This situation often occurred when a failure triggered a beam dump. Also, the recorded operation mode is not always descriptive of what happens with the machine. For example, injection mode can be present when the LHC is still practically within setup mode. Also, the sample of the studied equipment and the observation time are limited. So achieving statistically credible result where the failure rates are different for different operation phases is not apparent. Due to this reason, the phase-dependent failure rates were used mainly to model failures that were directly linked to the beam. This was relatively easy as these failures can only occur in phases where the beam is present.

As mentioned in the section 3.1.1 utilization of age-dependent failure rates can give valuable insight into a system. However, the effect of aging on the failure rates is hard to measure or prove as the beam parameters change over the year and between operations years. This makes it difficult to study if a particular failure was caused due to aging or due to beam parameters and to say that the effect is statistically credible. Also, the study of the effect of aging would require intrinsic knowledge on how individual systems were maintained during their life, as the maintenance can affect the system reliability. For example, if a system was renovated during a technical stop and this is not taken into account, an analysis could, in the worst case, interpret teething related failures as aging-related failures. Moreover, failure data collection practices were improved at CERN between 2012 and 2015. This change of practices can lead to higher failure rates as the failure events are more likely to be recorded.

Plan for the recovery time modeling was also ambitious. The idea was to divide the recovery time into four phases: (i) failure identification, (ii) failure diagnostics, (iii) logistics, and (vi) restoration. This division is shown in figure 5.7. The idea would have been possible to implement in ELMAS. It was also abandoned because the available data could not distinguish these kinds of details in recovery time.

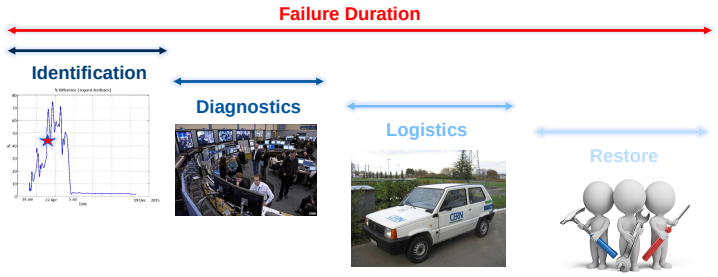

Figure 5.7: An idea to divide the recovery time into individual actions.

Finally, in all modeling, the data that was used was not too different from the high-level data presented in the table 4.2 and this data was used to calculate high-level MTTF and MTTB values.

5.3 Technical considerations on special model features

Multiple features of the model go beyond the normal scope of classical modeling methods and features that were supported by the ELMAS software during the project. All of these features required additional code to implement them that was one of the key tasks of this thesis.

5.3.1 Schedules

The schedules were implemented as Markov styled state models where states define different scheduled operations. However, in the project, this was seen as a low priority item, so the model is not highly developed. The model presented here is the only initial version that was developed to create LHC schedules.

The states and transitions of this model are shown in the figure 5.8. On the top, is the model for the long-term schedules. It contains just two states: shutdown and run. In a typical modeling scenario, exact deterministic transition times would be given. For example,a shutdown takes exactly two years and a run exactly 3 three years.

The lower part of figure 5.8 shows the model to create annual schedules. During a simulation, this model is active when the run state is active. The model contains scheduled operation modes: long technical stop (LTS), hardware commissioning (HWC), commissioning with a beam (CWB), proton physics, machine development (MD) and technical stop (TS). In the current model implementation the "END" state contains both the ion physics and the year-end break. The study elected not to model those modes in detail. After the "END" state, a new year starts from LTS state.

The inputs for the annual schedule model are: (i) length distributions for different operation modes and (ii) number of production stops (MD, TS and MD+TS periods) per year. The length of the

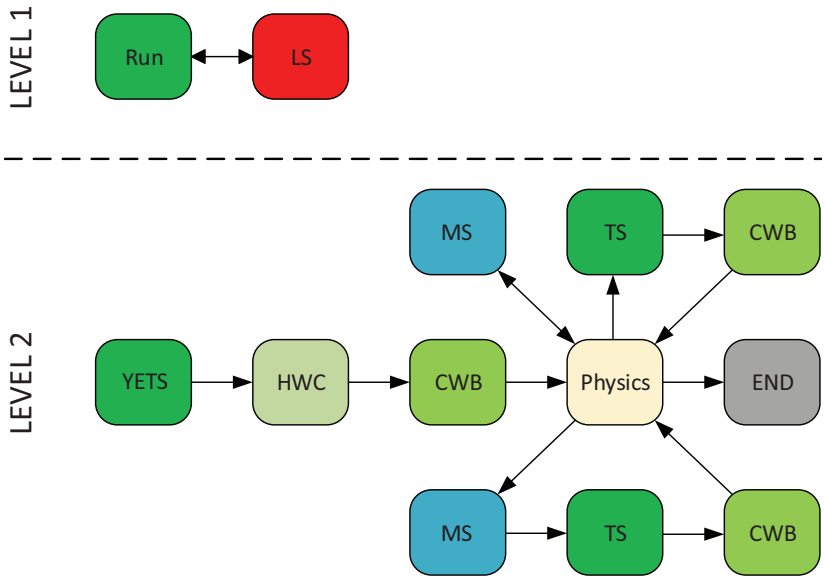

Figure 5.8: Operational phases for the schedule model [1]

physics period is not given it is calculated based on the lengths of the other operation modes. The operation schedule for a year is created in the Java library before the luminosity production is calculated.

The Java library forms a LHC schedule in the following way. First, the inputs are used to calculate the length of operation modes. If a specific mode is reached multiple times during a simulated year, multiple length values are calculated for this mode. The time for physics operations is calculated by subtracting the sum of simulated values from the length of a year (365 days). Time for physics is further divided into smaller periods based on the number of operation stops per year. Figure 5.9 shows how three operation stops divide the operations into four periods with equal lengths. For example, the table 5.1 shows an example of possible inputs for the model. From these inputs, the model contracts the schedule seen on the figure 5.10

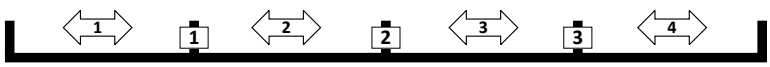

Figure 5.9: Three stops create four equally long operation periods.

Table 5.1: Example of potential inputs for the schedule model.

Year	1	2	3	4
Long shut down	0	1	1	0
Year-end technical stop	30	0	0	30
Variation	5	0	0	5
Hardware commissioning	30	0	0	30
Variation	5	0	0	5
Beam commissioning	30	0	0	30
Variation	5	0	0	5
Number of tech. stops	3	0	0	3
Number of machine dev.	1	0	0	1
Number of MD+TS	0	0	0	0
TS	5	0	0	5
Variation	1	0	0	1
MD	3	0	0	3
Variation	0.5	0	0	0.5
Short beam commissioning	2	0	0	2
Variation	0.5	0	0	0.5
End(Ions+TS)	60	0	0	60
Variation	0	0	0	0

The simulated values are stored as lists in a map where the key is the name of the operations mode. This procedure is repeated each time a year starts. Additionally, a list is formed based on type and number of production stops. Order of this is randomized at the start of every year. ELMAS uses the length value map and the stop list during simulations. Each time a state in the annual operation modes model starts, ELMAS accesses the map with the mode name, and the map returns the next mode duration from the specified list. Also, after each time physics mode ends, ELMAS accesses the production stop list to ask what is the next mode. Once the list has no more items, the "END" mode starts.

The current implementation of the schedule model is not in many ways ideal. The length of the modes is calculated before physics production is simulated. Due to this, the schedule is not a dynamically changed during simulation. In real life, the schedule could be changed as a result of long failures or the high amount of production. Also, the model assumes that the production periods are organized by an annual basis. However, this could be easily changed in further development.

This kind of model would be useful if technical stops were to be linked to specific tasks like deferred corrective maintenance or preventive maintenance that aim to improve item condition were modeled. Example of a case with deferred corrective maintenance can be found in [96]. However, this was not the case in this project, as so relatively low priority was given to the schedule model.

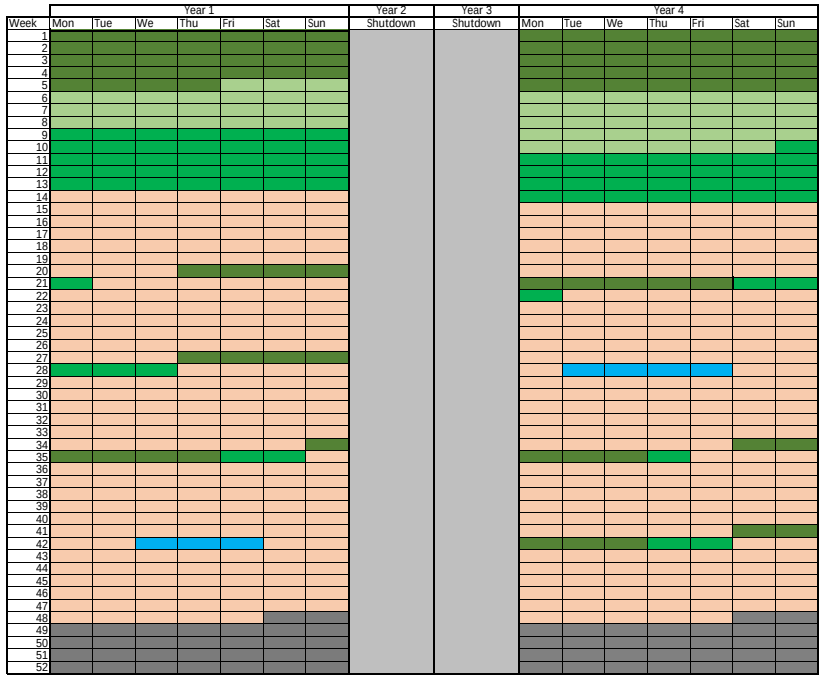

Figure 5.10: An example schedule based on inputs in the table 5.1.

5.3.2 Operation cycle

During a physics mode in a schedule model, operations cycle and system failures are simulated. Theoretically operation cycle model is a semi-Markov model where the state transition times depend on transition distributions in the Markov model and system failures that are simulated with fault trees depicting different accelerators in an accelerator complex.

The operation cycle model aims to produce paths shown in the figure 5.11 where individual turnaround phases are modeled and a failure can occur at any phase. This figure can be compared to the A. Apollonio's model shown by the figure 5.1. The main difference between models is the modeling of the turnaround where the new model takes into account all the phases, and if the failure rate is high, the process has a low probability to reach the collision delivery. Also, the failure occurrence time is truly randomized so that a failure can occur at any moment in each phase.

If no failures are present, the first version of the model functioned based on transition distributions in the Markov states. Figure 5.12 shows these transitions with green color. The latest version of the model adds more detail to the modeling the injection phase that is described in the section 5.3.5.

Failures in the fault trees affect the operation phase model by activating and inhibiting certain transitions. In practice, there are three cases where special rules are used. Figure 5.13 shows in red the transitions that are affected by failures. The cases are: (i) If a failure occurs in a collider when the beam is present, a beam dump is performed. In the model this means that a transition from an injection, ramp or stable beams phase to ramp down phase is performed; (ii) If a fault state is present in the collider or in an injector when the idle phase is active, the model waits

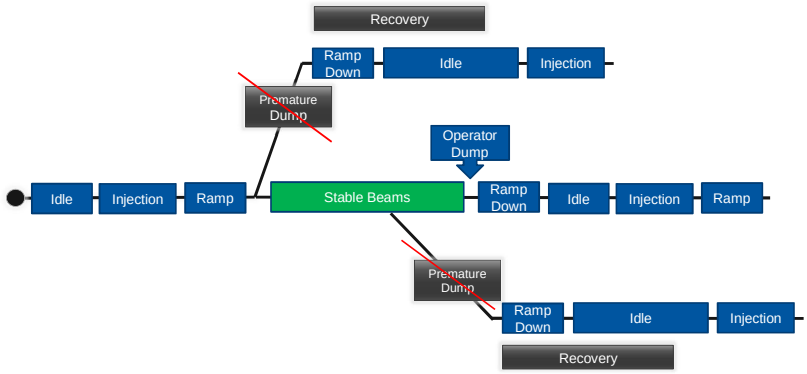

Figure 5.11: Example of possible phase sequences in my model.

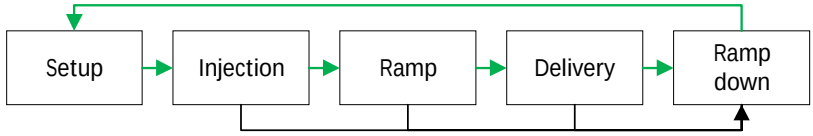

Figure 5.12: Green color highlights transitions that occur if no faults are present.

that the failure is repaired before injection starts. This feature is implemented by inhibiting the transition to injection phase when a fault state is present; (iii) Model waits that the collider is filled with beam before a ramp starts. This behavior is implemented by adding the fault time to the overall phase time.

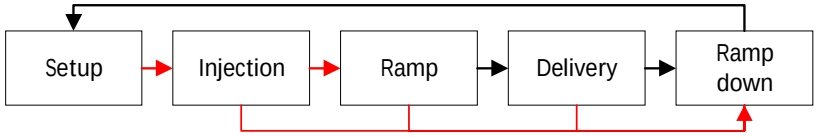

Figure 5.13: Red color highlights transitions that are affected by failures.

Modeling the operations with these rules do not fully depict the reality. The last case iii can be seen as unrealistic if the failure takes a long time to recover. The beam quality decays over time and waiting a long time in injection energy should affect the physics production rate. Also, at some point collider operator would dump the beam and restart the fill or go to stable beams phase with the partially filled collider. Implementing this would require Luminosity production model that takes into account this decay and new decision logic model.

Another way, the model does not depict reality is that if operators know that an injector is faulty, they can elect to keep the beam in the collider longer than the optimal time. In the model, this would mean that the transition from the stable beams to the ramp down would be inhibited if an injector is in a fault state. However, this rule could also be unrealistic if a fault is quick to recover. As before injecting the next beam, the collider will go through ramp down and setup phases. The time spent in these phases makes a minor delay unnecessary.

5.3.3 Luminosity production

The luminosity production modeling is linked to the operations cycle. After every stable beams phase, ELMAS sends the length of this phase to the Java library. The library calculates the luminosity production using production functions shown in section 4.3.3. The luminosity production evolution that is shown in figure 4.8 is implemented by tracking the stable beams hours. Different beam parameters are used based on total time spent in stable beams. These parameters can be different for separate years. Also, the intensity ramp after a short technical stop is modeled by counting the stable beams hours after the maintenance. The implementation is shown in the table 5.2 that shows potential input values for a simulation. Table 5.2 shows the evolution of the parameters for several years of operations. The evolution after a technical stop is defined in another table that similar to the table 5.2.

Table 5.2: Luminosity parameter input table. Del_Time parameter tells the total amount of hours the specific parameters are used; beamType tells which method is used for calculation. Here the parameters for a LHC type beam are shown. For a FCC beam, a file path would be given that contains the pre-calculated function values for different delivery times.

Year	Del_Time [h]	beamType	fillTime [s]	peakLumi[a]	lumiLife [h]	levelingTime [h]
1	50	LHC	4000	1	20	0
1	50	LHC	8000	5	15	0
1	50	LHC	12000	10	12	0
1	inf	LHC	36000	20	10	0
2	50	LHC	4000	1	20	0
2	50	LHC	12000	10	12	0
2	inf	LHC	36000	20	10	0

[a] Peak luminosity unit is [cm^{-2} s^{-1} $*10^{33}$]

A luminosity production for an individual delivery phase is calculated based on values coming from the simulation: i) the phase length; ii) number of the simulated year; iii) the total number of hours spend in delivery phase during that year; and number of hours, in delivery phase, after last technical stop. The code stores the values from the input data table in memory. This defines the luminosity production function and parameter values in combination with the hours spent in the considered delivery phase give the correct luminosity production value. After this is calculated, the code checks if the parameters need to be updated. If this is the case, the code accesses the input data tables where it chooses i) correct table regarding the hours since the last technical stop; ii) correct row based on year and hours in delivery. If the intensity ramp after short technical stops is modeled, the technical stop phase in the simulation updates the beam parameters to use values from the other table.

5.3.4 Phase-dependent failure rates

During the project, two versions were developed on how to implement phase dependent failure rates. The first implementation with multiple fault trees was deemed not ideal as it made the model unnecessary complex and lead to problems with fault recovery. The second version was developed once the ELMAS supported the functionality to have multiple failure rates per a fault node. The key use cases for this feature are the failures related to the stored energies in beam and magnets. One such case is beam losses that result in a beam dump to protect the machine.

Studies have been conducted on beam losses leading to quenches. During the first run, operational beam losses leading to quenches occurred exclusively during the during the injection phase [174].

During the second run, the beam energy was increased and beam-macroparticle interactions started causing quenches [130]. These quenches occurred mainly in the high energies (after ramp phase).

5.3.4.1 Multi-branch fault trees

The first version of the model had multiple fault trees to depict phase dependent failure rates. These trees were activated and deactivated based on the active phase in the operation cycle model. Each phase had a specific fault tree and additionally, there was a fault tree for failures whose failure rate did not depend on the active phase. Figure 5.14 shows how this was implemented.

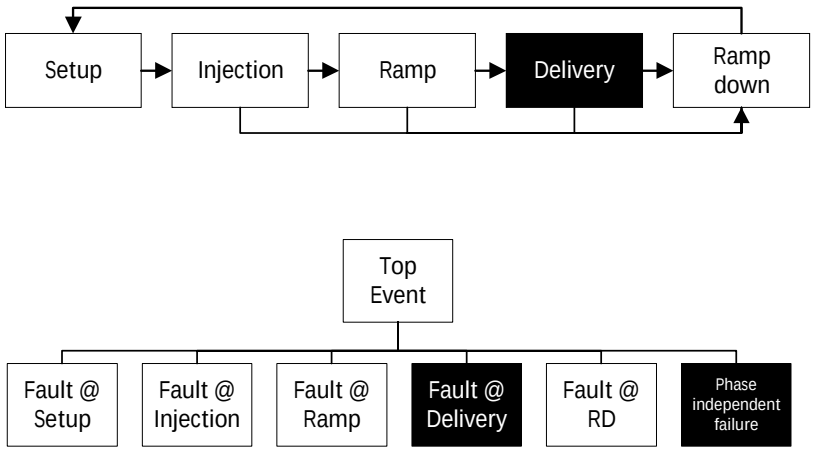

Figure 5.14: When the delivery phase is active in the Markov chain, the corresponding fault tree branch and the branch for phase independent failures are active in the fault tree.

In practice combining fault recovery and fault tree branch deactivation prove to be problematic. The idea was that once a phase ends, the related fault tree branch is set to be inactive. However, if the phase change was activated due to a phase dependent failure, then setting a branch in an inactive state fixed the failure. the problem is demonstrated in the figure 5.15. This problem limited the use of phase-dependent failure rate feature to the beam losses[2]. As a beam loss causes in the most cases an instantaneous beam dump that does not require repair or recovery time[3]. A decision was made not to fix this error. Instead, it was decided to wait that the multiple failure rates for a fault node were implemented as a software feature.

[2] Accelerator loses beam constantly. In this case, the term beam loss as a failure refers to a situation where the loss rate is too high to maintain safe operations. This triggers a beam dump and decreases the availability for physics.

[3] In some cases this is a gross simplification, as noted earlier a beam loss can lead to a magnet quench with long recovery time.

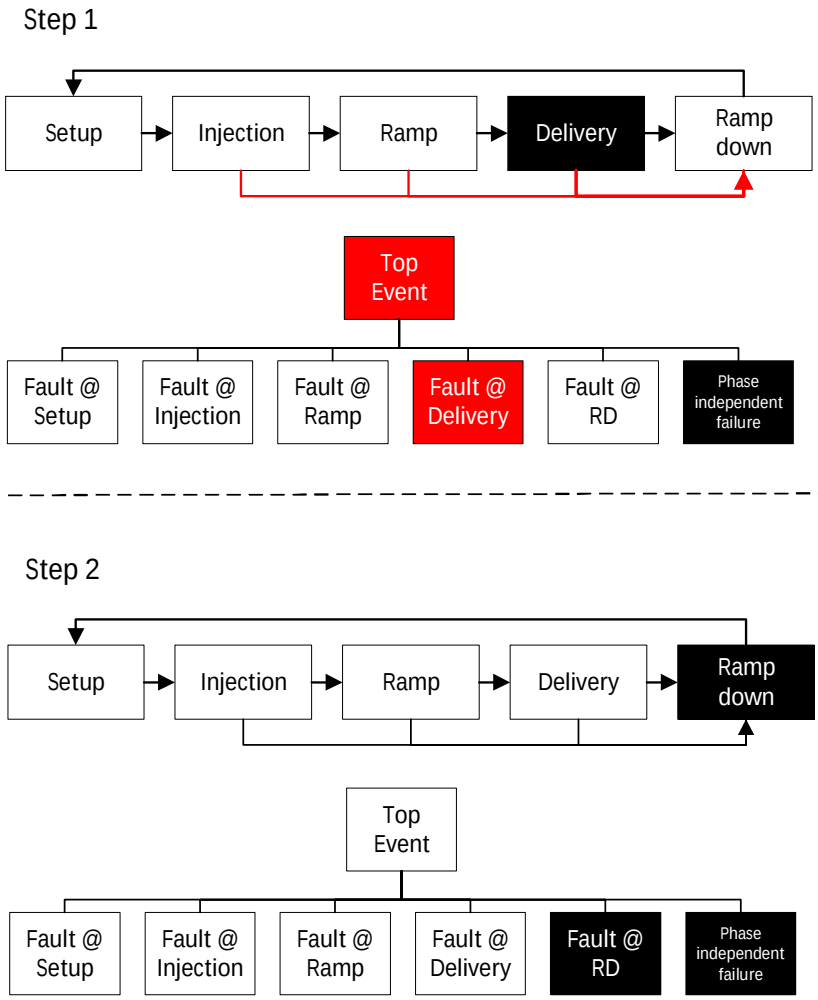

Figure 5.15: Step 1: Failure occurs during a delivery phase and the actual phase shifts to ramp down. At the same time, the fault tree changes the active branch. This leads to an erroneous deactivation of a fault state for the step 2.

5.3.4.2 Multiple failure rates for fault nodes

The second implementation allowed setting multiple failure rates for a fault node. Practically this is implemented by having multiple tracks that model failure behavior in different modes. In the simulation, only the track related to active mode is progressed and the other tracks are halted. Figure 5.16 shows the concept.

Each node also has a non-mode dependent failure rate that is always active, as the implementation does not limit how many modes are active. In practice, however, it was chosen that: (i) A node can only have mode dependent failure rates or a mode independent failure rate, not both; and (ii) Only one mode is active at any particular time for a specific accelerator.

The implementation works if constant failure rates are assumed. However, if the item's age affects the failure rate, this technique would not work. In this case, an item could be assumed to age in

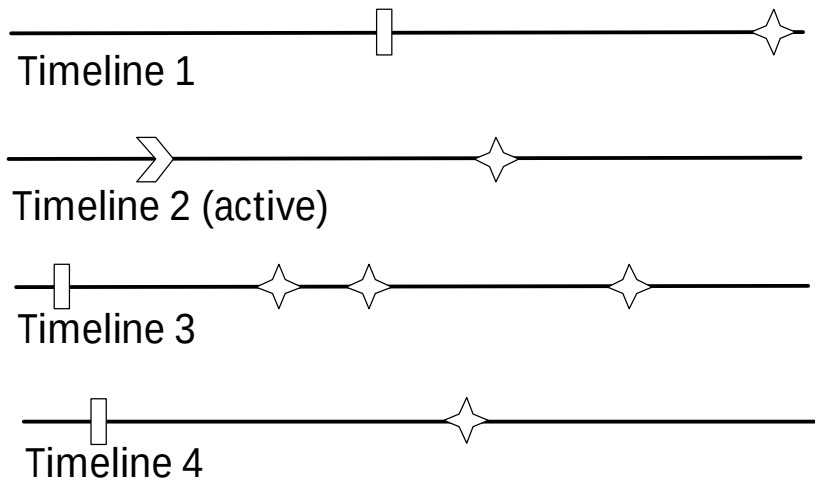

Figure 5.16: Timeline 2 is active and progresses, while others are halted. This approach allows modeling different failure rates in different operational phases. However, the timelines are not linked and do not share a common history, which complicates modeling age-related failures.

different rate during different phases (i.e., stress scale). Example of a case with this assumption can be found where stresses that system is subjected to affects the Weibull scale factor [175].

In practice, both assumptions can be useful. The ideal implementation would be to let the users choose the assumption used in the simulations. It is worth to understand that multiple time and stress scale assumptions are used to improve the calculation speed, as the already simulated events can be stored when a new mode is activated. If all failure events were recalculated each time a new mode is activated, it would allow more freedom to set assumptions.

5.3.5 Injection process

As described in section 5.3.2 injection phase starts by injecting pilot bunches and measurements after which certain amount of successful injections are required to fill the collider.

Chronologically the injection process modeling was the last feature to be added into the model. Implementing this benefited from features in the prior versions of the model. The model takes the concept of the operational cycle and tries to model how many successful cycles an injector needs to perform to fill the next level machine (collider or another injector) this is shown in the figure 5.17. Here the phase stable beams is replaced by term delivery. Also, it was assumed that an injection could cause a failure that results in a beam dump.

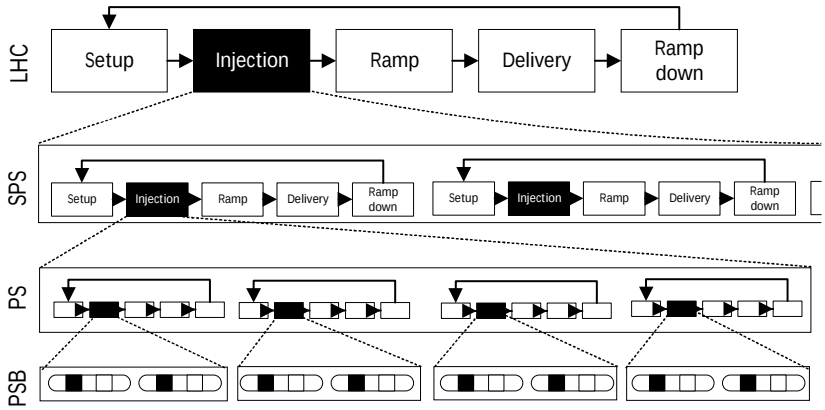

Figure 5.17: When injections are modeled an injection phase of the LHC consists of a sequence of SPS injections where an injection phase consists of a sequence of PS injections.

At the start of the injection phase, the Java library forms a list of boolean values. The length of the list denotes the number required cycles to fill an accelerator. In this first version, the boolean value specifies if the cycle causes a failure in the accelerator. This logic could be extended to have multiple failure modes that an injection could activate. In that case, the list could consist of enumerals that could depict if an injection is successful or causes a specified failure.

The list is formed with a method in the Java library that is presented in the listing 5.1. This method has the following inputs: (i) Number of required injections; (ii) Success rate of an injection; (iii) Probabilities that successful or unsuccessful injection causes a failure that leads to beam dump; (iv) Injector super-cycle, e.g. today only every other injection goes to LHC from SPS; and (v) Additional cycles that are used for taking into account e.g. time required for measurements with the pilots. These values are stored in additional columns with the beam parameter table 5.2.

This paragraph describes the code shown in the listing 5.1. The first line defines the method name and the fact that it returns a list of boolean values. The line three generates the list object that the method returns and the rest of the method fills it with boolean values. The majority of these values are created using a *"while"*-operator (rows 8 and 10-33)[4]. It generates more items to the list until the required number of successful injections is reached.

The first task of this operator is to check if a new injection is successful. The method generates a random number that is compared to the success rate in row 13. If an injection is successful in the lines 15-19 the number of successes is increased and a new boolean object is added to the list. This object is generated with an additional method shown in the listing 5.2. This method tests if an injection caused a failure using an input probability and random number generator[5]. If instead, an injection was unsuccessful a new injection object is added to the list with a similar fashion in line 25.

After an injection, the method checks, in line 29, if there is an operational super cycle and if the number of required successful injections has not been reached. If this is the case a support method shown in the listing 5.3 is called to add the required number of objects to the list. Once the required number of successful injections is reached in the main method list is returned in line 35.

[4]The row 6 and additional cycles are discussed later in this section.

[5]There might be an opportunity to optimize the code as the method is let to finish the list even-though it might contain a failure that stops the injection process.

Listing 5.1: Simplified Java method to form injection cycle list.

```
1  public ArrayList<Boolean> getRequiredCycles(){
2
3  ArrayList<Boolean> rList = new ArrayList<Boolean>();
4
5  //add additional cycles
6  returnList = fillList(addCycles, rList);
7
8  int successful = 0;
9
10 while(successful < required){
11
12 //Check if injection is successful
13 if(injectionSuccess > rng.nextDouble()){
14
15 //Increase the number of successes
16 successful++;
17
18 //check if a successful injection causes a beam dump
19 rList.add(faultCheck(successProb));
20 }
21
22 else{
23
24 //check if a unsuccessful injection causes a beam dump
25 rList.add(faultCheck(failProb));
26 }
27
28 //Check that there is a super cycle and the level is not yet filled
29 if(superCycle > 0 && successful < required){
30
31 //Add new cycles to list based on super cycle
32 rList = fillList(superCycle, rList);
33 }}
34
35 return rList;
36 }
```

Listing 5.2: Support method to check if injection causes a failure.

```
1  protected boolean faultCheck(double probability){
2
3  //Fault
4  if(probability > rng.nextDouble()){
5  return true;
6
7  //No Fault
8  }else{
9  return false;
10 }}
```

Listing 5.3: Support method to add injection to the list to resemble a super cycle.

```
1  protected ArrayList<Boolean> fillList(int number,
2   ArrayList<Boolean> list){
3
4  //Adds cycles to the list that cannot cause a fault
5  for(int i =0; i < number; i++){
6  list.add(false);
7  }
8
9  return list;
10 }
```

This version was made mostly to test and demonstrate the injection process functionalities. Due to this reason, it is not as finalized as other features. One relevant simplification is done in row 5 of the method where the support method seen in the listing 5.2 is called. The additional cycles are now added to the start of the list. These cycles resemble the time when the collider would be doing a calibration with pilot beams. Due to this simplification, pilot beams does not start the injection phase, and success rates of pilot and nominal injection are treated as equal. In the simulation, the only significant impact of not starting with pilot beams is that an injection-related failure cannot occur at the start of the phase. However, treating pilot and nominal injection success probabilities as equal is more significant [176] and should be corrected in the future. The reasons for the simplifications were that: (i) data from pilot and nominal injection success probabilities were not available while writing this code; and (ii) Adding this detail would have significantly increased the detail of the input file and code.

Once the list is produced, it is delivered to ELMAS software. During an injection phase, ELMAS reads the next item from the list each time an injector reaches the delivery phase. If the boolean item is "true" a specific node from the fault tree is set to the fault state and a beam dump is triggered in the machine that the injector feeds. Else if the boolean item is "false" a new cycle is started in the injector or if this item was the last one in the list the next level machine finishes its injection phase and starts the ramp phase.

The process is generic so it can be applied regardless whether the next level machine is the collider or another injector. The rules developed for a collider cycle could be copied to form an injector cycle: (i) if the injector chain for an injector is not available the injector waits in the setup phase; (ii) a failure while the beam is in an injector will cause a beam dump; and (iii) active phase of an injector can affect its failure rates.

The length of injector cycles is in the order of seconds. Modeling these cycles would have significantly increased the number of events in the simulation and so the calculation time. For this reason, a limit in resolution was chosen and detailed modeling of the injectors was stopped at a certain level. For the LHC only the SPS, and for the FCC the operation of LHC and SPS injectors is modeled in detail. Below this level, individual injectors cycles are not modeled and the injection phase length is a random value from a distribution. Additionally, to increase the calculation speed, the injector cycle modeling is active only when the collider requires beam.

6 Results on LHC operation modeling

6.1 Model validation

The model was validated using data from 2012 LHC operations. The choice was logical as at the start of the project this was best available operations data and A. Apollonio had used it to validate his model [18]. Following chapters present the model input data and the obtained results. Before this thesis, these results were presented in a summarized format in [1].

6.1.1 Simulation assumptions

We use the following assumptions to simulate the LHC luminosity production during 2012. The duration of different cycle phases is shown in the table 6.1. Fixed phase durations are used in the simulation and the length is an average of the measured phase durations in 2012.

Table 6.1: Time lengths assumed in the simulation for each successful phase in the operation cycle.

Phase	Idle	Injection	Ramp	Delivery	Ramp down
Length [min][a]	180[b]	52	49	576[c]	35

[a] The majority of the lengths are from [177].
[b] The phase length is an estimate to achieve the average turn around time of 5.5 h [178].
[c] Average duration of production phases ended by operators for luminosity optimization.

Luminosity production is modeled with the equation 4.6. We used two sets of parameters to model intensity ramp up. Parameters are presented in the table 6.2. The low-intensity beam is used for the first 40 days of the proton physics period.

Table 6.2: Beam parameters for simulation.

Beam[a]	Peak luminosity	Luminosity lifetime
Low-intensity	$2.5 * 10^{33}$ cm^{-2}s^{-1}	10.5 h
Nominal intensity	$6.3 * 10^{33}$ cm^{-2}s^{-1}	10.0 h

[a] The parameters are averages for given periods and not related to specific filling schemes.

Exponential distribution (equation 3.5) was used for generating random failure times and repair lengths. Failure statistics that are shown in table 6.3 were used in the model. The available LHC hardware failure data [179] could not reliably specify the phase in which a failure occurred. These failures were thus treated as phase-independent.

Table 6.3: Beam parameters for simulation.

System[a]	MTTF [d]	MTTR [h]
Access System	7.3	1.5
Beam dump system	9.8	1.8
Beam instab. injection	4[b]	0
Beam instab. ramp	12[b]	0
Beam instab. production	30[b]	0
Beam interlock system	207	1.3
Collimators	6.8	1.4
Controls	9.9	1.2
Cryogenics	6.3	8.7
Experiments	6.8	1.4
Injector Complex	0.5[b]	1.5[b]
Injection Magnets	5.2	2.3
Miscellaneous	5.1	2.5
Power Converters	3.8	1.9
Quench protection system	3.7	1.9
Radio frequency	3.0	1.7
Software Interlocks	207	0
Technical Services	6.3	3.8
Vacuum System	10.7	3.8

[a] The LHC hardware failure statistics are from CERN's failure database records for the 2012 proton physics periods [179].
[b] Estimates

Most of the failure and repair rates, in the table 6.3, are calculated directly from the data. For MTTF, this was done by calculating the number of failures in a specific category during the physics production period in 2012 and dividing the period with the number of observed failures. This procedure leads to a minor simplification. Technically in the model, each system was an individual failure node and it cannot fail again once it is in the fault state. However, a node represents a system consisting of multiple pieces of equipment which can fail at the same time and this is not now taken into account. The MTTR values were calculated by dividing the total downtime with the number of failures.

The occurrence rates for beam instabilities were however estimated as they were not reported as failures in 2012. The raw estimate was based on the fact that during 2012, beam related events caused 73 beam dumps at top energy [18]. This statistic was the basis for the instability rate in the production phase. The instabilities were assumed to be more common during injection and ramp phases leading to estimated higher occurrence rates. Also, the injector complex availability was estimated to be about 90% for LHC beams from SPS. In the statistics from [180] the injector availability for the LHC, during 2012, was more than 90%. However, these statistics do not take into account if the LHC requires the beam. For comparison in 2017, the SPS achieved about 90% availability [147]. This 90% availability value considers only the hardware availability of the injector chain. The insufficient beam quality caused issues both in run 1 and run 2 (see section 4.3.2). For this validation, this is taken into account by a longer duration of the injection phase.

6.1.2 Results

The results of the simulation using 2012 LHC input data allow validating the approach. The distribution of the simulated integrated luminosity productions resulting from 1000 simulations is shown in the figure 6.1. The mean integrated luminosity is 23.38 fb^{-1}. The measured production of 23.27 fb^{-1} [181] is well within the set 10% accuracy goal.

Aleatory uncertainty can be tested by fitting a normal distribution to the simulation result distribution. Confidence intervals (95%) can be derived for the mean value μ with the equation:

$$x - 1.96\frac{s}{\sqrt{N}} < \mu < x + 1.96\frac{s}{\sqrt{N}}, \tag{6.1}$$

where x is the simulation mean, s its standard deviation, and N is the number of simulation iterations. There is 95% confidence that the true value of the parameter is in the confidence interval. This gives $\pm$ 0.05 fb^{-1} confidence limits for the simulated result. That shows that 1000 simulations yield sufficient computational accuracy.

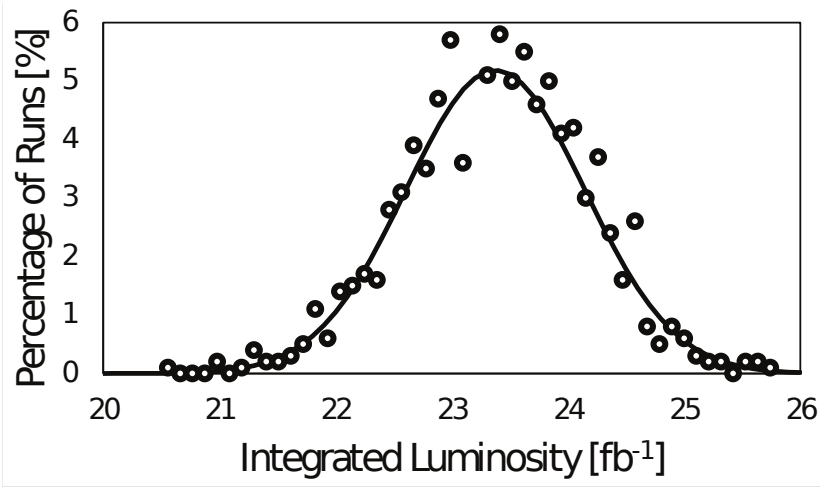

Figure 6.1: Result distribution of 1000 model iterations with a Gaussian fit that uses the data from the simulations. The mean value is 23.38 fb^{-1} and the sample standard deviation is 0.77 fb^{-1}.

Figure 6.2 compares the measured and the modeled luminosity production for one model simulation. At the start of the luminosity production period, the effect of the intensity ramp up is visible through a lower production rate compared to the rest of the year.

Simplifications introduced in the model implementation explain differences between the measured and the simulated luminosity productions. For example, the time between technical stops is assumed to be constant in the model. Nevertheless, this does not affect the availability nor the luminosity production calculations.

6.2 Luminosity production prediction

When the results that validate the model where send to Physical Review Accelerators and Beams journal, referees asked if the method can predict future performance. To assess this, we simulated the 2015 luminosity production from 7.9.2015 to 3.11.2015 (58 days). During this period the

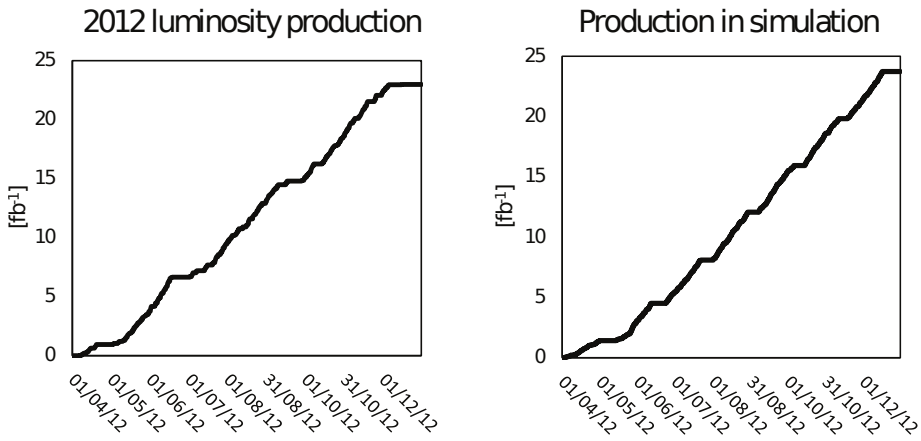

Figure 6.2: Actual LHC integrated luminosity production during 2012 compared to production in one model simulation

LHC produced 3.9 fb^{-1} of integrated luminosity [182]. The period was chosen as it was on-going during the writing of the paper [1] and it was also a relatively stable operation period during 2015. As, before this period LHC operated with different bunch spacing and after it LHC collided ions.

Our first goal was to create a credible estimate on LHC performance based on what information was available and published beforehand. 2015 was the first year of the second LHC run. The fault statistics from 2012 would have been the most relevant and the numbers presented in the table 6.3 would have been a good estimate. The operation energy increased for the second run to the 6.5 TeV. This meant that the lengths of the ramp and ramp down phases increase. An estimate was given in reference [177] that ramp lasts 67 minutes and ramp down 70 minutes. Other phase lengths were assumed to stay the same. Peak luminosity was assumed to increase to $8.1 * 10^{33}$ $cm^{-2}s^{-1}$ [183] and we assume an exponential decay with a lifetime of 10 h (similar to 2012). For a reference period of 58 days, these assumptions give 8.9 fb^{-1}, which is much more than the actual production value of 3.9 fb^{-1}. This difference is shown by scenario 1 in the figure 6.3.

Next task was to understand why the result was wrong and to test the effect of alternative assumptions. The unavailability of the LHC in the years 2012 and 2015 was similar. So failure and repair times were assumed to be correct. However, the beam parameters were different compared to the assumed values. Theoretical beam parameters for the period were: peak luminosity of $4.5 * 10^{33}$ $cm^{-2}s^{-1}$ and luminosity lifetime of 30 - 60 h [184]. These formed the scenario 2, which still resulted in significantly higher luminosity production. This was corrected when the peak luminosity value was changed to an average value calculated from operation statistics ($2.9 * 10^{33}$ $cm^{-2}s^{-1}$) [182]. This value was used in scenario 3 that is quite close to the actual production. Lastly, the lengths of the cycle phases were analyzed. Report [140] indicated that the average length of the injection phase was 90 minutes. This time is significantly longer than in 2012, and in scenario 4, this value and the beam parameters from the data were used which lead to achieving the correct production value.

Some thought can be given to understand why the original prediction and the actual production were so different. The year 2015 followed a two-year shutdown that caused uncertainties and challenges for operations and made performance predictions difficult. Key reasons why the peak

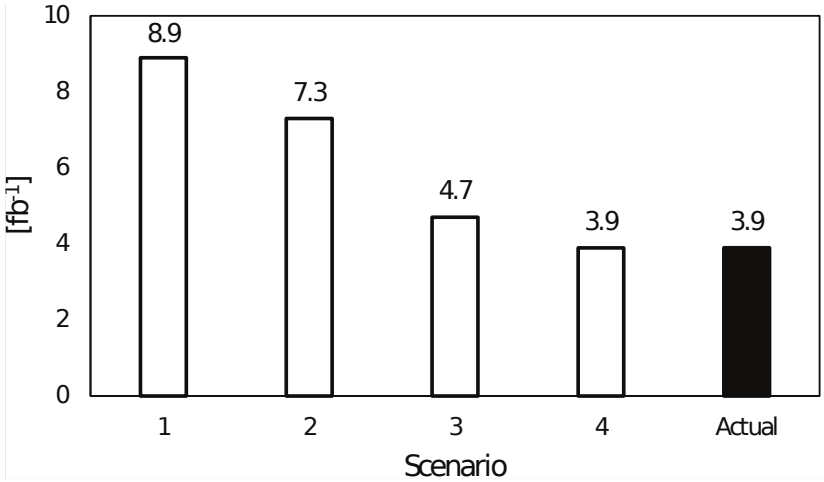

Figure 6.3: Predictions for integrated luminosity production in the LHC during 2015 reference period compared to actual value.

luminosity was overestimated were: i) delayed start of the proton-physics and a special physics run during the studied period that limited the luminosity production [128]; ii) technical difficulties encountered in August 2015 related to radiation effects [185]; and iii) performance after the long shutdown was limited by deconditioning effects (e.g., electron cloud [131] and interaction of the beam with dust particles (Unidentified Falling Objects (UFO)) [130]). The main reasons for the longer injections were injection intensity limitations due to electron cloud and insufficient beam quality [140].

6.3 Sensitivity analysis

The model validation presented in this thesis does not take into account epistemic uncertainty. It assumes that the mean values observed during the operations are correct. However, the observed values form a sample, and its mean value can be tested with equation 6.1 if the number of the observed equipment is known. Also, one can think that input parameters have systematic errors or other uncertainties. For example, operating conditions change between years and runs. So, it is important to be able to extrapolate the behavior of failure causes and assess their impact on operations. These can be handled with a sensitivity analysis where a range is given for the input values.

We created an example scenario to show how our model can be used to assess the sensitivity of luminosity production on the input parameters. In 2015 interaction of the beam with dust particles (Unidentified Falling Objects (UFO) [130]) caused 21 beam dumps from which 3 caused a quench in the superconducting magnets [186]. The failure mode where an UFO quenched a magnet was first observed in 2015 when the beam energy was increased to 6.5 TeV. The UFO occurrence rate and consequences correlate with the beam energy. Here it is assumed that, each UFO has a probability of producing beam losses that lead to a quench.

We performed a what-if scenario on the 2012 data, introducing quenches caused by UFOs. The amount of luminosity production is calculated with different UFO dump rates and quench probabilities. We assumed that a quench causes nine hours of downtime.

The results in the figure 6.4 show that the integrated luminosity production depends strongly on the UFO dump rate whereas it is less sensitive to the quench probability. Each dump causes the cycle to be re-executed. The downtime increases with higher quench probability, as the cryogenic conditions need to be recovered after a quench.

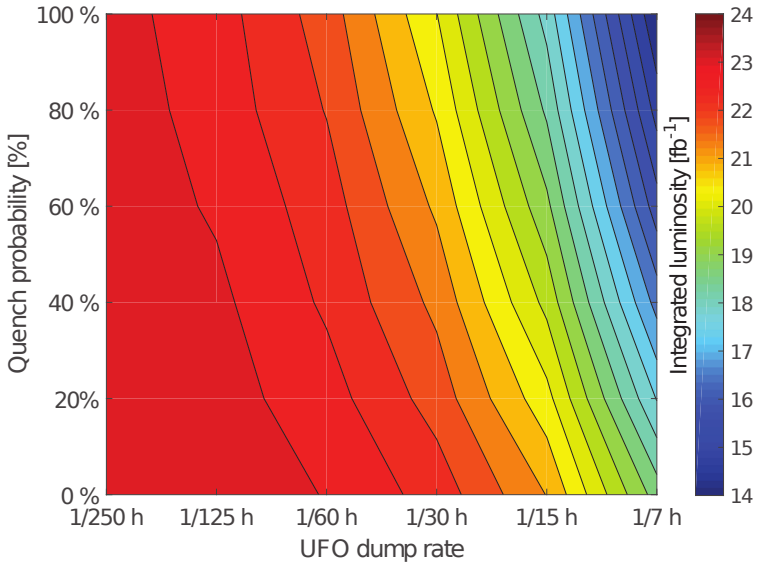

Figure 6.4: Sensitivity analysis of luminosity production to UFO dump rate and quench probability, based on LHC 2012 operational data. Quench probability is the probability of a quench to occur in each UFO dump.

Figure 6.4 is formed by simulating 36 different value pairs. In each point, 1000 rounds are simulated and the mean value is presented. This makes creating sensitivity analyses calculation heavy. The choice of points with the constant intervals is not generally encouraged as it might create a systematic error. Alternative methods such as Latin hypercube or completely random selection of points are usually preferred [187].

This sensitivity analysis confirms the decision taken in 2015 and 2016 when the beam loss monitor thresholds were increased to limit the UFO dump rate while accepting higher beam losses [130]. Our results indicate that decreasing the dump rates can increase the luminosity production, even if the downtime associated with a dump would be longer.

7 Availability goals for the FCC-hh

7.1 High-level availability goals

The preliminary baseline assumption for the FCC-hh lifetime is 25 years, consisting of five 5-year long periods [17] that have consist of shutdown and operation time. The exploitation is divided into two phases: baseline parameters will be used in the first ten years and ultimate parameters during the last 15 years of operations. Integrated luminosity production goals for proton physics are 2.5 ab^{-1} with baseline parameters and 15 ab^{-1} with ultimate parameters. With these assumptions, the production goal for baseline parameter run is 1.25 ab^{-1} and 5 ab^{-1} for a run with ultimate parameters.

The first run starts with an extensive commissioning period, which could overlap with the end of the construction since some of the systems can already be commissioned while the collider is still being completed. Later runs start with a 1.5-year shutdown. Within a run, the time allocated for proton physics is assumed to be 2.5 years, while three months are foreseen for ion physics. Combining proton and ion physics yields a total of 165 months of physics over the FCC lifetime. This leaves per run nine months for machine commissioning, machine studies, and scheduled stops. The number and strategy for managing stops can be reviewed, based on the needs imposed by individual system designs, but should not exceed the allocated time.

The challenge related to short maintenance, commissioning and machine development periods can be appreciated when comparing it to the planned operational schedule for the High-Luminosity Large Hadron Collider (HL-LHC). During a standard operational year, only about 6.5 months are foreseen for proton, ion and special physics [58]. Over a 25 year lifetime including shutdowns, such a schedule would result in 114 months for physics. In a 3-year run, it reserves about 17 months for machine commissioning, machine studies, and scheduled stops.

7.1.1 Examples of possible operational schedules

Several alternative ways can be conceived of how to arrange the stops during a run. These schedules should not be seen as definitive, but rather as potential options. Originally, the author presented these schedules in a report [3]. However, at the time of writing this these schedules might be obsolete as the report [188] states that FCC-hh would operate 160 days during an operational year.

Few simplifications were made in drawing up these schedules. The document does not differentiate between hardware and beam commissioning. It is assumed that straight after commissioning the machine is ready for operation with full beam intensity. In reality, the switch from commissioning to production is not binary; the intensity is typically ramped up gradually, as shown in the section 4.2. The ion runs are presented as 1 or 1.5 month long periods before each stop. However, it is currently unclear whether it is more optimal to arrange ion physics into a few long runs or several smaller periods [189].

Traditionally the LHC has had annual technical stops, scheduled at the end of the year. Figure 7.1 shows that with this arrangement the time reserved for the stop and the following recommissioning is only two months.

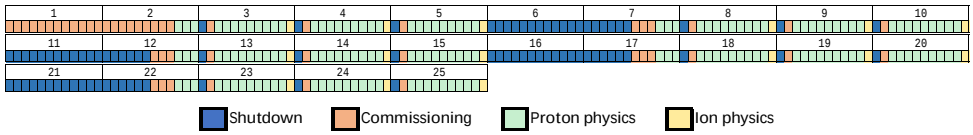

Figure 7.1: Schedule with annual technical stops.

If the LHC will be the final injector for FCC, the maintenance and recommission of the injector chain will take at least 3.5 months [3]. This can be accommodated if successive annual stops are combined into one longer stop. Figure 7.2 shows a schedule with a longer stop in the middle of each run. In this arrangement, the time reserved for the stop and recommissioning is five months. This version provides in total 170 months for physics operations.

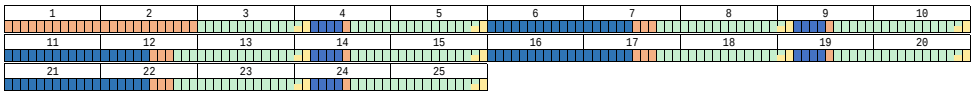

Figure 7.2: Schedule with one longer stop during each run.

If the preliminary requirement [17] for long shutdowns every five years is abandoned, a schedule with three runs can be imagined. Figure 7.2 shows a schedule with only two long shutdowns. During the first one, FCC is converted from phase 1 to operate with ultimate parameters. The second long shutdown is placed in the middle of phase 2. The cryogenic system is assumed to require maintenance every 2.5 years [190]. This schedule provides about 185-190 months for physics. The end of the schedule is not developed to detail; an additional stop might be required to operate through to the end of year 25. Such an additional stop, however, could also extend the FCC lifetime for another two years.

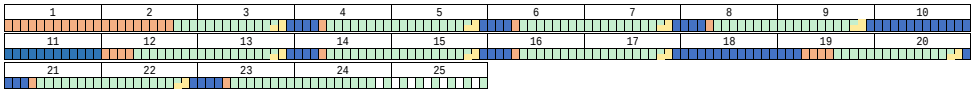

Figure 7.3: Schedule with only two long stops and shorter stops every other year. The final ion physics run could be at the end of the lifetime (not drawn).

After the report [3] was published, the author got further input on this topic. The four months long technical stops would not allow machine upgrades in the collider or in the injectors. Experience from LHC [191] and older accelerators [129] has shown that long shutdowns allow to upgrade and improve a machine which has led to significant gains in instantaneous luminosity production. In the LHC the collision energy has also increased over the years: the energy during the first run was 7 TeV, 13 TeV in the second run and the HL-LHC upgrade [58] is planned for the fourth run leading to a high increase in luminosity production. Also, increasing the collision energy to 14 TeV is being studied [192].

Another input relates to energy prices. In CERN's current energy contract, cost of the energy is higher during the winter months (October - March). Today the price difference is about 30-40% higher costs during these months [193]. This purchase price difference is caused by increasing energy demand during winter due to heating. A recent study [194] however, suggests that this might change if the climate change increases temperatures in Europe. Rising summer air temperatures would cause people to use more air conditioning and this will lead to high energy consumption during summers. If this is a case in the future or at the site location, the annual shutdown could be during the summer[1].

7.1.2 Maintenance during technical stops

The report [3] focused on the cryogenics system maintenance to study required maintenance interval and maintenance break length, as this system is the maintenance time driver for a superconducting accelerator. The presented scenarios assume that maintenance of the cooling water system is scheduled in parallel to cryogenics maintenance. The effect on the total intervention time from adding backup cooling towers and cryogenics interconnection boxes is studied. The study is divided in three cases: (1) with backup cooling towers and interconnection boxes; (2) with backup cooling towers but no interconnection boxes; and (3) no backup cooling towers and no interconnection boxes.

The backup cooling towers would provide cooling to the cryogenic system while the primary cooling facilities are being cleaned. The cryogenics interconnection boxes would be located in points B, L, F, and H in figure 7.4. These boxes allow distributing the cryogenics to the adjacent sectors when the primary cryogenic plant in point A or G is unavailable. Additional basic assumptions for all scenarios are: (i) 10 cryogenics plants with a cooling tower in each of the six technical points; (ii) cryogenics plant maintenance is carried out by 5 individual teams and takes 2 weeks; and (iii) cooling tower maintenance is carried out by 3 individual teams and takes 3 weeks.

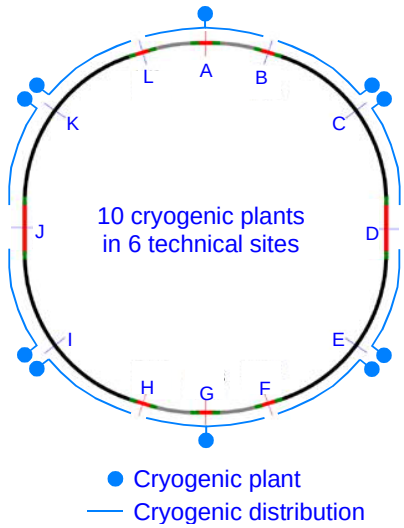

Figure 7.4: Layout of the FCC-hh cryogenics system [195].

[1] Author suspects that this is a case already in some accelerator, but could not find a reference that would directly state this.

Table 7.1 shows the maintenance time in the three different cases; the associated Gantt charts are shown in the Appendix A. In case 1, the downtime is caused solely by the cryogenic plant maintenance and magnet powering tests which take seven weeks. In case 2, due to the lack of interconnection boxes, liquid helium needs to be emptied during the maintenance from sectors L-A-B and F-G-H. This, combined with re-cooling and resulting electronic integrity tests, lead to 11 weeks of maintenance time. In case 3, without backup cooling towers, all sectors need to be warmed up during maintenance, which results in 15 weeks of downtime.

Table 7.1: Study cases and minimum time for cryogenic maintenance

Case	Interconnection boxes	Backup cooling	Maintenance time
1	Yes	Yes	7 weeks
2	No	Yes	11 weeks
3	No	No	15 weeks

The key result is that with the backup cooling towers the cooling tower maintenance does not stop the cryogenics system operations. So, the cases 1 and 2 give more freedom to schedule the cooling tower and associated system maintenance. This depends on the level of redundancy provided by the backup cooling towers. It is possible to design a system where both cooling towers would be required to provide adequate cooling in an operational scenario.

7.2 Operational cycle

Operational cycles consist of a collision phase, or 'physics production phase', and a turnaround phase, required to reach collisions. The minimum turnaround is 1.8 h (Table 7.2). Table 7.2 shows the breakdown of the cycle phases, including the LHC experience. The recorded LHC ramp-down time is artificially longer, as it is dominated by failure recovery times, which are mostly cleared during this phase.

Table 7.2: Technical performance targets for FCC-hh turnaround cycle [7], theoretical limit for LHC turnaround [10] and observed minimum and mean turnaround times in 2017 [133] in minutes.

Phase	FCC target	LHC theoretical	LHC min 2017	LHC mean 2017
Setup	10	10	-	-
Injection	40	16[a]	28.0	77.1
Prepare ramp	5	-	2.3	5.0
Ramp-Squeeze-Flat top	20+ 5+3	20	20.2+13.4+2.8	20.5+18.1+4.5
Adjust	5	-	3.3	7.9
Ramp down	20	20	36	153.2[b]
Total	108 (1.8 h)	$\approx$ 70 (1.2 h)	106.0 (1.8 h)	286.3 (4.8 h)

[a] This assumes 20 seconds-long SPS cycles.
[b] The ramp down phase includes the recovery time from failures.

Given the fact that the optimum collision time will be only 3.5 hours for ultimate parameters, the production efficiency will be highly dependent on the average turnaround time. This implies the increasing importance of the injection phase and of the injectors' performance. Based on LHC

experience, the current injection process could be improved by (i) adding beam diagnostics in the injectors to help identify beam quality issues as close as possible to the source, (ii) having fast diagnostics to understanding the cause of rejected injections, and (iii) improved synchronization and coordination with the injectors [140].

The FCC turnaround time is presented in the report [7] that describes the assumptions that were used to derive it. The filling of the FCC is the most interesting phase. Further identified challenges in achieving the turnaround time are the high requirements for power converters set by 20 minutes ramp time [196] and the transient heat loads to the cryogenic system [197].

During the study three injector options have been considered: i) using LHC as injector and ii) superconducting machine in the SPS tunnel, scSPS [198] and normal conducting machine in the FCC tunnel. However, lately, concerns have been raised that the energy difference between scSPS and FCC would be too high to use scSPS as the final injector [199]. The filling process with different options is described in the presentation [200]. If the LHC is used as an injector, the injection energy is 3.3 TeV. The energy of a single injection from the LHC to FCC is limited due to machine protection reasons, as it is assumed that one should be able to deposit the energy safely if an injection kicker misfires. This leads a staggered transfer scheme where there are 33 individual injections from the LHC to FCC in rapid intervals.

FCC requires 4 LHC cycles to fill it. The first cycle is assumed to last 12.5 minutes and the later ones 10.5 minutes. The first cycle takes longer due to pilot bunch injection and measurement. The last ramp down of the fourth cycle does not need to be counted to the fill time, as at that point the FCC is already filled. In total, this results in 42 minutes fill time. Achieving a 10.5 minute LHC cycle requires five times increase in the LHC ramp speed [201] that requires 50 A/s ramp rate. For other injectors, cycle requirements are: SPS 32 cycles, cycle time 10.8 s and PS 300 cycles, cycle time 3.6 s.

If scSPS is used as an injector, the injection energy is 1.3 TeV. Again the injection energy is limited and there would be four individual injections per cycle [202]. FCC requires 33 scSPS cycles to fill it. The first cycle is assumed to last 3 minutes and the later ones 1 minute. The first cycle is again longer due to pilot bunch injection and measurement. In total, this results in 37 minutes fill time. The normal conducting machine in the FCC tunnel could fill the FCC in 29 minutes [200]. However, concerns have been raised that this option might interfere with the experiments or would require 15.5 kilometers of tunnel to bypass them [200].

7.3 Availability

This section presents calculation results that test the FCC baseline assumptions on availability and sensitivity analyses that study failure, recovery and injection success rates' effect on the collider availability for physics. Section 7.3.3 presents the current understanding on how system complexity scales from the LHC to the FCC. However, this information is not used in calculations nor to form a system availability budget, as system availability depends heavily on system design. Providing a credible allocation would require detailed studies of individual systems to forecast what availability could be achieved in the FCC.

7.3.1 Testing FCC baseline assumptions

FCC preliminary baseline [17] assumes that if LHC like 70% availability is reached, the physics goals are reached. To validate this, the following assumptions were used to set up a simulation. Failure and repair rates were based on LHC 2016 operations. Operational issues were assumed to be mitigated thanks to increased automation in the future. Also, beam instabilities were excluded.

Injector chain availability was assumed to be 90% and the injection success rate was assumed to be 50%. This conservative estimate on success rates was based on source [138]. More recent results [140] show that the portion of the rejected injections is lowered to 20%. Cycle phase times from 2016 [203] were used as inputs.

Luminosity production was modeled with luminosity production curves shown in section 4.3.3. The key difference between luminosity production with baseline and ultimate parameters is the significantly higher instantaneous luminosity with the ultimate parameters. This leads to about 3.5 hours optimal collision time before the beam is dumped. Integrated luminosity functions were not available in an analytical form. Thus the simulation used lists that had the integrated luminosity value with 30 seconds interval that were provided by X. Buffat. The intensity ramp-up (see section 4.2) was assumed to reduce the total luminosity production with 10%.

Figures 7.5 and 7.6 show the distribution of the amount of integrated luminosity produced within 900 days period. This time corresponded to the schedule shown in figure 7.1 where life-cycle of the FCC contained five 1.5 year long shutdown periods. The results show that with these assumptions FCC could exceed the production goals. However, this assumes that the whole 900 days period could be operated without major failures.

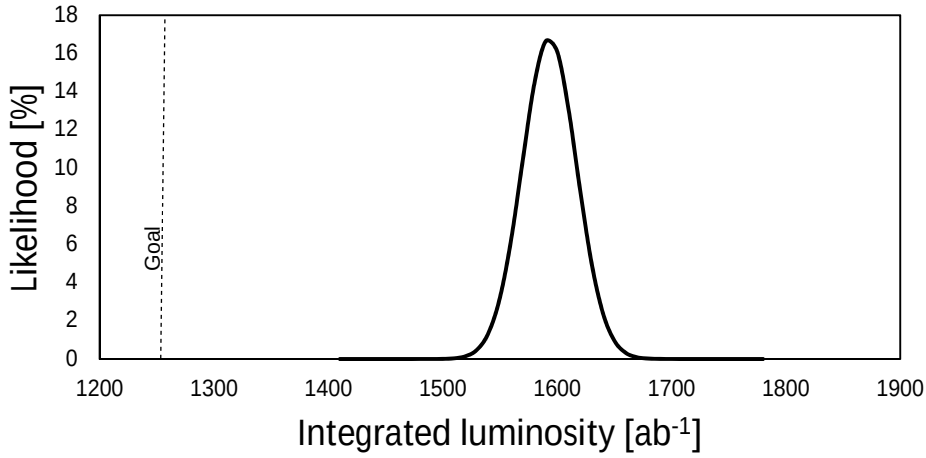

Figure 7.5: Result distribution of integrated luminosity production in 100 simulations with nominal FCC beam parameters compared to 1.25 ab^{-1} production goal.

Further analyses were conducted to study how operations would look like with injector failure rates that reflect actual LHC operations. These numbers were further used to estimate failure rates for scSPS. If the LHC were the final injector for the FCC, it would not have experiments, the beam energy would be limited to 3.3 TeV, and the bunch intensity would be less $1.0 * 10^{11}$ which is less than half of the HL-LHC bunch intensity $2.2 * 10^{11}$. Due to this, failures related to the experiments and high energy beam were removed. Failures related to the high energy beam include UFO related beam losses [130] and high stored energy in the magnets include magnet training quenches [150]. If the scSPS is the final injector, it will have superconducting NbTi magnets, quench protection system, new power converters and estimated three cryogenics plants.

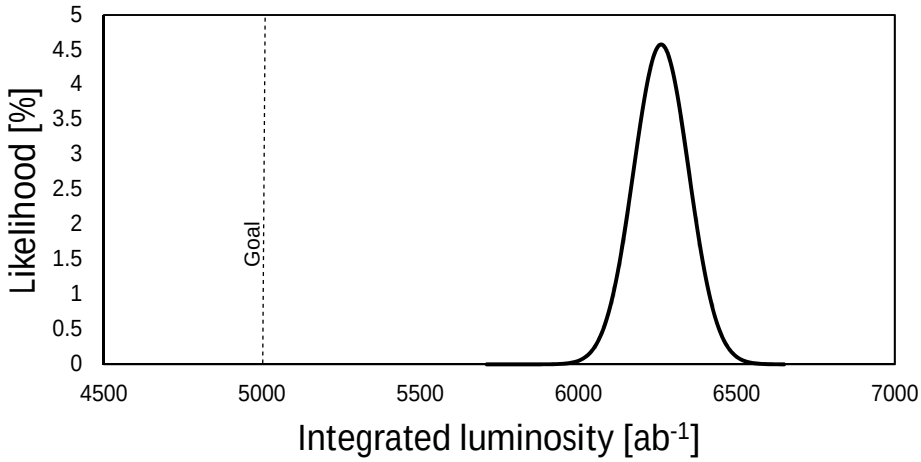

Figure 7.6: Result distribution of integrated luminosity production in 100 simulations with nominal FCC beam parameters compared to 5 ab^{-1} production goal.

Figure 7.7 shows the operations time allocation with ultimate parameters with the two injector options compared to the earlier "Goal" scenario presented by the figure 7.6. In the scenario with the LHC as the final injector FCC would on average wait three times longer for the injector chain to be ready than in the other scenarios. Notably, with the assumptions, the scSPS would be better than both LHC scenarios. Main reasons for this are a less complex injector chain and slightly shorter fill time.

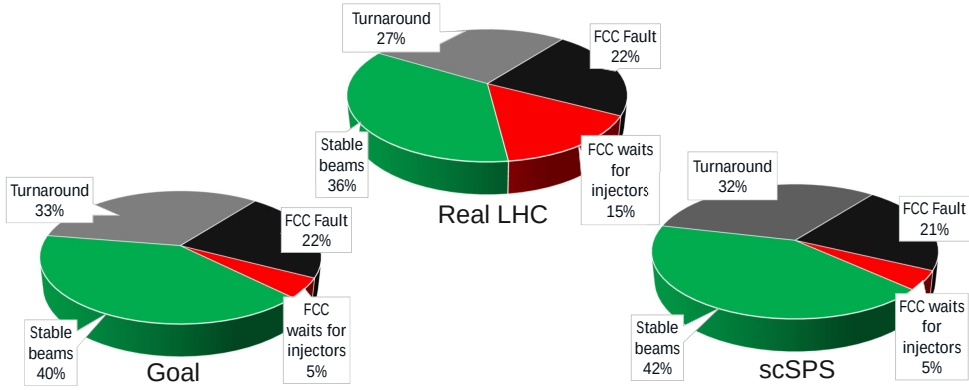

Figure 7.7: The "Goal" scenario shown in the figure 7.6 is compared to scenarios where the final injector has a) LHC 2016 failure rates and b) estimated scSPS failure rates. The results show that if the 2016 LHC failure rates are assumed FCC would be unavailable due to injectors failures on average three times longer than in the other scenarios.

To put the unavailability caused by the injector chain in perspective, the figure 7.8 shows a bar chart that compares how long time it would take on average to produce one ab^{-1} of integrated luminosity in the scenarios shown by the figure 7.7.

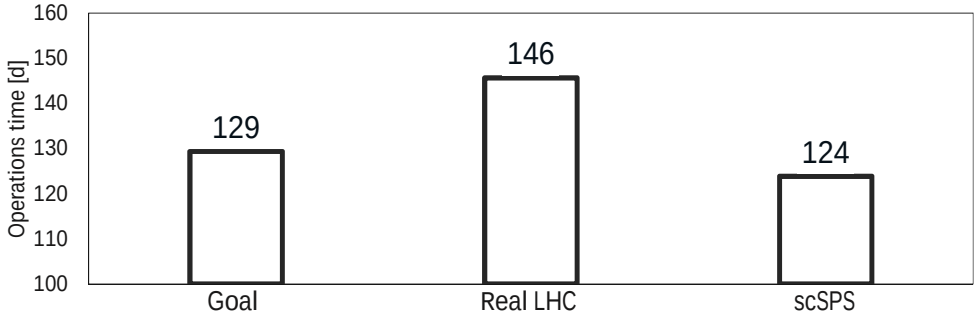

Figure 7.8: With present assumptions using LHC as an injector would lead to about 12-15% slower production rate compared to "goal" and scSPS scenarios.

7.3.2 Sensitivity analyses

As predicting FCC system availability is highly uncertain, it is more interesting to test how certain factors would affect the FCC availability and luminosity production. As presented for HL-LHC [18] the MTTF and MTTR values are one the most interesting parameters to play with. Figures 7.9 and 7.10 show the results for nominal and ultimate scenarios. For these plots, injector availability is 90% and injection success rate is 50%.

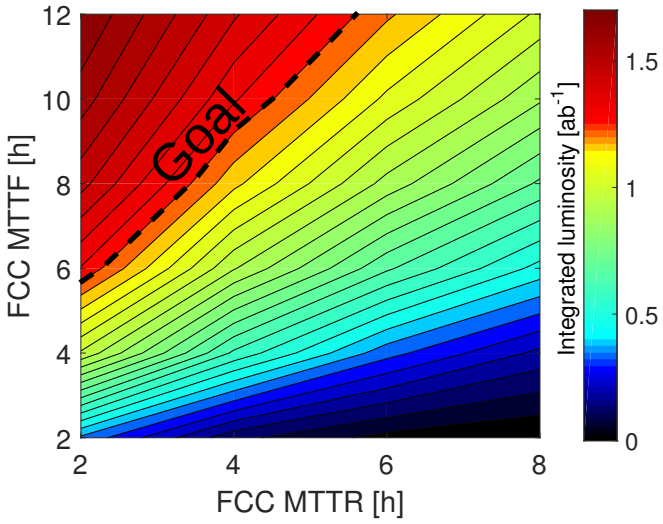

Figure 7.9: Integrated luminosity production's sensitivity to collider MTTR and MTTF with the baseline parameters.

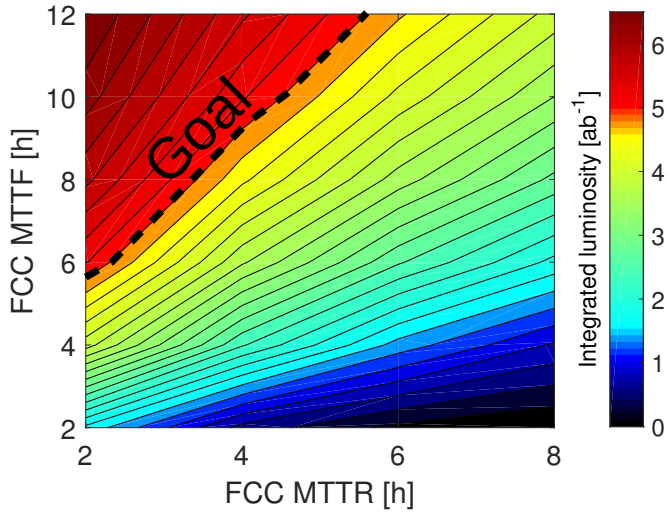

Figure 7.10: Integrated luminosity production's sensitivity to collider MTTR and MTTF with the ultimate parameters.

The injector chain availability is an uncertain factor due to different injector options. The effect of the injector availability on the FCC availability is tested in the figures 7.11 and 7.12. The MTTR is assumed to be 4 h.

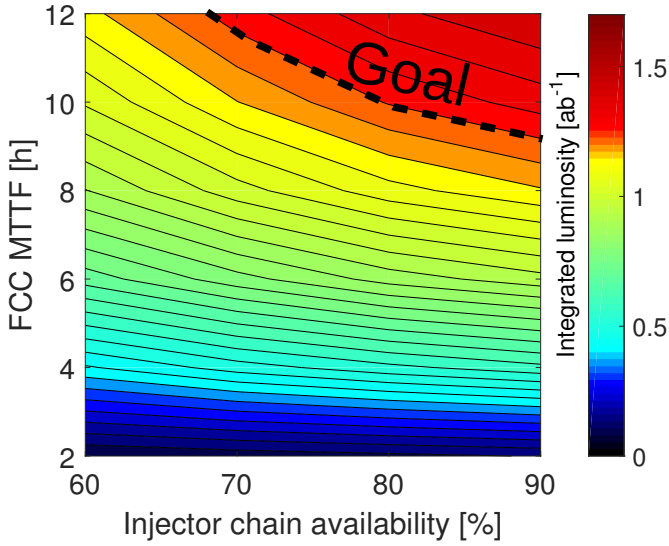

Figure 7.11: Integrated luminosity production's sensitivity to injector chain availability and collider MTTF with the baseline parameters.

The simulation model also allows changing the injection success rate. In practical terms, lower injection success rate equals to longer time spent in the injection. So, the results could also be seen to indicate how extending the injection time would affect the availability for physics. The

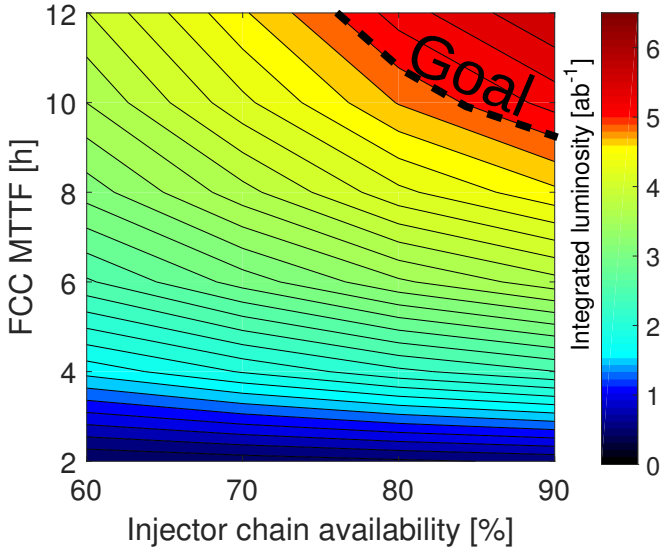

Figure 7.12: Integrated luminosity production's sensitivity to injector chain availability and to collider MTTF with the ultimate parameters.

MTTF is an interesting parameter pair for the injection success rate. High failure rate value leads to a situation where it is likely that the collider does not reach the collision phase. This likelihood depends on the failure rate and the turnaround time. Figures 7.13 and 7.14 show the results for nominal and ultimate scenarios.

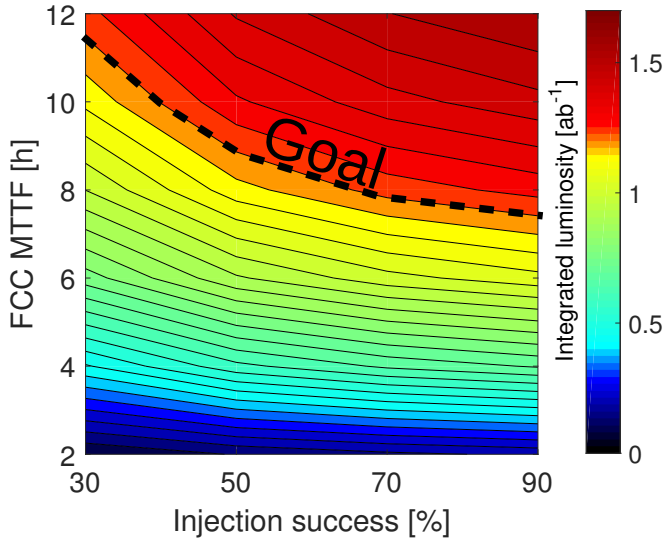

Figure 7.13: Integrated luminosity production's sensitivity to injection success rate and collider MTTF with the baseline parameters.

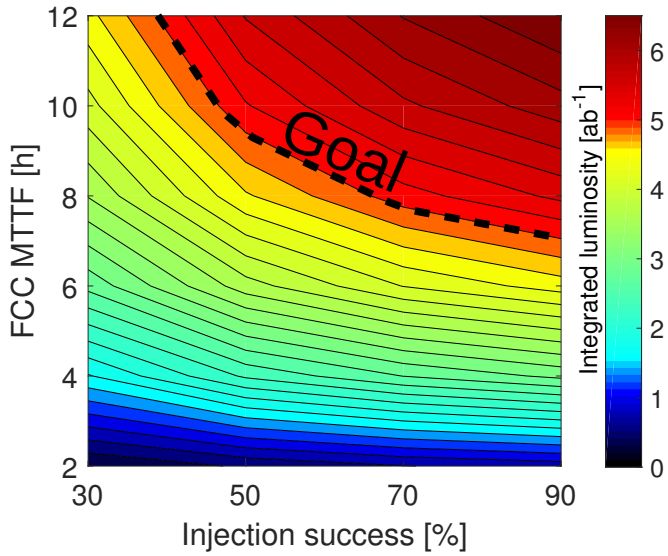

Figure 7.14: Integrated luminosity production's sensitivity to injection success rate and collider MTTF with the ultimate parameters.

The results show that the failure rate and repair rate affect the baseline and ultimate scenarios nearly identically, despite the different luminosity production functions. The likely reason is that with both parameters the integrated luminosity increases relatively linearly as a function of the fill time, as can be seen on the right side of the figure 4.16. This leads to a situation where the integrated luminosity production is practically linearly correlated to the time spend in collisions. However, the injector availability and injection success rates affect more the ultimate scenario. The reason for this is the shorter collision time that leads to more frequent fills with the ultimate parameters. Also, any failure that would be linked to the turnaround phase would affect more the scenario with ultimate parameters, due to more frequent turnarounds.

7.3.3 Scaling complexity

This section presents how system complexity will scale from the LHC to the FCC. Following considerations were made with A. Apollinio based on LHC experience and the present understanding of the FCC baseline. The list is ordered based on system's effect on the LHC availability.

Cryogenics system has inherently long recovery times as it needs to cool down to the operating temperature after failures. Management of the transient heat loads and the heating caused by electron cloud has been a challenge in the LHC. For the FCC the number of cryogenic plants will increase from 8 to 10 and the higher heat loads and longer distance between cooling plants will increase the plant requirements. However, new neon-helium compressors [195] might be more reliable than the currently used skew compressors thanks to oil-free active magnetic bearings and the fact that they do not need gear-boxes or shaft-seals .

Injector complex Today injector complexes availability for the LHC is over 90%. However, a relatively high number of injections is rejected due to beam quality or operational issues. For the FCC, both LHC and scSPS would be superconducting machines resulting in more complex injector chain. Also, in both options, there will be a challenge related to aging

legacy equipment within the injector complex. However, new, advanced beam diagnostics may improve help identifying factors that affect the beam quality. This can help to improve the quality and thus improve the injection success rates. FCC operations will likely be more automated. This can decrease the number of operator-related errors as automated control system performs more consistently repetitive tasks [204]. Operations will also benefit from inherited knowledge on LHC operations.

Quench protection system is complex and distributed. Today new quench protection techniques are under consideration and FCC QPS might use so-called Coupling-Loss Induced Quench (CLIQ) technology [205]. FCC might have over 100000 interlocks for the QPS [6], if the number is assumed to scale linearly from the LHC.

Power converters and magnet circuits LHC has eight powering circuits for main dipoles. Today, the radiation effects on electronics significantly contribute to the number of spurious beam aborts. FCC will have about 100 powering circuits due to machine protection requirements that limit the stored energy within a circuit.

Beam dump system LHC dump system consists of 15 extraction kickers and ten dilution kickers. Extraction kickers have two redundant generator branches each with ten switches. In the latest plan, FCC kicker system will have 150 kickers each with two redundant switches and about 80 dilution kickers [206]. So, the number of switches in the kicker magnets will not increase, but the number of dilution kickers will increase significantly.

Electrical network Unavailability mainly due to isolated, high impact events and electrical glitches resulting in spurious beam dumps. FCC will have three connections to the power grid (vs. 1 for the LHC). Infrastructure is expected to scale with the size of the machine.

Cooling and ventilation causes infrequent failures with long recovery time. Complexity will scale with the machine size. Constraints for periodic maintenance of cooling towers may lead to increased complexity.

Radio frequency LHC has 16 cavities in 4 cryogenic modules powered by 16 klystrons. FCC will have 40 cavities in 10 cryogenic modules powered by 40 klystrons.

Injection Systems LHC has four kickers and 20 m of septum magnets per beam. FCC will have 40 kicker magnets and 90 m of septum magnets per beam.

Beam Instrumentation LHC has 4000 beam loss monitors that have caused majority of the beam instrumentation related downtime [155]. For the FCC, complexity is expected to scale with the size of the machine.

Beam Losses and instabilities LHC has no major problems with beam instabilities. UFO occurrence rate strongly depends on machine conditioning. Large UFOs can lead to magnet quenches. For FCC, a desire is that there will be no major problems with UFO's generation mechanism understood and solved. This could be achieved, for example with clean room installation [207].

Vacuum system Vacuum system has not caused frequent failures in the LHC. For the FCC, complexity will scale with the size of the machine and a new design is required for the beam screen to cope with the synchrotron radiation [208].

The list shows that scaling depends on the increasing size and the energy of the collider. Several of these details changed during the project and will likely change again in the future. What makes

it difficult to allocate credible availability requirements for individual systems as it does not only depend on scaling complexity, but also design practices and developing technology. To fully understand what the credible availability goals should be for an individual system would require the knowledge on how the system was designed in the LHC and could its reliability be improved with existing technologies or with technologies developed in next two decades.

This requires a thorough knowledge of the system and technology field to be able to forecast what reliability and recovery times can be achieved in the FCC. Also, system maintenance intervals and use profile will likely change from what they currently are in the LHC. In total, this adds too many variables to be able to produce a credible system availability budget that is allocated to the system level.

7.3.4 Recommendations for FCC-hh

The recommendations on FCC availability are divided into three different levels 1) schedule, 2) operational cycle and 3) availability. The challenge in the operation schedule is to reduce the frequency of planned maintenance. Operations with annual maintenance are unlikely to reach luminosity production goals [3]. The best option is to reduce the frequency of maintenance. Reducing the length of the maintenance or commissioning will be challenging as the machine complexity will grow. It is also notable that the beam commissioning cannot be conducted in parallel in different machines in the complex. The upstream machines need to be commissioned before injecting the beam into the next accelerator within a complex.

The key driver for maintenance time and frequency is the cryogenic system and its rotating elements. Further studies should be conducted to understand if the useful life of this system can be increased. An analysis also shows that increasing the maintenance interval can decrease the operational costs as the cost of a cryogenic system overhaul is assumed to be 3% of the system's capital expenditures [209]. Savings can be achieved if these overhauls need to be conducted less often.

The time required for cryogenics maintenance can be decreased if the magnets that system cools can stay in cold temperatures during the maintenance. This can be achieved if other cryogenics plants can provide an alternative cooling while the primary source is under maintenance. In addition, studies should be conducted to understand what maintenance activities require cryogenics system to be warmed. For example, in the LHC the cooling towers are cleaned annually to prevent a threat of water born bacteria [210]. In this case, backup cooling towers could provide cooling during the cleaning. There are also alternative means to prevent the bacteria thread, such as chemical treatment or cooling the water to a temperature where the bacteria cannot grow [211].

Although the initial study focused on the maintenance of the cryogenic system, similar studies should be conducted in the other systems to understand the system maintenance requirements. This is a crucial ingredient for forming an operational scenario which is essential for forming system design requirements. However, based on already available information, less frequent maintenance stops will require all accelerator systems and experiments to sustain longer periods without maintenance. The maintenance of the injectors needs to fit the schedule such that the injector chain is operational when FCC recommissioning starts. Additional work is required to understand if and how this can be accomplished with the current injector complex and the LHC and or with the scSPS [198] as final FCC injector. In both options, the cost of the operations and the maintenance should be one of the deciding factors.

For the operational cycle, the key question is the length of turnaround time. The turnaround will become much more important with the ultimate parameters where an optimal fill last just 3.5 h, which is short compared to over ten hour fills in the LHC. In the turnaround, the most

complex operation is the filling of the collider. In the LHC this has been the operational phase where the studies have found the most inefficiencies [140]. Some of these inefficiencies are due to communication between different accelerator operations groups. For example, the beam is not prepared to be injected when it is needed. More automated operations will likely handle some of this. Further studies should still be conducted to find ways to improve operational performance.

Concerns have also been raised for cryogenics systems ability to handle the dynamic heat loads [212]. A study should be conducted to understand if the best solution is to increase the cryogenics operation margin or to relax the requirements that are linked to the heat load, such as ramp times, beam energy or intensity. Furthermore, the ramp and ramp down cause fluctuations in energy consumption. For example, during a ramp, the consumption can increase with 100 - 280 MW [196]. There is a risk this high variability in the consumptions affect the energy price, as customers with predictable and relatively flat demand profiles are cheaper to serve than customers with peaky and/or unpredictable consumption patterns [213]. Storing the energy between ramp down and ramp can smooth the energy consumption [196].

For achieving high performance, the key recommendation is to acknowledge that the availability will be an issue that should be addressed from the early design stages of the project. Many guidelines exist on how to design reliable systems. For example, reliability centered maintenance approach [102] encourages on use of condition-based maintenance tasks to focus the maintenance activities. Implementing this will require designing systems to be inspectable. For a collider located in a 100 km long tunnel, this will depend on remote monitoring. In the best case monitoring can detect early symptoms of a fault which can be used for preparing a planned maintenance intervention. An alternative would be to let the system fail and leading to an unplanned corrective maintenance intervention.

Another way to increase the availability is to design fault-tolerant systems with built-in redundancy. This was used for example in the RF system of the PS Booster where system consists of 36 RF cells of which 30 needs to be operational for the system to work [96]. The plan with the redundancy is that a single failure does not stop the operations. The redundancy can be restored with deferred corrective maintenance at an occasion that causes minimal interruption for the operations. In an ideal world, this maintenance could be performed online while the system is operating and without impact on its operations.

For this, the maintainable items need to be accessible during the operations. Due to the radiation, the tunnel will be inaccessible during operations. So, to achieve accessibility, the maintained items should be located on the surface. This idea has been studied for HL-LHC power converters where the energy would be transferred from the surface via superconducting cables [214]. Some items need to be still in the tunnel and ideas have been raised if robotics could be used for maintenance tasks [215]. This might also reduce logistics time as the robot could be on site. These technologies might not be ready today, but it is worth remembering the FCC is planned to be built 30 years in the future.

Locating systems on the surface has additional benefits as it limits the exposure to the radiation. This is important for electronics. The failures caused by single event upsets were a cause for concern in LHC operations during the run 1 [216]. If the equipment cannot be located to the surface, there also exists an idea to locate systems to separate alcoves in the tunnel to limit the exposure [217].

Another new idea that has been proposed to improve the availability is a system to reduce power voltage dips. Reference [218] lists potential solutions on how to implement a voltage dip mitigation system. Slight fluctuations in voltage cause 20 - 30 beam dumps in the LHC that lead to a loss in integrated luminosity production due to the turnaround time. In the FCC the power will be supplied through three points. Naively scaling this means that power dips will cause three times more problems. Thus, a power dip reduction system could significantly increase the FCC availability.

8 Studies inspired by the research on the FCC availability

Work for the FCC availability study spurred two follow-up studies: collaboration between CERN, Tampere University of Technology and Ramentor Oy to establish OpenMARS approach for system reliability and availability modeling; and Accelerator reliability information system study in a frame of ARIES EU project. This chapter introduces these studies that are connected to each other as shown by the figure 8.1. The reliability information system provides the input data for models and OpenMARS approach covers the model definition and analysis.

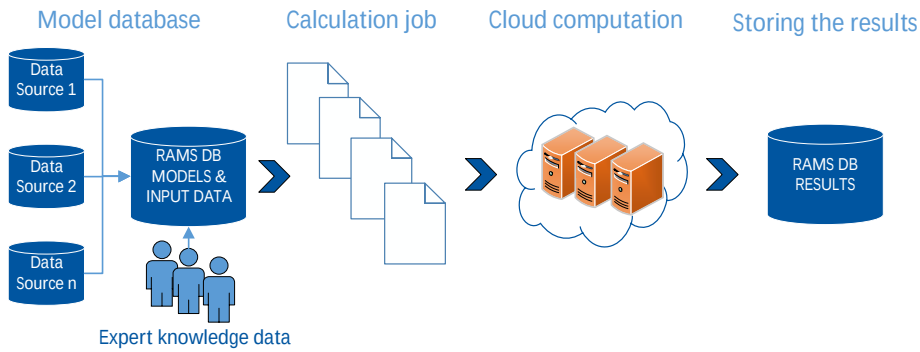

Figure 8.1: Vision where models and model data would be stored in a database that could be used for creating calculation jobs for a cloud or distributed computing cluster.

In the vision, the reliability models would be stored to the database in the OpenMARS format and new calculation jobs could be created for a cloud or distributed computing cluster. An interesting possibility arises if the results are systematically stored in a result database. This data could form Online analytical processing (OLAP) cube [219]. OLAP cube forms a dataset that stores its relations. This would allow a user to explore a pre-calculated the result data-space, instead of creating new calculation jobs to obtain new results.

The idea is similar to providing pre-aggregated data in databases [219] such as sums, row counts, and calculated averages. This can save time if the results were calculated automatically before a user needs them. For example, section 7.3.2 now shows the sensitivity of luminosity production to injector availability and collider MTTF. An automated system could also independently calculate results for sensitivity to injector availability and injection success or collider MTTR just based on the fact that a user has used these parameters in other sensitivity analyses. Also, a user could

provide confidence bounds for the input parameters and the case would be studied with different parameter value combinations to study the effect of the epistemic uncertainty.

It is worth to understand that some aspects of this vision are limited compared to what is achieved in the field of reliability engineering in the industry. For example in the nuclear industry, already two decades ago risk monitoring was used for determining instantaneous risk based on the actual status of systems and components for daily risk assessments [220]. Achieving this would require linking online system status information to a fault tree. Also, sharing system reliability information started in USA nuclear industry in late 1970's [221] and Norway's oil industry 1980's with the OREDA project that now comprises eight oil and gas companies [105].

8.1 OpenMARS approach

The OpenMARS stands for an open modeling approach for availability and reliability of systems. It supports the most common risk assessment and operation modeling techniques. Uniquely OpenMARS allows combining and connecting models defined with different techniques. The idea is to ensure that a modeler has a high degree of freedom to accurately describe the modeled system without limitations imposed by an individual technique. OpenMARS models are defined with a tool independent tabular format. This format was chosen as this format is used by databases and spreadsheet software to present the data. Thus, storing the data in this format allows using these applications to store model data. The approach is presented in documents [2, 5], but a small overview is also given here. The chapter 8.1.1 gives an introduction to the basic features of the OpenMARS table format and the chapter 8.1.2 shows how the collider operations model can be implemented in the OpenMARS approach.

8.1.1 Table format for OpenMARS model definition

The most distinctive feature of the OpenMARS is that models are defined in a tabular format. The motivation for using tables as a format to store models came from the fact that spreadsheet software and databases present information in tables. This way models are relatively easy to store in these applications. The original idea was to define model elements, connections, and parameters in a table format. Thanks to the work of J.-P. Penttinen all classes and key class features are also defined in tables. OpenMARS also has tables for storing model meta-data, as each model needs to be identifiable in a collaborative environment [222].

8.1.1.1 Model meta-data

Model meta-data is defined with two tables. The first table is called the Model table. It contains columns GROUP PROPERTY and VALUE. Property column defines what information is provided in the VALUE column such as model name, date, author, author's email address, organization and organizational unit. Example of this is shown by the table 8.1. There are also fields to provide information on the purpose and scope of the model and to provide a link to further documentation. These properties are defined in different groups defined by the GROUP column. These groups include general information, legal information, access related information, and other groups.

Table 8.1: Example of information in the model property table.

GROUP	PROPERTY	VALUE
General	Name	Collider availability model
General	Author	Arto Niemi

The second meta-data table is used for tracking changes. It has column REV to indicate revision number, DATE for the date of the revision, AUTHOR, STATUS for the status of the model (e.g., in work) and a COMMENT column to give a summary of changes. Table 8.2 shows the concept.

Table 8.2: Concept revision table.

REV	DATE	AUTHOR	STATUS	CHANGES
0.01	xx:xx:xxxx	Author A	IN WORK	initial model
1.0	xx:xx:xxxx	Author A	RELEASED	finished model

8.1.1.2 Class definition

OpenMARS model consists of elements that have classes. A class defines what attributes that an element has. Each modeling technique has a catalog of available classes and the goal is that most cases should be covered with them. The first version of OpenMARS supports fault tree and Markov models, reliability block diagrams, Petri nets, and failure mode and effect analyses. Special features include already mentioned radio-listener concept for connecting models and tools to define mathematical functions. In addition, expert users can extend available classes to define new classes.

Classes are mainly defined in two specific tables, but further definitions can be provided in tables designed for defining the model and its parameters. The class definition table defines the class name and its inheritance. In object-oriented programming a class of an object has a parent and child classes can inherit features of the parent. For this purpose, the class definition table contains two columns: CLASS, that gives the name of the class and IS A, that defines the parent class.

Class attribute table defines attributes that describe the class and its behavior. The table contains columns: CLASS, the name of the class; TYPE, type of the defined attribute, such as a number or a text string; and ATTR the name of the defined attribute. For example, the node class in OpenMARS is the parent for fault tree's fault nodes, Markov model class, and Petri net places. Node class' attributes include initial state at the start of calculation and information for a potential user interface, such as title and background color. As these attributes are defined for the parent all the child classes also have them.

Class attribute table is further used for defining containers. The idea was that models coherence is established with folders that can contain different model elements. This initial thought was generalized so that classes can contain instances of specified classes. For example, a Markov model class contains states and transitions of a specific Markov model.

Example: Tables 8.3 and 8.4 show how examples discussed in this section are defined in OpenMARS tables. Table 8.3 shows how a Markov model is defined as a child of the node class and Table 8.4 shows how background color is defined for nodes and how a Markov model is defined as a container for states and transitions.

8.1.1.3 Model structure

OpenMARS model structure is defined with two tables. Element creation table describes the elements that a model consists of, and the element connection table describes their connections. The element creation table contains columns: CONT, element that contains the newly defined element, ELEMENT, the name of the new element and CLASS, class of the new element. The element connection table contains columns: SOURCE, the source element of the connection and TARGET, the target element of the connection.

Table 8.3: Markov class is defined as a child of Node class and StateMarkov as a child of State class. Node and State classes are children of Element class. In OpenMARS Element class is a fundamental class that means that it does not have a parent.

CLASS	IS A
Node	Element
State	Element
Transition	Element
Markov	Node
StateMarkov	State

Table 8.4: A background attribute is defined for Node class and for all classes that inherit it, such as Markov and Markov is defined as a container for Markov states and transitions with the syntax "*".

CLASS	TYPE	ATTR.
Node	Color	background
Markov	StateMarkov	*
Markov	Transition	*

The element creation table and connection table can be further used for defining predetermined elements for classes. For example, a fault tree node contains a Markov model with two states and transitions between them. This example is illustrated in the figure 8.2 and in the tables 8.5 and 8.6.

Figure 8.2: A fault node in the OpenMARS contains a Markov model with fault and normal states and transitions between them [2].

Table 8.5: Predefined elements of the fault class are here defined in the element creation table. Alternatively, the same definitions could be made in the class attribute table.

CONT	ELEMENT	CLASS
Fault	normal	State
Fault	fault	State
Fault	failure	Transition
Fault	restoration	Transition

Table 8.6: Connections within the fault class elements.

SOURCE	TARGET
Fault/normal	Fault/failure
Fault/failure	Fault/fault
Fault/fault	Fault/restoration
Fault/restoration	Fault/normal

8.1.1.4 Comment on element connection table

A design decision was made during OpenMARS development that an element is an object but a connection between elements is not. The OpenMARS development started with the fault tree technique and the other techniques were added later during the development. The fact that a connection is not an object complicates defining models with other techniques. If a new version of OpenMARS were to be developed, the author would define connections as objects, as it would allow defining parameters for them.

The reasons why connections were not defined as objects are that it practically works for fault trees. For example, connections for cryogenics fault tree in the figure 5.3 are shown in the table 8.7. This shows that defining connections only by their source and target elements is straightforward. This is good as the ability to define a model manually with the table format was one of the design goals. The concept is similar to the adjacency list [223] that is one of the simplest ways store hierarchical data in a relational database. This assured us to continue with the approach.

Table 8.7: Connection definitions for the fault tree shown in the figure 5.3.

SOURCE	TARGET
cryoTopOr	cryoNotAvail
priFault	cryoTopOr
secFault	cryoTopOr
planMaint	cryoTopOr

Complications arose when Markov models and Petri net where implemented. A Markov transition has a source, a target, and a transition distribution. Due to this, a Markov transition must be an object. Same is true for Petri net arcs that connect Petri net places to Petri net transitions. An arc can have a weight attribute [224] that determines how many tokens an arc takes or adds to a place when the associated transition is activated. This is shown in the image 8.3.

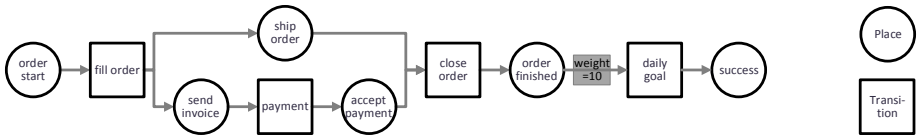

Figure 8.3: Petri net example from [2], reaching the daily goal requires ten handled orders. This is modeled by setting a weight in the transition from the order finished place. Once the daily goal transition activates, it removes ten tokens from the order finished place and adds one to the success place.

Defining models with the current implementation of these techniques is a bit burdensome. As an example, a connection between Markov states is defined with a Markov transition that is connected to the Markov places. Ideally, the Markov transition itself should be the only connection between the Markov places. To demonstrate this, connections in 8.6 could be defined only with two rows if a connection were an object, as shown by the table 8.8. Also, the weight of a Petri net arc is now defined with a separate stream element between Petri net places and transitions 8.9. If an OpenMARS connection were an object, it would allow defining other properties for the connections such as arrow color or thickness. Although, defining connections as objects seems obvious improvement at the moment, a proper study would be needed to determine this.

Table 8.8: Connections within the fault class elements if a connection were an object.

CONNECTION	SOURCE	TARGET
Fault/failure	Fault/normal	Fault/fault
Fault/restoration	Fault/fault	Fault/normal

Table 8.9: In the example presented by the figure 8.3 reaching the daily goal requires ten handled orders. In the table, the dailyGoalWeight is a separate object as the connection from orderFinished to dailyGoal is not an object and cannot have attributes.

SOURCE	TARGET
orderFinished	dailyGoalWeight
dailyGoalWeight	dailyGoal
dailyGoal	success

8.1.1.5 Model attributes

A model needs attributes that define for example failure and repair rates, node colors, the initial state of the Markov model, or the initial number of tokens in Petri net places. The attributes are defined with the attribute value table. It contains columns: OBJECT that is the path to the object that has the attribute, ATTR. is the name of the attribute that is being defined, and VALUE is the defined attribute value. Table 8.10 shows an example on defining a Weibull distribution for a fault node.

The most interesting features of the OpenMARS allow defining attribute values that are dependent on some external condition like the active phase of the collider or connecting different models with radio emitters and listeners. These are further elaborated in the section 8.1.2 that show how OpenMARS is used for defining the collider operations model.

Table 8.10: Definition of a Weibull distribution for a faultNode object.

OBJECT	ATTR	VALUE
faultNode	failure	TrWeibull
faultNode/failure	shape	1.2
faultNode/failure	scale	1000 h

8.1.1.6 Shortcuts for the definition of vast models

The idea to use table format as a way to define models came from a parallel study of Proton Synchrotron Boosters (PSB) RF-system [96]. PSB consists of four rings that each have 36 accelerating cavities. Figure 8.4 shows the vastness of the PSB RF-system fault model that results from the repetitive structures. This large number of components made the model definition burdensome in a graphical user interface. This problem led us to think about ways to allow users to define multiple elements or connections within a single row.

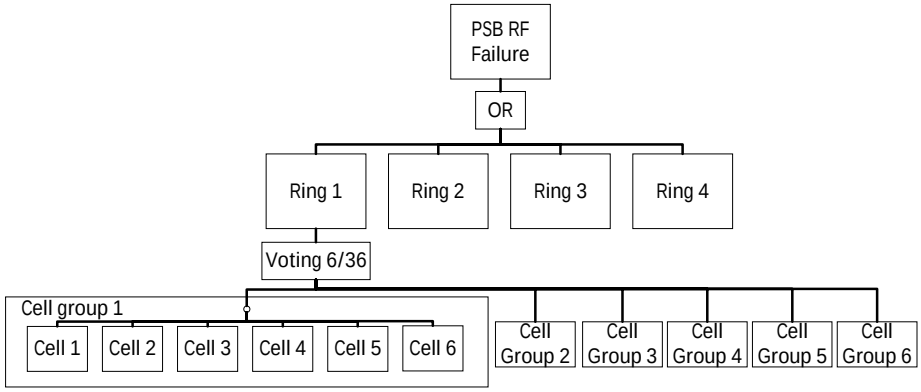

Figure 8.4: Fault tree of the PSB RF-system. This figure shows the fault tree structure only for the first ring's first cell group. The other cell groups and rings have a similar structure. In the full model that is presented in [96], each cell and cell group has internal elements that are not shown here.

The following example shows how the model definition and previously mentioned short-cuts work. The table 8.11 shows how the PSB consists of four rings that each have six cell groups each containing six cells. The table uses folder concept, array definition and reference definition to structure and limit the required rows. For example, four ring folders are defined in row 3 with the array definition. Row 5 uses reference definition to define a ringFault element to each ring folder. In comparison, if an element were to be defined to a specific folder, this would require input *"/ring[x]"*, where the *x* is the number of the folder. Without these shortcuts, over 100 rows would be needed to define the elements.

Table 8.11: Start of the element definition table for PSB RF model. Shortcuts are used for defining repetitive structures.

#	CONT	ELEMENT	CLASS
1		psb_RF_Fault	FaultTree
2	/psb_RF_Fault	psb_RF_OR	OR
3	/psb_RF_Fault	ring[1-4]	Folder
4	ring	ringFault	Fault
5	ring	ringVote	Vote
6	ring	cellGroup[1-6]	Folder
7	cellGroup	cell[1-6]	Folder
8	cell	cellFault	Fault

The table 8.12 shows the connection definition for the example. Again as the source and target elements are not specified connections are made for all elements. Defining this rule was unexpectedly difficult. There was an open question: Why the definition like the one shown in row 3 does not connect all ring vote elements to all ringFaults? This would lead to a situation where a failure of a single RF cell would affect all rings which would be wrong. This problem is shown in the picture 8.5. Finally, a rule was invented that with the reference definition a connection is made only between elements in a common container object. The effects of this rule can be seen, for example, in the second row. In the definition: ring/ringFault, the notation of the container ring is needed to define reference that is at the same level as the target element psb_RF_OR.

Table 8.12: Start of the element connection table for PSB RF model. Reference definitions are used for defining repetitive connections.

#	SOURCE	TARGET
1	psb_RF_Fault/psb_RF_OR	psb_RF_Fault
2	ring/ringFault	psb_RF_OR
3	ringVote	ringFault
4	cellGroup/cell/cellFault	ringVote

The model parameters are defined in the attribute value table. This limited example in the table 8.13 defines MTTF and MTTR values for the RF cell failure and the required number of simultaneous failures for the voting gate. The MTTF and MTTR values for the RF cell failures needs to be set only once if the cavities are assumed similar. Here MTTF value for all cells is set in row one, but this is overwritten in the second line for the first RF cell in the first ring. The value 7 for a voting gate means that at least 7 out of n simultaneous faults are required to cause a ring failure.

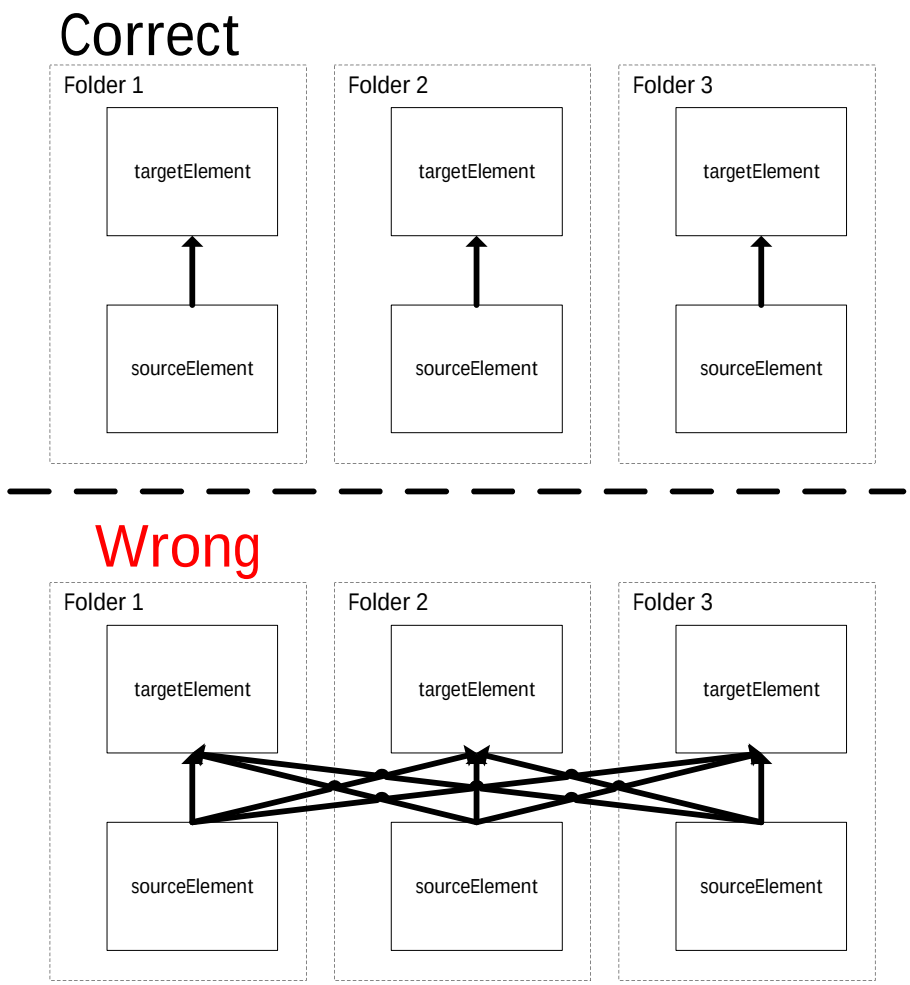

Figure 8.5: Reference definition to connect source elements to targets makes connections within folders but not between them.

The example shows rows in logical order for easy readability. However, this is not necessary, as one of the design features OpenMARS assumes that lines may not be in any order. This caused a specific challenge in defining logic for the attribute table. In the example table 8.13, the definition in the second row is more specific, so it must override the definition in row one. The following rule was composed: A definition for a specific element will always override a definition for an unspecified element group and any definition for an element will override a specification for a class. Further, a definition for a child class overrides a definition for a parent class.

Table 8.13: Start of the element connection table for PSB RF model. Reference definitions are used for defining repetitive connections.

#	OBJECT	ATTR.	VALUE
1	cellFault/failure	mean	4000a
2	ring[1]/cellGroup[1]/cell[1]/cellFault/failure	mean	1000a
3	cellFault/restoration	mean	72h
4	ringVote	atLeast	7

8.1.2 Application in the collider operations model

The driving motivation for the OpenMARS was to formalize the methodologies used for modeling collider operations and provide a tool that gives native support for them. The OpenMARS approach permits model definition with any of the most common risk assessment modeling techniques. The collider operations model uses fault tree analysis (FTA) and Markov analysis. The luminosity production is modeled with a specific function technique that formalizes the triggering mechanism that activates the function.

The most interesting feature is how the OpenMARS allows combining and connecting models made with different techniques. OpenMARS implements the connections between models with a radio-listener concept. In this concept, radios emit information on state changes in different models in a specific channel and listeners tuned to this channel receive the information. Later during the project, the author discovered that Altarica modeling formalism has synchronization mechanisms that can create logics like broadcast (one emitter and several receivers) which enable the creation of modular models [225]. This discovery occurred well after the design phase of OpenMARS, so it is unclear if the approaches are similar in the detail level.

In the collider operations model, the radio listener concept is used in four specific instances that are introduced in sections 8.1.2.1-8.1.2.4. This section shows an initial version of the collider model in the OpenMARS format that does not implement all aspects of the original ELMAS version.

8.1.2.1 Links from fault trees to Markov models

As explained in the chapter 5, the core of the collider operations model consists of Markov models and fault trees that are interlinked. To recap, fault trees affect the Markov models such that: 1) if a beam is present in a machine a failure causes a beam dump; and 2) during a fault in the collider or an injector new beam cannot be injected to a machine.

OpenMARS radio-listener concept allows defining these links formally. Radios are connected the top events of the collider and injector fault trees and a message is sent when the state of the top node changes. A failure in the collider activates the transition from the collider injection, ramp and delivery states to the ramp down state. A failure in the collider or in an injector inhibits the transition from the setup state to the injection state.

An example of this the connection between fault tree and the beam dump transition is shown in the figure 8.6. Table 8.14 shows definitions for relevant connections. Here, the beam dump is modeled with a single transition that has multiple entry points. In practice, this is a shortcut that saves a modeler from defining individual connections from each beam phase to the ramp down phase. Also, a phase "Wait" was introduced between ramp down and setup. If a fault exists in the collider or in one of the injectors collider will stay in the wait phase until the state is normal again.

Implementing this logic requires a new transition from setup to wait to model situations where a failure occurs during the setup.

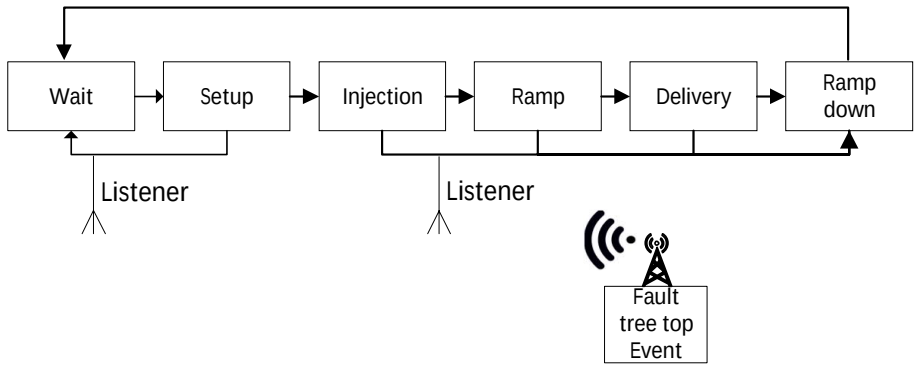

Figure 8.6: The top event of a fault tree sends a message that a listener of the beam dump connection hears.

Table 8.14: Definitions for the beam dump transition from beam phases to the ramp down phase and the relevant transitions.

#	SOURCE	TARGET
1	injection	beamDump
2	ramp	beamDump
3	delivery	beamDump
4	beamDump	rampDown
7	wait	waitToSetup
8	waitToSetup	setup
7	setup	setupToWait
8	setupToWait	wait

Table 8.15 shows radios and listeners required for the beam dump transition and the transition from setup to wait state that triggers if a failure occurs. A LHC failure triggers the beam dump transition, but any machine in the complex can trigger the transition from wait to setup. This table shows these links for the LHC, but similar connections exist for other accelerators. In the LHC and FCC models, the operation cycles were modeled up to the SPS.

Table 8.15: Radios and listeners used in the LHC to implement beam dump transition and transition from setup to wait when a failure occurs.

#	OBJECT	ATTR.	VALUE
1	/lhcFault/fault	radio	lhcFaultStartChannel
9	phases[lhc]/beamDump	listener	lhcFaultStartChannel
10	phases[lhc]/setupToWait	listener	cFaultStartChannel

Next, a rule is defined that the active cycle phase stays in the wait state during the time that the collider or an injector is being repaired. Table 8.16 shows radios and listeners required to

implement this. This implementation uses mode dependent transition rates that are formally introduced in the next section. For simplicity, here the implementation is shown only for the LHC failure behavior. A LHC failure starts the mode "faultMode" and restoration starts the mode "normalMode". The radios are defined in rows 1 and 2 and the mode changes in rows 3, 4, 6 and 7. Mode dependent transition rates from wait to setup are defined in rows 5 and 8. With these rules transition from the wait to setup is instantaneous when a fault is not present and when a fault is present the transition time is infinite[1].

Table 8.16: The radios and listeners to implement the inhibit rule from wait to setup when the collider or an injector is in a fault state.

#	OBJECT	ATTR.	VALUE
1	/complexFault/fault	radio	cFaultStartChannel
2	/complexFault/restoration	radio	cFaultEndChannel
3	simulator	wakeListener[faultMode]	cFaultStartChannel
4	simulator	waitListener[faultMode]	cFaultEndChannel
5	phases[lhc]/waitToSetup[faultMode]	mean	1000a $\approx \infty$
6	simulator	waitListener[normalMode]	cFaultStartChannel
7	simulator	wakeListener[normalMode]	cFaultEndChannel
8	phases[lhc]/waitToSetup[normalMode]	mean	0

More complex rules are required when the failures in the injector chain are linked to the cycle model. Transitions from setup to wait and the beam dump transition change, as a failure in an injector will delay the start of the injection phase but cannot cause a beam dump. Setup to wait transition is triggered by a failure in the collider or in an injector, and a failure in the collider triggers the beam dump transition. The transition rule from wait to setup is also different as both the collider and the injector chain needs to be available before the setup starts. Figure 8.7 shows logic trees to implement these rules.

The figure 8.7 shows a logic tree for a collider impact. This tree is linked to a radio that sends a message for the beam dump and setup to wait transition. Logic shows that this is always triggered when a LHC fault starts or if an injector fault occurs when the collider is in the setup phase. This tree is also linked to the mode dependent transition from wait to setup. This transition is inhibited if a collider fault or injector fault is present.

8.1.2.2 Mode dependent failure rates

The second instance, where the fault tree and the Markov models are linked, are the mode dependent failure rates. They were implemented with adding radios to every Markov state to send a message for state start and state end. Listeners are added to the fault trees to transmit the information about the state change to the fault tree nodes. In the OpenMARS, the concept of the mode dependency was generalized. Almost every parameter can be mode dependent. This is described in the section D.4.2 of the OpenMARS specification [2]. A mode dependent value forms an entity that has similarities to a Map object in Java [226], where values are mapped to specific keys. In OpenMARS a mode name is used as the key which is linked to specific value.

In the collider model, each Markov state forms a mode and the fault nodes can have a specific failure rate value for each mode. In this example, the LHC is a single fault node with phase dependent failure rates. This simplifies the model but maintains the concept of phase-dependent

[1] Here the mean transition time is set to 1000 years that is near infinite in the context of the model.

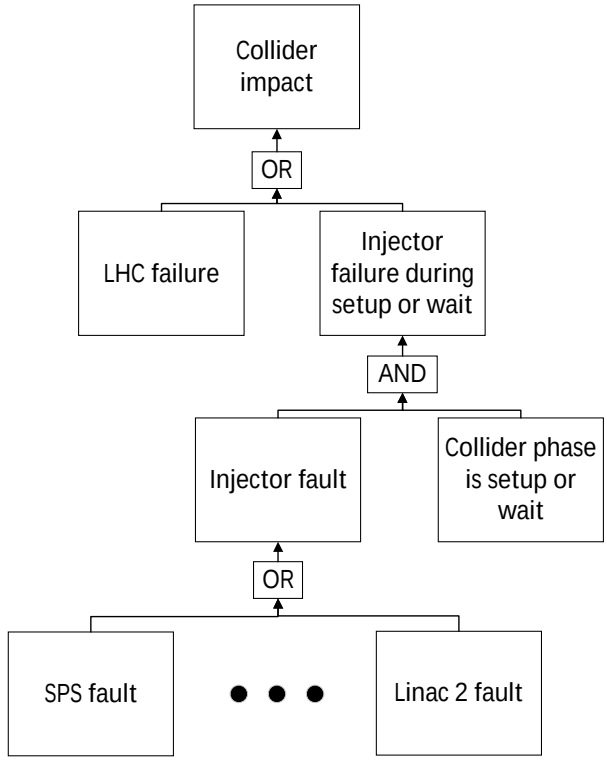

Figure 8.7: The logic tree that links collider and injector failures so that an injector fault can trigger the setup to wait transition and inhibit wait to setup transition but cannot cause a beam dump.

failure rates. Table 8.17 shows how emitters are set for phases, how they activate modes and how mode dependent failure rates are set for the LHC node. As the aim is just to introduce the concept, only the definitions for injection and ramp phases are shown. The phase dependent failure rates were used only in accelerators with significant phase lengths. In my models, I have only used them for the LHC and the FCC.

8.1.2.3 Simplified injection model

The injection phase modeling was also formalized with the OpenMARS approach. OpenMARS models Markov model state changes with transition classes. A specific class was created to model the injection process that is not presented in the specification [2]. This transition uses the radio-listener concept to link the collider state transition from injection to ramp to the injector operations. Each time the downstream injector completes a delivery phase a radio message is sent. These messages are listened by the transition to determine if the collider is filled.

There are two ways to implement this. In the original ELMAS model, the required number of injections was counted in a Java class and delivered to the model. This can be modeled with a transition where a specified number of messages needs to be received before the transition occurs. Alternatively, the logic built into the Java class could be formalized described as an OpenMARS model. At the time writing this, such a model was not fully realized. The transition between injection and ramp states can have a success chance. This is visualized in the figure 8.8 and it

Table 8.17: Rows 1-4 define the radios to send messages when a phase starts or end; rows 5-8 define commands for the modes to activate or de-activate and the rows 7 and 8 define the mode dependent failure rates.

#	OBJECT	ATTR.	VALUE
1	phases[lhc]/injection	radio	lhcInjectionChannel
2	phases[lhc]/injection	endRadio	lhcInjectionEndChannel
3	phases[lhc]/ramp	radio	lhcRampChannel
4	phases[lhc]/ramp	endRadio	lhcRampEndChannel
5	simulator	wakeListener[injectionMode]	lhcInjectionChannel
6	simulator	waitListener[injectionMode]	lhcInjectionEndChannel
7	simulator	wakeListener[rampMode]	lhcRampChannel
8	simulator	waitListener[rampMode]	lhcRampEndChannel
9	lhcFault/failure[injectionMode]	mean	5h
10	lhcFault/failure[rampMode]	mean	20h

can model a situation where only a certain percentage of the injections succeed. Also, a similar element can be attached to a fault tree. This can model a situation where an event causes a failure with a certain likelihood.

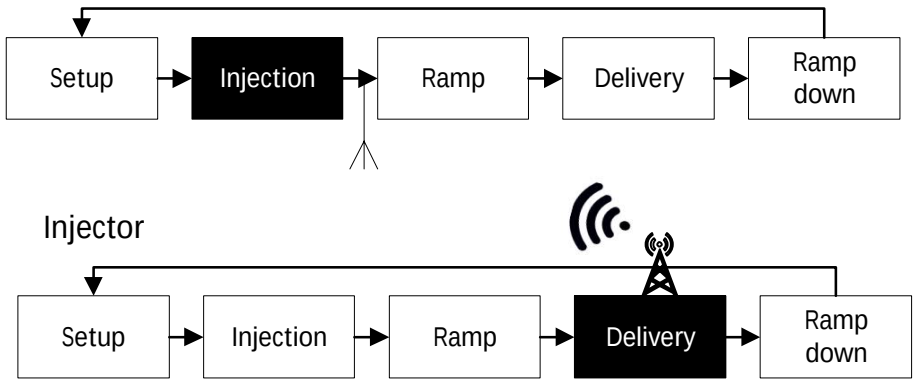

Figure 8.8: An end of a delivery phase in an injector triggers a radio message that is listened by downstream accelerator's transition between injection and ramp phase.

Table 8.18 shows definitions for a simple injection model. Rows 1-2 define the radios listener pair to transmit a message to transition between LHC injection and ramp phase each time a delivery phase ends in the SPS. Rows 3-5 define the count transition. Here 14 successful injections are needed and the probability of failing is 20%. This still leaves two features that are not modeled: i) injector super-cycles and ii) pilot beam operations. Super-cycles are required to model the SPS operations where currently every other SPS cycle delivers beam to the CERN north-area. Modeling these features will require a more detailed model.

One option that was not studied was to model the injection process Petri net based approach. This might result in an elegant solution as the required number of injection could be modeled using the tokens. The choice of modeling the operation cycle as a Markov chain came from the fact that it was the most simple solution for the modeling of the collider operation cycles. However, when

Table 8.18: Attributes to define a simple injection model.

#	OBJECT	ATTR.	VALUE
1	phases[sps]/delToRD	radio	spsDeliveryChannel
2	phases[lhc]/injToRamp	listener	spsDeliveryChannel
3	phases[lhc]	injToRamp	TrCount
4	phases[lhc]/injToRamp	count	14
5	phases[lhc]/injToRamp	prob	0.2
6	phases[sps]/waitToSetup	ifActive[fault]	/spsFault/normal
7	phases[sps]/waitToSetup	ifActive[lhc]	/phases[lhc]/injection
8	phases[sps]/waitToSetup	listener[fault]	spsFaultEndChannel
9	phases[sps]/waitToSetup	listener[lhc]	lhcInjectionChannel

the injectors are taken into account in the model, the current way of linking individual models with radios might not be the most prudent solution to this modeling problem.

8.1.2.4 Modeling integrated luminosity production

The last feature formalized with the OpenMARS approach is the luminosity production modeling. Figure 8.9 illustrates the implementation. In the collider model, the end of the collider's delivery phase sends a message that triggers the luminosity calculation function. The listener is connected to the "output" folder that contains the function. A message in the LHC delivery end channel triggers the calculation. Notably, OpenMARS supports Java styled script code that is shown in the table 8.21. Alternatively, a user could provide a link to the file that contains the code. Here, the code is the equation 4 from the reference [142] that model the integrated luminosity production.

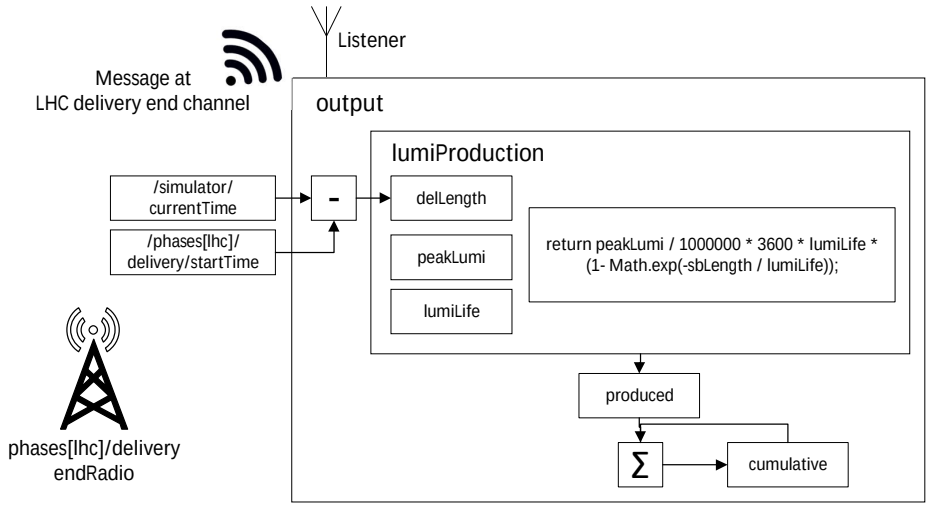

Figure 8.9: Elements and the folder structure for luminosity production calculation with the OpenMARS

Tables 8.19 - 8.21 show how the model is defined with the OpenMARS tables. The table 8.19 defines the model elements. The model contains elements of graphical programming and a user function. Usually, the graphical programming elements are always connected. However, elements of a user function are recognized as a part of this function if the user function folder contains them.

Table 8.19: Elements of the luminosity production model.

CONT	ELEMENT	CLASS
	output	Folder
output	difference	Subtraction
output	lumiProduction	UserFunction
lumiProduction	peakLumi	Value
lumiProduction	lumiLife	ValueDuration
lumiProduction	delLength	Value
output	produced	Value
output	sum	Addition
output	cumulative	Value

During the development, a long debate was held on how the connections to and from a user function and other functions should be defined. The problem is not trivial. For example in subtraction the order of the inputs matter as:

$$\text{difference} = \text{minuend} - \text{subtrahend}. \tag{8.1}$$

However, for a sum function, the order of the inputs does not matter. Table 8.20 shows a choice where minuend is defined for the difference function, but subtrahend is not. Also for a sum, the gate is not specified. In this discussion, the author held a position that each input and output should be defined. As this is close to the way, the graphical programming is implemented in the National Instruments' LabVIEW software [227]. Current OpenMARS implementation defines the connections when they are necessary.

Table 8.20: Definitions for the beam dump transition from beam phases to the ramp down phase and the relevant transitions.

SOURCE	TARGET
/simulator/currentTime	/output/difference/minuend
/phases[lhc]/delivery/startTime	/output/difference
difference	lumiProduction/sbLength
lumiProduction	produced
produced	sum
cumulative	sum
sum	cumulative

During the OpenMARS development, the inspiration on how to include user-defined code to a model came from on how National Instruments' LabVIEW defines formula nodes [227]. Figure 8.10 shows a formula node in LabVIEW that can be compared to the OpenMARS concept shown the figure 8.9. Difference between approaches is that the mechanism that triggers an

OpenMARS function is specified with radio listener pair as shown by the table 8.21 where a listener is defined for the output folder. There is a difference how the code is set up a LabVIEW formula node could have multiple named outputs. In the first implementation of the OpenMARS, the code has a return statement and the connection is made from the function to the target item.

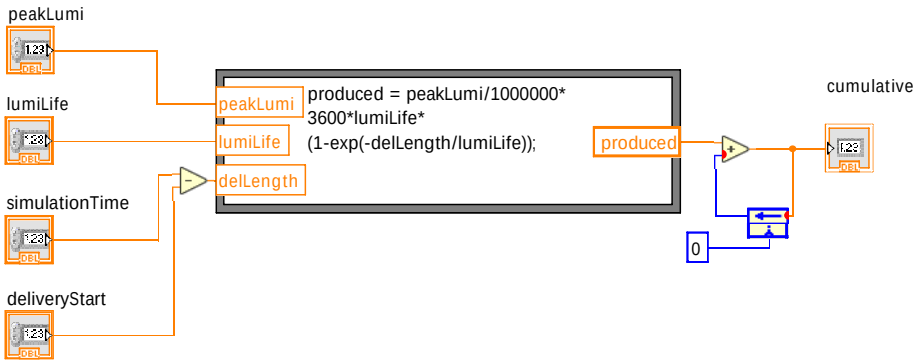

Figure 8.10: Formula node with integrated luminosity production function in LabVIEW.

Table 8.21: Attributes to define simple luminosity production model.

OBJECT	ATTR.	VALUE
output	listener	lhcDeliveryEndChannel
lumiProduction/peakLumi	value	5.33
lumiProduction/lumiLife	value	10h
output/cumulative	initialValue	0
lumiProduction	code	**return** peakLumi / 1000000 * 3600 * lumiLife * (1-Math.exp(-sbLength / lumiLife));

8.2 The accelerator reliability information system in the ARIES project

Collecting and sharing reliability data has been proposed as a solution to the lack of reliability data in various industries. For example, IAEA symposium on nuclear plants reliability in 1975 [228] stated that then high reliability was achieved in safety-related systems and components, and non-safety related failures caused the most availability issues. However, the reliability analyses for these systems were hindered by the lack of the reliability statistics. The symposium recommended establishing a reliability database.

Later, in the 1980s, OREDA project started to collect reliability data in the oil industry [105] and recently wind energy industry was recommended to adopt this approach [229]. Today wind industry has several national projects that share reliability data within individual countries like WInD-pool in Germany [230]. This idea is also used in fusion energy applications [231]. In this case, the idea was that the database would include data from various sources to provide reliability information on the relevant systems for fusion energy projects.

The aim of the accelerator reliability information system (ARIS) [4] in the ARIES project is to study if the reliability information sharing is plausible within the accelerator community[2]. The

[2]Accelerator reliability database has been proposed prior in [232].

OREDA concept was chosen as the most appropriate starting point for the reliability statistics sharing. The table 8.22 shows example of this data. Although, similarly to the fusion reliability database idea is to include data from various sources like shown in the figure 8.11.

Table 8.22: A sample of OREDA data showing failure rate for compressors. The top row shows the failure rates based on calendar time and the bottom row based on the operations time, SD stands for standard deviation [105].

Population: 47	Installations: 11	Calendar time 2.3142 10^6 hours		Operation time 2.1255 10^6 hours	
Failure mode	No of failures	Failure rate			
		Low	Mean	Up	SD
Critical	281	6.92	154.91	465.14	156.54
	281	7.2	194.4	586.72	132.21

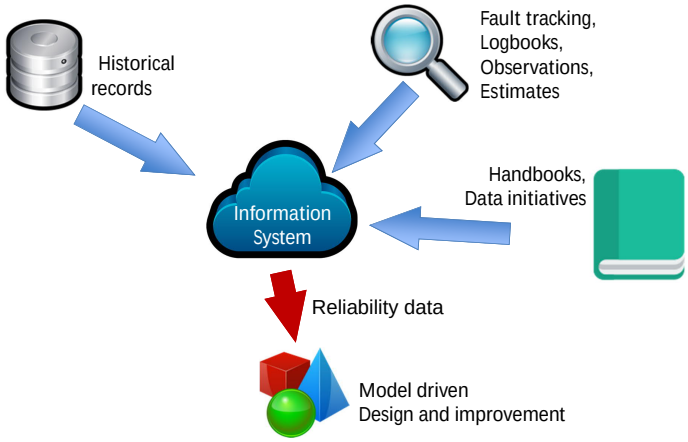

Figure 8.11: Concept of an information system where inputs from different sources are stored in a database to provide reliability data for end users [233].

Sharing reliability information has a specific challenge of data anonymity. One should not be able to distinguish specific information like the component manufacturer from the shared reliability data. For example, ISO 14224 standard [106] states that a study where competitors exchange information so that suppliers/contractors can be "ranked" incurs a real risk that the parties to the benchmarking study will arrive at a common conclusion not to use particular suppliers/contractors. The standard further states that this could be seen as collective boycott arrangements between competitors where competitors agree not to deal with individual suppliers/contractors, and this should be avoided. It is unclear how this is transformable to accelerator community. However, the plan for ARIS is only to share anonymized data.

The challenge related to the anonymized data is that the reliability information user should still understand the data after the anonymization. For this purpose, the information system should include a taxonomy that describes and classifies data from different components. For example, reference [234] uses switch technology, power and voltage rating and quality attributes to describe different power converters used in accelerators.

9 Discussion

9.1 Availability of electron-positron colliders

The availability of the FCC-ee was not formally studied during the project. Only a limited analysis of lepton collider availability was published [235]. FCC-ee would have normal conducting magnets and superconducting RF-cavities. Due to this the FCC-ee would be less reliant on the cryogenics system availability and would not need a magnet quench protection system. The plan is to operate the FCC-ee in a top-up injection mode where the beam is ramped to the collision energy before it is injected into the collider. Thus the collider is not constantly cycled.

The operational experience from electron-positron colliders is available from Large Electron-Positron Collider (LEP), Positron-Electron Project (PEP-II), and KEKB. LEP is the predecessor of the LHC and was located in the same 27 km long tunnel where LHC is now. Beam energy ranged from 45 to 104 GeV during the LEP operations. Like the hadron colliders, LEP had an operational cycle. Reports like [236] provide details on LEP availability from 1994 to 2000. During those years the physics production time was 35 to 59% of the time the machine was scheduled for physics operations. The filling and preparing for physics took 26 to 40% of the scheduled time and the rest of the time was spent in calibration and downtime.

PEP-II was 2.2 km long collider that had 3.1 GeV positron beam and 9.0 GeV electron beam. The availability of PEP-II is reported in [237] and it says that in scheduled operations during 2000 - 2004 the machine availability for physics was 61%, 26% of the time was spend in machine development, injection and tuning. Corrective maintenance took 13% of the scheduled time. Interestingly, the operational mode of the PEP-II was changed in 2004 to top-up injection [238]. Comparing references [237] and [238] shows that this increased luminosity production, but did not seem to affect the machine availability.

KEKB is a 3 km long collider that from 2003 to 2009 had 3.5 GeV positron and 8 GeV electron beams. In that time the machine generally reached 70 - 85% physics availability [239] and the downtime due to failures was 2 - 12%. In 2007, crab cavities where installed to the KEKB [240]. These cavities aim to align the bunches at the collision point and thus to increase the luminosity. Effect of this can be seen in KEKB availability as in 2007 time was spend on commissioning of the cavities. Cavities also caused failures. On average, a crab cavity failure occurred 0.5 to 1.3 times per day [240].

Top-up injection scheme and normal conducting magnets could allow FCC-ee to reach easier its physics production goals. The plan is that FCC-ee would be developed in stages [241]. The first four years for studies of Z-bosons, one year for W-bosons, three years for Higgs bosons and 4-5 years for top quarks. Also, the compact linear collider project (CLIC) plans stages energy upgrades for the linear electron-positron collider [56]. FCC-ee and CLIC energy upgrades are different. The length of the CLIC is extended each time accelerating new cavities are added while

for FCC-ee an upgrade mainly means adding new cavities as the length tunnel circular tunnel cannot be increased.

The key physics difference between linear and circular e-e+ colliders is the optimal luminosity production regime, as seen in the figure 9.1. Both designs have disadvantages. In a linear collider, particles cross the interaction point only once before they are dumped. This limits the luminosity production as the beam is used only once. The FCC-ee is limited by synchrotron radiation. In circular accelerators, the amount of synchrotron radiation increases with the beam energy. Due to this, the FCC-ee beam intensity will be lowered each time the energy is increased. When CLIC and FCC-ee are compared, FCC-ee would be more efficient up to the energy 500 GeV. Another upside of the FCC-ee is that the same tunnel could be later used for FCC-hh. The FCC-hh aims to reach 100 TeV energy while the CLIC plans 3 TeV as its highest energy. However, in a hadron collision, the energy of the colliding particles varies, when in contrast, in a lepton collider this mass is well known. So, luminosity production in a lepton and a hadron collider should be valued differently.

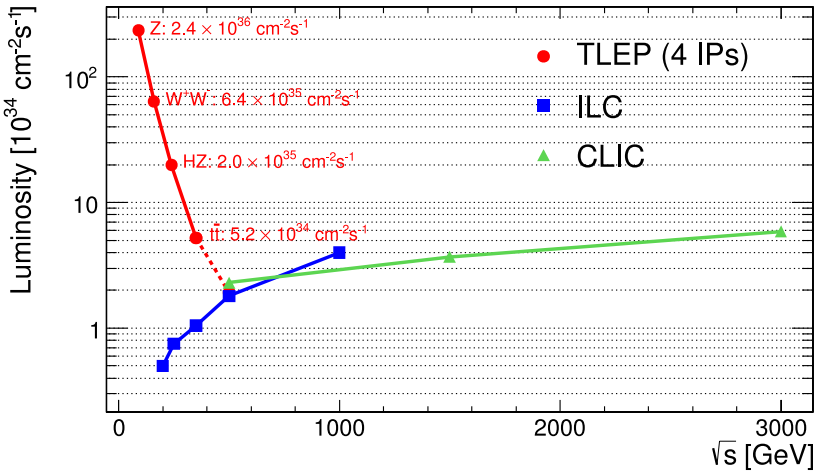

Figure 9.1: The red curve shows the FCC-ee luminosity production with different centre-of-mass energies, the green curve shows this for CLIC and the blue one for ILC, which is another linear electron-positron collider project [242].

9.2 Comments on modeling activities

There two key factors that had an adverse effect on the modeling activities. These were: lack of reliability data and lack of suitable modeling platform. Both of these were tried to be addressed by the consecutive projects on OpenMARS and accelerator reliability information database that were presented in the chapter 8.

9.2.1 Modeling data

The work shows that at CERN the reliability data different machines and subsystems lacks information on failure events' root causes and related operation modes. The efforts to improve this are going on within the accelerator fault tracking project [117]. This data would be still manually collected by the operators which makes it prone to errors. Especially the idea of linking failures

to operation modes depends heavily on the correct fault start time. As this information is essential for linking the fault to the machine operation mode.

Another problem is the statistical significance of the data due to the relatively small sample size. The idea in the reliability information database was to increase the size of the dataset by sharing the reliability data collected from different accelerator facilities. However, if collecting and sharing such data has been difficult within CERN [117], collecting and sharing this data between different institutions will likely be exponentially more difficult.

The problem with the data collection is that this data is not accumulated naturally. The author has experience in a prior project that studied the reliability of aircraft components [8], this data was readily available as the data collection was mandatory to ensure safe aircraft operations. Alternatively, one could ask what are the economic benefits of the failure data collection. A WInD-Pool business case [230] estimated that collecting maintenance data in an 80 unit offshore wind park or a 200 unit onshore park would lead to significant savings thanks to increased availability. This saving can be achieved by reducing the number of failures by 5%.

There is a problem also in assessing if the data relevant, if one imagines that failure data were available from multiple accelerators. It is an open question on how descriptive this data would be to model operations of another accelerator that potentially has different operational conditions. This is also relevant for LHC and FCC: To what degree the current LHC experience is relevant for the FCC that potentially uses new technology and design practices?

Based on this it is clear that there will always be uncertainty on the input values. The section 7.3.2 addressed this problem by testing luminosity productions sensitivity to different input values. Practically taking into account the uncertainty of the input values requires determining confidence bounds for individual inputs. This creates a multidimensional space of possible input combinations. Various techniques exist to determine the uncertainty in the results, when input confidence bounds are defined, e.g., Latin hypercube sampling is a popular option for this [187]. This, however, is calculation heavy and a practical application of this will require deploying the simulation engine in a cluster computing setup with input parameter database as proposed in the chapter 8.

9.2.2 Modeling approach

This availability study has focused on frequently occurring failures. Individually these failures have relatively limited consequences, but their occurrence rates hinder the operations. Our approach, as is, cannot be applied to rare failures with severe consequences. For example, such incident occurred in LHC in 2008 due to a defective joint between superconducting cables delaying the physics program for more than a year [150]. Prevention of these severe failures requires detailed studies of specific systems.

With these limitations, the model itself works fine. However, ideally, the shift away from the ELMAS platform should have occurred earlier. As chapter 5 shows, a significant development work that was carried out to implement the model in ELMAS and at this stage the OpenMARS implementation contain only the key features. It is an interesting thought experiment what would have been the course of action if the collaboration would not have taken place.

When faced with this situation the IFMIF accelerator study chose to develop an in-house simulator model based on the AvailSim that was originally developed for ILC study [243]. Alternatively, the study could have searched for open source simulator engines and built the model using one of them as a platform. As this was not the case, a thorough study of the potential options was not conducted. For example, the RBD driven Petri net [244] and the conjoint system model (CSM) [245] both combine Petri net and RBD techniques.

However, the most interesting platforms seem to be so-called model-based safety assessment formalisms (MBSA) [82]. Examples of MBSA are Altarica and Figaro languages that are both available as open source projects[1]. Altarica was originally developed in the LaBRI research unit and has been adapted to oil and aviation industry [246]. Figaro was developed in Électricité de France (EDF) for nuclear safety studies [247]. Further research would be required to understand if one of these languages could be used as the platform for the collider operations model. Referee comments for the OpenMARS publication [5] suggested that these approaches have the same expressive power to develop models.

9.2.3 Potential further applications of OpenMARS

Studying industry cases gives ideas for further OpenMARS applications. The closest matches are the industry cases where the link between availability and production performance is not trivial. The following cases are described in detail: (i) a nickel reduction plant production line [248], (ii) dynamic process simulation of a LNG fuel storage tanks risk assessment [249], (iii) availability of offshore installation [250] and manufacturing lines [251]. These cases show that there is a demand for complex operations and production modeling in industry.

In the nickel reduction plant, the production process has multiple stages [248]. Between these stages are tanks for liquids and silos for powder material that can act as a buffer in cases where a failure or maintenance halts part of the production. Thus, the production modeling has to take into account the filling levels of the buffers, which is non-trivial. OpenMARS's function modeling tools could be used for simulating the production logic and the buffer filling levels, while the fault trees could be used to model failures of different process stages. The actual model was used for choosing optimal times for planned maintenance and helped to improve the operational reliability.

The dynamic process simulation of a LNG fuel storage tanks risks assessment [249] shows a more complex example where failure and process modeling are combined. In this case, failure modeling was relatively simple. The model defined the main components and used exponential repair and failure rates from the OREDA handbook. However, this model was combined with a dynamic process model to simulate liquid and vapor flows of natural gas within the system and thermodynamic modeling of the LNG fuel tanks. This model was used for assessing process availability, LNG fuel supply and risk of overpressure within a fuel storage tank.

In the model for offshore installation simulates the production of gas, oil, and water in an offshore plant. Gas is produced for export, power generations and so-called gas lift system that injects highly pressurized gas back to the oil well to lift the fluids. The analyses found that due to maintenance activities and failures system can have more than 400 different states that lead to 7 different production levels for gas, oil, and water. It is also notable that failure modeling takes into account system degeneration. If such a model were to be created in the OpenMARS, the key feature would be linking the failure states to the production functions. One possible solution would be to make production rates mode dependent variables and let different failures activate different modes.

The selected manufacturing line case [251] models system reliability and production quality degradation; preventive maintenance and production quality inspections; and an inventory that acts as a buffer between production stages. The model is based on the SIMAN simulation language used in ARENA software. This example shows that it is likely that the collider operations model could have also been implemented with SIMAN language. The model consists of high amounts of custom logic. For example, preventive maintenance is linked to the amount of cumulative

[1] https://altarica.labri.fr/wp/ , https://sourceforge.net/p/visualfigaro/wiki/Home/

production, or it is carried out during corrective maintenance. Furthermore, the production rates depend on the buffer fill levels. The answer for the question could this model be implemented with OpenMARS is likely yes, but would require a fully realized model to prove it.

10 Contributions of the thesis and conclusions

This thesis includes and extends material from several other publications to which the author has contributed. The contributions can be grouped into three categories: collider operations model, requirements for the FCC-hh, and the OpenMARS approach and Accelerator reliability database that were inspired by this work. This chapter presents a summary of these contributions with conclusions.

The collider availability and operations model was first presented in the publication [1] that was written by the author and revised together with the co-authors. The publication presented the first version of the collider availability model, its validation with LHC operations data, and studies of different operation scenarios. This model is also the main contribution of this thesis that extends the findings of the publication [1]: chapter 3 provided the basic background on reliability engineering methods, chapter 4 provided background on the collider operations, and chapter 5 presented the model as it was implemented in Ramentor Oy's ELMAS software. Chapter 6 showed the model validation. Compared to the publication [1], this thesis provides more details on the background, model implementation and injection modeling that was not developed when the publication was written.

Availability and operations requirements for FCC-hh: The author has presented requirements for FCC-hh operations and availability in references [3] and [6]. The report [3] was written by the author and revised together with the co-authors and the chapter in the reference [6] was written in collaboration with A. Apollonio. Findings from these contributions were repeated and elaborated in the chapter 7. In brief, the high integrated luminosity production goals create demanding requirements of the operations schedule, short optimal delivery phase time with ultimate parameters increase the importance of turnaround time and injector availability, and increasing beam energy and vastness of infrastructure result in challenging system availability requirements. However, a concrete system level availability allocation was not provided.

OpenMARS approach: Desire to improve the ELMAS modeling platform led to a collaboration to develop OpenMARS approach. This approach has been described in the report [2] and in a publication [5]. It supports the most common risk assessment and operation modeling techniques. OpenMARS allows combining and connecting models defined with different techniques. This ensures that a modeler has a high degree of freedom to accurately describe the modeled system without limitations imposed by an individual technique. Section 8.1 of this thesis gave an introduction to this approach and presented some praise and critique on the OpenMARS approach. More importantly, this thesis showed a version of the collider availability model implemented with the OpenMARS approach that has not been published previously.

Accelerator reliability database: The author also contributed to the early development of the accelerator reliability database that has been documented in the report [4]. Although the author did not write this report, it is a result of collaboration where the author was a key contributor. The report presents a literature review on maintenance data collection and sharing. One key issue discovered during the project was the limited quantity and quality of reliability data from accelerator systems. This has been an issue in practically every accelerator availability study. Section 8.2 of this thesis provides some comments on this activity. Notable, for a future project like the FCC, the most important data is not what availability can be achieved today, but how this availability could be improved. Assessing this will require information on the failure root causes. Non-subjective information on the root causes will only be available if accelerator facilities further commit to collecting reliability data.

This thesis started with three questions:

- How to estimate and model collider's availability while taking into account the whole collider complex and the collider operations?
- Can the modern simulation approach be applied to modeling reliability and availability of research infrastructures?
- Is this model able to produce useful insights for the FCC-hh?

The answers to these questions can be evaluated based on the contributions. This thesis and the earlier publication [1] demonstrate that the developed model works and can capture the essential aspects of collider operations that affect the luminosity production and availability for physics. However, the lack of reliability data from research infrastructures hinders its applicability. This led to the study on reliability information sharing. This lack of data also prevented providing a system level availability allocation, as it would have required information on system failure root causes to assess how the system reliability can be improved in the future.

In retrospect, the project management focused on long-term goals of quickly securing collaborations, establishing the modeling platform and reliability information system. Alternatively, the project could have focused more on short-term goals and try to study more the reliability aspects of the FCC design. However, this probably would have been too early in the FCC study. For example, contents of the chapter 7 are mostly from the last year of the project and before that, many aspects of the collider designs were open. The project might have been best served with a slower approach where the project would have tried to understand what is feasible in term of reliability engineering in a research environment and to establish what are the use cases for the modeling, before committing to activities.

Bibliography

[1] A. Niemi, A. Apollonio et al. Availability modeling approach for future circular colliders based on the LHC operation experience. *Phys. Rev. Accel. Beams*, 19:121003, 2016. http://dx.doi.org/10.1103/PhysRevAccelBeams.19.121003.

[2] J.-P. Penttinen, A. Niemi et al. An open modelling approach for availability and reliability of systems - OpenMARS. Technical Report CERN-ACC-2018-0006, CERN, Geneva, 2018. https://cds.cern.ch/record/2302387.

[3] A. Niemi, A. Apollonio et al. Considerations on operation schedule and maintenance aspects of FCC-hh. Technical Report CERN-ACC-2018-0013. https://cds.cern.ch/record/2317097.

[4] A. Preinerstorfer, H. Humer et al. Bibliography and state of the art of reliability information systems. Technical report, 2018. https://doi.org/10.5281/zenodo.1292068.

[5] J.-P. Penttinen, A. Niemi et al. An open modelling approach for availability and reliability of systems. *Reliability Engineering & System Safety*, 183:387–399, 2019.

[6] M. Benedikt, M. Capeans Garrido et al, editor. *Future Circular Collider Study*, volume 3: The Hadron Collider (FCC-hh) Conceptual Design Report, 2018. Submitted to Eur. Phys. J. ST.

[7] R. Alemany Fernandez, A. Apollonio et al. FCC-hh turn-around cycle. Technical Report CERN-ACC-2016-0341, CERN, Geneva, 2016. https://cds.cern.ch/record/2239138.

[8] A. Niemi. Applying statistical methods to improve maintenance of aircraft system components. Master's thesis, Aalto university, Espoo, Finland, 2013. Written in Finnish language.

[9] K. Mahlamäki, A. Niemi et al. Importance of maintenance data quality in extended warranty simulation. *Int. J. COMADEM*, 19:3–10, 2016.

[10] O. Brüning, P. Collier et al., editor. *LHC design report: Vol. I The LHC main ring*. CERN, 2004. http://dx.doi.org/10.5170/CERN-2004-003-V-1.

[11] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(1):1 – 29, 2012. http://dx.doi.org/10.1016/j.physletb.2012.08.020.

[12] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716(1):30 – 61, 2012. http://dx.doi.org/10.1016/j.physletb.2012.08.021.

[13] O. Brüning and L. Rossi. Introduction to the HL-LHC project. In *The High Luminosity Large Hadron Collider*, chapter 1, pages 1–17. World Scientific Publishing Co. Pte. Ltd., 2015. http://dx.doi.org/10.1142/9789814675475_0001.

[14] S. Myers and W. Schnell. Technical Report CERN-LHC-Note-1, CERN, Geneva, Apr 1983. http://cds.cern.ch/record/443058?ln=en.

[15] F. Zimmermann. High-energy physics strategies and future large-scale projects. *Nucl. Instrum. Methods Phys. Res. B*, 355:4 – 10, 2014. http://dx.doi.org/10.1016/j.nimb.2015.03.090.

[16] R. Heuer. Future Circular Collider Study - FCC Mandate, December 2013.

[17] D. Schulte. Preliminary collider baseline parameters: Deliverable D1.1. Technical Report CERN-ACC-2015-0132, CERN, Geneva, 2015. https://cds.cern.ch/record/2059230.

[18] A. Apollonio. *Machine Protection: Availability for Particle Accelerators*. PhD thesis, Vienna University of Technology, Vienna, May 2015.

[19] A. Apollonio, M. Jonker et al. HL–LHC: Integrated luminosity and availability. In *IPAC2013: Proceedings of the 4th International Particle Accelerator Conference*, pages 1352 – 1354, Geneva, Jun 2013. JACoW.

[20] M. Valette, A. Apollonio et al. Requirements for crab cavity system availability in HL-LHC. In *Proc. of International Particle Accelerator Conference (IPAC'17), Copenhagen, Denmark, 14-19 May, 2017*, number 8 in International Particle Accelerator Conference, pages 2097–2100, Geneva, May 2017. JACoW. https://doi.org/10.18429/JACoW-IPAC2017-TUPVA022.

[21] Ramentor Oy. ELMAS - event logic modeling and analysis software. http://www.ramentor.com/products/elmas/.

[22] W. E. Vesely, F. F. Goldberg et al. *Fault Tree Handbook*. U.S. Nuclear Regulatory Commission, Washington, D.C., January 1981. https://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0492/.

[23] V. Barbu and N. Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications: Their use in Reliability and DNA Analysis*. Springer, New York, 2008. https://doi.org/10.1007/978-0-387-73173-5_1.

[24] J.-P. Penttinen and T. Lehtinen. Advanced fault tree analysis for improved quality and risk assessment. In *Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015)*, pages 471–478. Springer International Publishing, 2016. http://dx.doi.org/10.1007/978-3-319-27064-7_45.

[25] CERN. The European strategy for particle physics update 2013. In *16th Session of European Strategy Council*, number CERN-Council-S/106, Geneva, May 2013.

[26] M. Benedikt and F. Zimmermann. Status of the Future Circular Collider Study. In *25th Russian Particle Accelerator Conference, RuPAC2016*, pages 34 – 38, St. Petersburg, Nov 2016. http://dx.doi.org/10.18429/JACoW-RuPAC2016-TUYMH01.

[27] V. D. Shiltsev. High-energy particle colliders: past 20 years, next 20 years, and beyond. *Phys.-Uspekhi*, 55(10):965, 2012. https://doi.org/10.3367/UFNe.0182.201210d.1033.

[28] D. Tommasini, B. Auchmann et al. The 16 T dipole development program for FCC. *IEEE Trans. Appl. Supercond.*, 27(4):1–5, Jun 2017. https://doi.org/10.1109/TASC.2016.2634600.

[29] S. Ritz, M. Demarteau et al. *Building for Discovery: Strategic Plan for U.S. Particle Physics in the Global Context*. U.S. Department of Energy, Office of Science, 2014.

[30] T. Golling, M. Hance et al. Physics at a 100 TeV pp collider: beyond the standard model phenomena. In *Physics at the FCC-hh*, chapter 3. CERN, 2016.

[31] Physics at a 100 TeV pp collider: Higgs and EW symmetry breaking studies. In R. Contino, D. Curtin et al., editor, *Physics at the FCC-hh*, chapter 2. CERN, 2016.

[32] M.L. Mangano and G. Zanderighi, et al. Physics at a 100 TeV pp collider: Standard model processes. In *Physics at the FCC-hh*, chapter 1. CERN, 2016.

[33] A. Franklin. *Shifting standards: experiments in particle physics in the twentieth century*. University of Pittsburgh Press, 2013.

[34] E. Gibney. Hopes for revolutionary new LHC particle dashed. *Nat.*, 536(7615):133–134, 2016. https://doi.org/10.1038/nature.2016.20376.

[35] W. Herr and B. Muratori. Concept of luminosity. In *CAS - CERN Accelerator School: Intermediate Course on Accelerator Physics*, pages 361 – 377. CERN, 2003. http://dx.doi.org/10.5170/CERN-2006-002.361.

[36] J. Baglio, A. Djouadi et al. Prospects for Higgs physics at energies up to 100 TeV. *Rep. Prog. Phys.*, 79(11):116201, 2016. http://stacks.iop.org/0034-4885/79/i=11/a=116201.

[37] G. Apollinari, I. Béjar Alonso et al., editor. *High-Luminosity Large Hadron Collider (HL-LHC): Technical Design Report V. 0.1*. CERN Yellow Reports: Monographs. CERN, Geneva, 2017. https://cds.cern.ch/record/2284929.

[38] B. C. Paulson. Designing to reduce construction costs. *J. Constr. Div.*, 102(4):587–592, 1976.

[39] M. Florio, S. Forte et al. Forecasting the socio-economic impact of the Large Hadron Collider: A cost–benefit analysis to 2025 and beyond. *Technol. Forecast. Soc. Chang.*, 112:38 – 53, 2016. https://doi.org/10.1016/j.techfore.2016.03.007.

[40] International Electrotechnical Commission. International electrotechnical vocabulary - Part 192: Dependability. International standard IEC 60050-192:2015, 2015.

[41] J. D. Kalbfleisch and R. L. Prentice. *The Statistical Analysis of Failure Time Data*, chapter 1, pages 1–30. Wiley-Blackwell, 2011.

[42] B. Bertsche. *Reliability in Automotive and Mechanical Engineering*. Springer, 2008. https://doi.org/10.1007/978-3-540-34282-3.

[43] T. Kanti Agustiady and E.-A. Cudney. *Total Productive Maintenance: Strategies and Implementation Guide*, chapter 6, pages 111 – 122. CRC Press, 2015.

[44] Mechanical Engineering and Metals Industry Standardization in Finland. Maintenance. Maintenance key performance indicators. European standard EN 15341:2007, 2007.

[45] S.-B. Lee, A. Katz et al. Getting the quality and reliability terminology straight. *IEEE Trans. Compon. Packag. Manuf. Technol. Part A*, 21(3):521 – 523, Sep 1998. https://doi.org/10.1109/95.725217.

[46] M. Daňhel and H. Kubátová. Dependability or reliability in the real world history, terminology, prediction. In *2017 6th Mediterranean Conference on Embedded Computing (MECO)*. IEEE, Jul 2017. https://doi.org/10.1109/MECO.2017.7977177.

[47] R. Smith and R. Keith Mobley. *Rules of Thumb for Maintenance and Reliability Engineers*, chapter 5, page 66. Elsevier, Burlington, 2007.

[48] V. P. Nelson. Fault-tolerant computing: fundamental concepts. *Comput.*, 23(7):19–25, July 1990. https://doi.org/10.1109/2.56849.

[49] J. Gutleber. Accelerator reliability and availability training. *Accelerating News*, (16), 2016.

[50] Y. Cho. Reliability/Availability considerations of the SNS. In *Second Workshop on Utilisation and Reliability of High Power Accelerators*, pages 71 – 78, Aix-en-Provence, France, November 1999. https://doi.org/10.1787/9789264194847-en.

[51] J. L. Biarrotte, A. C. Mueller, et al. Accelerator reference design for the MYRRHA European ADS demonstrator. In *Proceedings of the 25th International Linear Accelerator Conference, LINAC-2010*, pages 440 – 442, Tsukuba, Japan, September 2010.

[52] A. Apollonio, S. Gabourin et al. Availability studies for LINAC4 and machine protection requirements for LINAC4 commissioning. In *IPAC 2014*, pages 3807 – 3809, Dresden, Germany, June 2014.

[53] E. Bargalló and R. Andersson et al. ESS reliability and availability approach. In *IPAC 2015*, pages 1033 – 1035, Richmond, USA, May 2015.

[54] E. Bargalló Font. *IFMIF Accelerator Facility RAMI Analyses in the Engineering Design Phase*. PhD thesis, Polytechnic University of Catalonia, Barcelona, Spain, February 2014.

[55] T. Himel, J. Nelson et al. Availability and reliability issues for ILC. In *PAC 2007*, pages 1966–1969, Albuquerque, June 2007. IEEE. https://doi.org/10.1109/PAC.2007.4441325.

[56] M. Aicheler and P. Burrows et al., editor. *A Multi-TeV linear collider based on CLIC technology: CLIC Conceptual Design Report*, page 789. CERN, October 2012.

[57] K. T. Dixon and J. Franciscovich. SSC accelerator availability allocation. In *Supercollider 3*, pages 1109 – 1123. Springer, 1991. https://doi.org/10.1007/978-1-4615-3746-5_102.

[58] G. Apollinari and I. Béjar Alonso et al., editors. *High-Luminosity Large Hadron Collider (HL-LHC) Technical Design Report V. 0.1*, volume 4 of *CERN Yellow Reports: Monographs*. CERN, Geneva, 2017. https://doi.org/10.23731/CYRM-2017-004.

[59] The CEPC-SPPC Study Group. Machine layout and performance. In *CEPC-SppC Preliminary Conceptual Design Report*, volume II: Accelerator, chapter 3, page 39. Institute of High Energy Physics, CAS, Beijing, China, March 2015.

[60] B. Todd, L. Ponce et al. LHC availability 2017: Standard proton physics. Technical Report CERN-ACC-NOTE-2017-0063, CERN, Dec 2017. https://cds.cern.ch/record/2294852.

[61] J. Galambos. SNS performance and the next generation of high power accelerators. In *Proceedings of PAC2013*, pages 1443 – 1447, Geneva, 2013. JACoW.

[62] A. Lüdeke. Operation event logging system of the swiss light source. *Phys. Rev. ST Accel. Beams*, 12:024701, Feb 2009. https://doi.org/10.1103/PhysRevSTAB.12.024701.

[63] A. Lüdeke, M. Bieler et al. Common operation metrics for storage ring light sources. *Phys. Rev. Accel. Beams*, 19:082802, Aug 2016. https://doi.org/10.1103/PhysRevAccelBeams.19.082802.

[64] F. Zimmermann. Collider beam physics. *Rev. Accel. Sci. Technol.*, 7:177–205, 2014. https://doi.org/10.1142/S1793626814300096.

[65] Finnish Standards Association SFS. Risk management – Risk assessment techniques. European standard IEC/ISO 31010:2009, 2009.

[66] Isograph Limited. *Isograph Reliability Workbench® user guide*, 12.1 edition, 2015.

[67] International Electrotechnical Commission. Analysis techniques for system reliability – Procedure for failure mode and effects analysis (FMEA). International standard IEC 60812, 2006.

[68] M. Rausand. *Reliability of Safety-Critical Systems: Theory and Applications*, chapter 5.3. Wiley, 2014.

[69] F. Salfner and M. Malek. Using hidden semi-Markov models for effective online failure prediction. In *2007 26th IEEE International Symposium on Reliable Distributed Systems (SRDS 2007)*, pages 161–174, 2007.

[70] M. Reščič and R. Seviour. Modelling accelerator reliability. In *4th International Workshop on ADSR systems and Thorium*, 2016.

[71] K.E. Moore and S.M. Gupta. Petri net models of flexible and automated manufacturing systems: a survey. *Int. J. Prod. Res.*, 34(11):3001, 1996.

[72] T. Rieker, P. Zeiler et al. Reliability analysis of a hybrid car drive system with EC-SPN. In *2015 Annual Reliability and Maintainability Symposium (RAMS)*, pages 1–7. https://doi.org/10.1109/RAMS.2015.7105197.

[73] Z. Wang, M. Atli et al. Coloured stochastic petri nets modelling for the reliability and maintenance analysis of multi-state multi-unit systems. *J. Manuf. Tech. Manag.*, 25(4):476–490, 2014. https://doi.org/10.1108/JMTM-04-2013-0045.

[74] P. Škňouřilová and R. Briš. Coloured Petri nets and a dynamic reliability problem. *Proc. Inst. Mech. Eng.*, 222:635–642, 12 2008.

[75] Q. Gong, K.K. Lai et al. Supply chain networks: Closed Jackson network models and properties. *Int. J. Prod. Econ.*, 113(2):567 – 574, 2008. https://doi.org/10.1016/j.ijpe.2007.10.013.

[76] T. Kurtoglu and I. Y. Tumer. A graph-based fault identification and propagation framework for functional design of complex systems. *J. Mech. Des.*, 130(5):051401, 2008. https://doi.org/10.1115/1.2885181.

[77] H. Langseth and L. Portinale. Bayesian networks in reliability. *Reliab. Eng. Syst. Saf.*, 92(1):92 – 108, 2007. https://doi.org/10.1016/j.ress.2005.11.037.

[78] A. K. Verma, s. Ajit et al. *Reliability and Safety Engineering*, chapter 4, pages 123–159. Springer London, London, 2016. https://doi.org/10.1007/978-1-4471-6269-8_4.

[79] E. Zio. Reliability engineering: Old problems and new challenges. *Reliab. Eng. Syst. Saf.*, 94(2):125 – 141, 2009. https://doi.org/10.1016/j.ress.2008.06.002.

[80] C. G. Cassandras and S. Lafortune, editors. *Introduction to Discrete Event Systems*, chapter 10, pages 557–615. Springer US, Boston, 2008. https://doi.org/10.1007/978-0-387-68612-7_10.

[81] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? Does it matter? *Struct. Saf.*, 31(2):105 – 112, 2009. https://doi.org/10.1016/j.strusafe.2008.06.020.

[82] O. Lisagor, T. Kelly, and R. Niu. Model-based safety assessment: Review of the discipline and its challenges. In *The Proceedings of 2011 9th International Conference on Reliability, Maintainability and Safety*, pages 625–632, 2011. https://doi.org/10.1109/ICRMS.2011.5979344.

[83] A. Alrabghi and A. Tiwari. State of the art in simulation-based optimisation for maintenance systems. *Comp. Ind. Eng.*, 82:167 – 182, 2015. https://doi.org/10.1016/j.cie.2014.12.022.

[84] M. Lipaczewski, F. Ortmeier et al. Comparison of modeling formalisms for safety analyses: SAML and AltaRica. *Reliab. Eng. Syst. Saf.*, 140:191 – 199, 2015. https://doi.org/10.1016/j.ress.2015.03.038.

[85] L. Hardy. Accelerator reliability – availability. In *Proceedings of EPAC 2002*, pages 149 – 153, Geneva, Aug. 2002. European Physical Society Interdivisional Group on Accelerators (EPS-IGA) and CERN.

[86] O. Rey Orozco, E. Bargalló et al. Reliability and availability modeling for accelerator driven facilities. In *IPAC2014: Proceedings of the 5th International Particle Accelerator Conference*, pages 3803 – 3806, Geneva, Jul. 2014. JACoW. https://doi.org/10.18429/JACoW-IPAC2014-THPRI019.

[87] E. S. McCrory and P. W. Lucas. A model of the Fermilab collider for optimization of performance. In *Proceedings Particle Accelerator Conference*, volume 1, pages 449 – 451, Dallas, May 1995. IEEE.

[88] E. S. McCrory. Monte Carlo of Tevatron operations, including the Recycler. In *PAC 2005*, pages 2479 – 2481, Knoxville, May 2005.

[89] I. Balan. *Adjoint sensitivity analysis procedure of Markov chains with application on reliability of IFMIF accelerator system facilities*. PhD thesis, Karlsruhe Institute of Technology, Karlsruhe, May 2005.

[90] C. Tapia, J. Dies et al. Exploration of reliability databases and comparison of former IFMIF's results. *Fusion Eng. Des.*, 86:2726 – 2729, 2011. http://dx.doi.org/10.1016/j.fusengdes.2010.12.078.

[91] E. Bargalló, P. Sureda et al. Availability simulation software adaptation to the IFMIF accelerator facility RAMI analyses. *Fusion Eng. Des.*, 89:2425 – 2429, 2014. https://doi.org/10.1016/j.fusengdes.2013.12.004.

[92] N. Phinney, T. Himel et al. Reliability simulations for a linear collider. In *EPAC 2004*, pages 857 – 860, Lucerne, July 2004.

[93] M. Motyka. Impact of Usability for Particle Accelerator Software Tools Analyzing Availability and Reliability. Master's thesis, Blekinge Institute of Technology, Blekinge, 2017.

[94] L. Burgazzi and P. Pierini. Reliability studies of a high-power proton accelerator for accelerator-driven system applications for nuclear waste transmutation. *Reliab. Eng. Syst. Saf.*, 92(4):449 – 463, 2007. https://doi.org/10.1016/j.ress.2005.12.008.

[95] A. Pitigoi and P. Fernández Ramos. Modeling high–power accelerators reliability – Reliability of SNS LINAC (SNS–ORNL); reliability modelling for MAX LINAC (Myrrha project). In *AccApp 2013*, Bruges, Belgium, August 2013.

[96] O. Rey Orozco, A. Apollonio et al. Dependability studies for CERN PS Booster RF system upgrade. In *IPAC 2016*, pages 4159 – 4162, Busan, Korea, May 2016.

[97] O. Rey Orozko, A. Apollonio et al. Methodology applied for dependability studies on the Compact Linear Collider. In *Proc. of International Particle Accelerator Conference (IPAC'17), Copenhagen, Denmark, 14-19 May, 2017*, number 8 in International Particle Accelerator Conference, pages 1943–1946, Geneva, May 2017. JACoW. https://doi.org/10.18429/JACoW-IPAC2017-TUPIK100.

[98] P. Gupta. Maintenance budget estimation for a particle accelerator system under its contextual conditions: A case study. *Int. J. Syst. Assur. Eng. Manag.*, 8(2):1143–1153, Nov 2017. https://doi.org/10.1007/s13198-017-0580-1.

[99] R. Schmidt, R. Assmann et al. Protection of the CERN Large Hadron Collider. *New J. Phys.*, 8(11):290, 2006. https://doi.org/10.1088/1367-2630/8/11/290.

[100] P. Ninin. IEC 61508 experience for the development of the LHC functional safety systems and future perspective. In *ICALEPCS*, pages 400 – 402, Kobe, Japan, October 2009.

[101] F. S. Nowlan and H. F. Heap. *Reliability-centered maintenance*. United Airlines, 1978.

[102] J. Moubray. *Reliability Centered Maintenance*. Industrial Press Inc., 2 edition, 2005.

[103] P.-E. Hagmark and J. Laitinen. Modeling of age-dependent failure tendency from incomplete data. In J. Lee, J. Ni et al., editor, *Engineering Asset Management 2011*, pages 449–459, London, 2014. Springer London. https://doi.org/10.1007/978-1-4471-4993-4_40.

[104] H. Ma and W. Q. Meeker. Strategy for planning accelerated life tests with small sample sizes. *IEEE Trans. Reliab.*, 59(4):610–619, 2010. https://doi.org/10.1109/TR.2010.2083251.

[105] SINTEF and NTNU, editor. *OREDA Offshore and onshore reliability data handbook*, volume 1. OREDA participants, 6 edition, 2015.

[106] Finnish Standards Association SFS. Petroleum, petrochemical and natural gas industries. Collection and exchange of reliability and maintenance data for equipment. European standard EN ISO 14224:2016, 2016.

[107] A. Peter, D. Das, et al. Critique of MIL-HDBK-217. In *Reliability Growth: Enhancing defense system reliability*, pages 203 – 246. The national academies press, Washington, D.C. USA, 2015.

[108] M. Held and K. Fritz. Comparison and evaluation of newest failure rate prediction models: FIDES and RIAC 217Plus. *Microelectron. Reliab.*, 49:967 – 971, 2009. https://doi.org/10.1016/j.microrel.2009.07.031.

[109] Advisory Group on Reliability of Electronic Equipment. Allocation of system failure rates. In *Reliability of military electronic equipment*, pages 52 – 57. U.S. Government Printing Office, 1957.

[110] E. D. Karmiol. *Reliability apportionment, preliminary report EIAM-5, Task II*, pages 10–22. General Electric, 1965. as cited in [120] and [113].

[111] R.W. Saaty. The analytic hierarchy process - what it is and how it is used. *Math. Model.*, 9:161 – 176, 1987. https://doi.org/10.1016/0270-0255(87)90473-8.

[112] S. Virtanen. Reliability in product design-specification of dependability requirements. In *Annual Reliability and Maintainability Symposium. 1998 Proceedings. International Symposium on Product Quality and Integrity*, pages 82–88, 1998. https://doi.org/10.1109/RAMS.1998.653596.

[113] C.-S. Liaw, Y.-C. Chang et al. ME-OWA based DEMATEL reliability apportionment method. *Expert Syst. Appl.*, 38:9713 – 9723, 2011. https://doi.org/10.1016/j.eswa.2011.02.029.

[114] *The first national particle accelerator conference*, volume 12 of *IEEE Trans. Nucl. Sci.*, 1965.

[115] R. C. Sass and H. Shoaee. CATER: An online problem tracking facility for SLC. In *Proceedings of the 1993 particle accelerator conference*, pages 1946 – 1948, 1993.

[116] CERN. *The workshop on machine availability and dependability for post-LS1 LHC*, Nov 2013. https://indico.cern.ch/event/277684.

[117] B. Todd, A. Apollonio et al. Availability working group & Accelerator Fault Tracker – Where do we go from here? In *LHC Performance Workshop (Chamonix 2016)*, Chamonix, January 2016. https://indico.cern.ch/event/448109/contributions/1942065.

[118] C. Roderick, D. Martin et al. Accelerator Fault Tracking at CERN. In *Proc. of International Conference on Accelerator and Large Experimental Control Systems (ICALEPCS'17), Barcelona, Spain, 8-13 October 2017*, pages 397–400, Geneva, Jan 2018. JACoW. https://doi.org/10.18429/JACoW-ICALEPCS2017-TUPHA013.

[119] K. Smedley. Reliability analysis for LEB ring magnet power system in SSC. *IEEE Trans. Nucl. Sci.*, 39:1170–1174, 1992. https://doi.org/10.1109/23.159778.

[120] J. C. Schryver and M. J. Haire. Spallation Neutron Source availability top-down apportionment using characteristic factors and expert opinion. Technical Report ORNL/TM-1999/206, Oak Ridge national laboratory, Oak Ridge, USA, 1999. https://www.osti.gov/servlets/purl/14313.

[121] O. Rey Orozco, A. Apollonio et al. Availability allocation to particle accelerator subsystem by complexity criteria. In *Proceedings of IPAC2018*. 2018.

[122] C. Zhang and S. X. Fang. Synchrotrons. In A. W. Chao, K. H. Mess et al., editor, *Handbook of accelerator physics and engineering*, chapter 1.6.21, pages 58 – 59. World Scientific, 2 edition, 2013. https://doi.org/10.1142/8543.

[123] R. Bailey and P. Collier. Standard filling schemes for various LHC operation modes. Technical Report LHC-project note 323, 2003. http://cds.cern.ch/record/691782.

[124] E. Jensen. Cavity basics. In *CAS - CERN Accelerator School: RF for Accelerators*, pages 259 – 275. CERN, 2012. https://doi.org/10.5170/CERN-2011-007.259.

[125] J. Rossbach and P. Schmüser. Basic course on accelerator optics. In *CAS - CERN Accelerator School : 5th General Accelerator Physics Course*, pages 17–88. CERN, 1993. https://doi.org/10.5170/CERN-1994-001.17.

[126] S. Baird. Accelerators for pedestrians. Technical Report AB-Note-2007-014, CERN, Geneva, Switzerland, Feb 2007. https://cds.cern.ch/record/1017689.

[127] J. Jowett and M. Lamont. LHC Report: far beyond expectations, Dec 2016. http://cds.cern.ch/record/2240029.

[128] G. Papotti. Lessons from LHC operation in 2015. In *LHC Performance Workshop*, Chamonix, France, January 2016.

[129] V. D. Shiltsev. On performance of high energy particle colliders and other complex scientific systems. *Mod. Phys. Lett. A*, 26(11):761–772, 2011. https://doi.org/10.1142/S0217732311035699.

[130] B. Auchmann, J. Ghini, L. Grob et al. How to survive a UFO attack. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 81 – 86, Evian, France, December 2015. CERN.

[131] G. Iadarola, H. Bartosik et al. Performance limitations from electron cloud in 2015. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 101 – 109, Evian, France, December 2015. CERN.

[132] O. Brüning. Turnaround time in modern hadron colliders & store-length optimization. In *CARE-HHH-APD Workshop proceedings*, pages 34–45. CERN, 2008. https://doi.org/10.5170/CERN-2008-005.34.

[133] M. Pojer. LHC operation. In *8th LHC Operations Evian Workshop*, 2017. https://indico.cern.ch/event/663598/contributions/2781463.

[134] F. Zimmermann, M. Benedict et al. *Future Circular Collider*, volume 4 : The High-Energy LHC (HE-LHC). CERN, 2018.

[135] F. Marcastel. CERN's accelerator complex. Technical Report OPEN-PHO-CHART-2013-001, CERN, Geneva, October 2013.

[136] G.-H. Hemelsoet. Time spent at injection. In *Proceedings of the 2017 Evian workshop on LHC beam operations*, pages 129 – 132, 2019.

[137] A. Lechner, W. Bartmann et al. TDI - past observations and improvements for 2016. In S. Dubourg and B. Goddard, editors, *6th Evian Workshop on LHC beam operation*, pages 123 – 130, Geneva, 2015. CERN.

[138] L. Norderhaug Drøsdal. *LHC Injection Beam Quality During LHC Run I*. PhD thesis, University of Oslo, Oslo, Norway, August 2015.

[139] G. Papotti, T. Bohl et al. The SPS beam quality monitor, from design to operation. In *IPAC2011*, pages 1849 – 1851, San Sebastián, September 2011.

[140] D. Jacquet. Injection. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 49 – 52, Evian, France, December 2015. CERN.

[141] I. Lamas Garcia, N. Biancacci et al. LHC injection protection devices, thermo-mechanical studies through the design phase. In *Proceedings of the 7th International Particle Accelerator Conference*, pages 3698 – 3701, Geneva, 2016. JACoW.

[142] J. Wenninger. Simple models for the integrated luminosity. Technical Report CERN-ATS-Note-2013-033 PERF, CERN, Geneva, Jun 2013. (Internal CERN note) https://cds.cern.ch/record/1553471?ln=en.

[143] M. Benedikt, D. Schulte et al. Optimizing integrated luminosity of future hadron colliders. *Phys. Rev. ST Accel. Beams*, 18:101002, Oct 2015. https://link.aps.org/doi/10.1103/PhysRevSTAB.18.101002.

[144] S. Myers. The engineering needed for particle physics. *Philos. Trans. Royal Soc. Lond. A: Math. Phys. Eng. Sci.*, 370(1973):3887–3923, 2012. https://doi.org/10.1098/rsta.2011.0053.

[145] B. Todd, L. Ponce, et al. LHC Availability 2016: Proton Run. Technical Report CERN-ACC-NOTE-2016-0067, CERN, Geneva, Dec 2016. https://cds.cern.ch/record/2237325?ln=en.

[146] B. Todd, L. Ponce et al. LHC Availability 2017: Proton Run. Technical Report CERN-ACC-NOTE-2017-0063, CERN, Geneva, Dec 2017. https://cds.cern.ch/record/2294852?ln=en.

[147] B. Mikulec and V. Kain. Injectors: Unavailability by machine, root causes, strategy and limitations. In *7th Evian workshop*. 2016. https://indico.cern.ch/event/578001/contributions/2366305/.

[148] P. Lebrun. Large cryogenic helium refrigeration system for the LHC. Technical Report LHC Project Report 629, CERN, Apr 2003. https://cds.cern.ch/record/605468?ln=en.

[149] J.-P. Burnet. Requirements for power converters. In *CAS - CERN Accelerator School: Power Converters*, pages 1 – 14. CERN, 2015. https://doi.org/10.5170/CERN-2015-003.1.

[150] L. Rossi. Manufacturing and testing of accelerator superconducting magnets. In R. Bailey, editor, *CAS - CERN Accelerator School: Course on Superconductivity for Accelerators*, pages 517 – 546, Geneve, 2014. CERN. https://doi.org/10.5170/CERN-2014-005.517.

[151] A. Siemko. What happens during and after a quench ? In *11th Workshop of the LHC*, pages 245 – 249, 2001.

[152] R. Schmidt. Introduction to machine protection. In *2014 Joint International Accelerator School: Beam Loss and Accelerator Protection*, pages 1 – 20, 2016.

[153] E. Ciapala, L. Arnaudon et al. Commissioning of the 400 MHz LHC RF system. In *EPAC 2008 Contributions to the Proceedings*, pages 847 – 849. European Physical Society Accelerator Group, 2008. https://accelconf.web.cern.ch/accelconf/e08/papers/mopp124.pdf.

[154] M.J. Barnes, J. Borburgh et al. Injection and extraction magnets: septa. In *CAS - CERN Accelerator School: Specialized course on Magnets*, pages 167 – 184. CERN, 2010.

[155] V. Schramm, W. Viganó, and C. Zamantzas. System's performances in BI. In *Proceedings of the 2017 Evian workshop on LHC beam operations*, pages 59 – 62, 2019.

[156] M. Stockner. *Beam Loss Calibration Studies for High Energy Proton Accelerators*. PhD thesis, Vienna University of Technology, 2007.

[157] M. Gasior, R. Jones et al. Introduction to beam instrumentation and diagnostics. In *CAS - CERN Accelerator School: Advanced Accelerator Physics Course*, pages 23 – 60. CERN, 2014. http://dx.doi.org/10.5170/CERN-2014-009.23.

[158] L. Ponce, R. Jung et al. *LHC proton beam diagnostics using synchrotron radiation*. CERN, 2004.

[159] O. Brüning, P. Collier et al., editor. *LHC design report: Vol. II The LHC infrastucture and general services*. CERN, 2004.

[160] J. Nielsen and L. Serio. Technical services: Unavailability root causes, strategy and limitations. In *7th Evian Workshop on LHC beam operation*, pages 59 – 62, 2017.

[161] M. Brice. Replacement of the heart of the CMS experiment - the pixel detector. Technical Report CERN-PHOTO-201702-052-74, CERN, Feb 2017. https://cds.cern.ch/record/2253519.

[162] M. Brice and A. Pantelia. CERN Computer Center. Technical Report CERN-CO-1207154-01, CERN, Jul 2012. https://cds.cern.ch/record/1462573.

[163] D. Forkel-Wirth, S. Roesler et al. Radiation protection at CERN. In *CAS - CERN Accelerator School: Course on High Power Hadron Machines*, pages 415–436, 2013.

[164] M. Tammi, V. Vuorela et al. Advanced RCM industry case — Modeling and advanced analytics (ELMAS) for improved availability and cost-efficiency. In *Proceedings of the 10th World Congress on Engineering Asset Management (WCEAM 2015)*, pages 581–589. Springer International Publishing, 2016. http://dx.doi.org/10.1007/978-3-319-27064-7_58.

[165] D. Imran Khan, S. Virtanen et al. Functional failure modes cause-consequence logic suited for mobile robots used at scientific facilities. *Reliab. Eng. Syst. Saf.*, 129:10–18, 2014. http://dx.doi.org/10.1016/j.ress.2014.03.012.

[166] S. Virtanen, J.-P. Penttinen et al. Application of design review to probabilistic risk assessment in a large investment project. In *Proceedings of the Probabilistic Safety Assessment and Management (PSAM) 12 Conference - Volume 9*, pages 200–213, 2016. http://psam12.org/proceedings/paper/paper_319_1.pdf.

[167] E. Rogova. *Reliability assessment of redundant safety systems with degradation*. PhD thesis, Delft University of Technology, 2017. https://doi.org/10.4233/uuid:17606183-f86f-45c4-8333-417dce87392f.

[168] P.E. Dodd and L.W. Massengill. Basic mechanisms and modeling of single-event upset in digital microelectronics. *IEEE Trans. Nucl. Sci.*, 50(3):583 – 602, 2003. https://doi.org/10.1109/TNS.2003.813129.

[169] L. Chadwick and S. Jeffcott. A brief introduction to human factors engineering. *Transfus.*, 53(6):1166 – 1167, 2013. https://doi.org/10.1111/trf.12177.

[170] B. Bradu, E. Rogez et al. Compensation of beam induced effects in LHC cryogenic systems. In *Proc. of International Particle Accelerator Conference (IPAC'16), Busan, Korea, May 8-13, 2016*, pages 1205–1208, 2016. https://doi.org/10.18429/JACoW-IPAC2016-TUPMB048.

[171] S. Claudet, K. Brodzinski et al. Rotating machinery for LHC cryogenics: First analysis of reliability and origins of downtime. *Phys. Procedia*, 67:72 – 76, 2015. https://doi.org/10.1016/j.phpro.2015.06.013.

[172] Safety related terms for advanced nuclear plants. Technical Report IAEA-TECDOC-626, International Atomic Energy Agency, 1991.

[173] L. Tavian. Cryogenics overview. In *FCC week 2016*, 2016. https://indico.cern.ch/event/438866/contributions/1085144.

[174] M. Sapinski, T. Baer et al. Beam induced quenches of LHC magnets. In *Proceedings of the 4th International Particle Accelerator Conference*, pages 3243 – 3245, 2013.

[175] D. Johnson and D. W. Coit et al. System level reliability analyses and predictions in a varying stress environment. In *2013 Proceedings Annual Reliability and Maintainability Symposium (RAMS)*. IEEE, 2013. https://doi.org/10.1109/RAMS.2013.6517617.

[176] G. H. Hemelsoet. Time spent at injection. In *8th LHC Operations Evian Workshop*, 2017. https://indico.cern.ch/event/663598/contributions/2782431.

[177] M. Solfaroli Camillocci. Operational cycle. In *Evian Workshop on LHC Beam Operation*, number CERN-ATS-2013-045, pages 23 – 25, Evian, France, December 2012. CERN.

[178] M. Lamont. How to reach the required availability in HL–LHC era. In *RLIUP: Review of LHC and Injector Upgrade Plans*, number CERN-2014-006, pages 139 – 144, Archamps, France, October 2013. CERN. https://doi.org/10.5170/CERN-2014-006.139.

[179] CERN. Accelerator fault tracking. aft.cern.ch, Accessed 2015. An internal website publishing LHC failure data.

[180] K Cornelis. SPS availability. In *The Workshop on Machine Availability and Dependability for Post–LS1 LHC*, Geneva, Switzerland, November 2013. CERN.

[181] A. Macpherson. LHC availability and performance in 2012. In *Evian Workshop on LHC Beam Operation*, number CERN-ATS-2013-045, pages 1 – 6, Evian, France, December 2012. CERN.

[182] CERN. Accelerator performance and statistics. lhc-statistics.web.cern.ch, Accessed 2016. An internal website publishing LHC operational data.

[183] J. Wenninger. Strategy for run 2. In *LHC Performance Workshop*, pages 129–133, Chamonix, France, September 2014. https://doi.org/10.5170/CERN-2015-002.129.

[184] F. Antoniou, G. Arduini et al. Luminosity modeling for the LHC. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 77 – 80, Evian, France, December 2015. CERN.

[185] M. Pojer. QPS operational aspects. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 29 – 31, Evian, France, December 2015. CERN.

[186] A. Apollonio, B. Auchmann et al. 2015 availability summary. In *Evian Workshop on LHC Beam Operation*, number CERN-ACC-2015-376, pages 19 – 28, Evian, France, December 2015. CERN.

[187] J.C. Helton and F.J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Safe.*, 81(1), 2003. https://doi.org/10.1016/S0951-8320(03)00058-9.

[188] F. Bordry, M. Benedict et al. Amachine parameters and projected luminosity performance of proposed future colliders at CERN. Technical Report CERN-ACC-2018-0037, 2018. https://cds.cern.ch/record/2645151.

[189] M. Schaumann. private communication, 2018.

[190] L. Tavian. private communication, 2017.

[191] J. Wenninger. LHC operation in 2015 and prospects for the future. In *Proc. 51st Rencontres de Moriond, 2016 Electroweak Interactions and Unified Theories*, pages 239–245. ARISF, 2016.

[192] O. Brüning, R. Garcia Alia et al. LHC full energy exploitation study: Operation at 7 TeV. Technical Report CERN-ACC-2017-0086. https://cds.cern.ch/record/2284549.

[193] S. Claudet. CERN procurement strategy. In *Energy for Sustainable Science at Research Infrastructures*, 2015. indico.desy.de/indico/event/11870.

[194] L. Wenz, A. Levermann et al. North–south polarization of european electricity consumption under future warming. *Proc. Natl. Acad. Sci.*, 114(38):E7910–E7918, 2017. https://doi.org/10.1073/pnas.1704339114.

[195] M. Chorowski, H. Correia Rodrigues et al. Towards the conceptual design of the cryogenic system of the Future Circular Collider (FCC). *IOP Conf. Ser. Mater. Sci. Eng.*, 278:012097, 2017. https://doi.org/10.1088/1757-899X/278/1/012097.

[196] F. R. Blánquez, T. Höhn et al. FCC magnet powering concepts and requirements. In *FCC week 2018*. https://indico.cern.ch/event/656491/contributions/2947279.

[197] L. Tavian. Transient modes and their impact on the cryoplant size and operation margins. In *FCC week 2018*. https://indico.cern.ch/event/656491/contributions/2921137.

[198] F. Burkart, W. Bartmann et al. Conceptual design considerations for a 1.3 TeV superconducting SPS (scSPS). In *Proc. International Particle Accelerator Conference (IPAC'17), Copenhagen, Denmark, 14-19 May, 2017*, pages 3323–3326, Geneva, 2017. JACoW. https://doi.org/10.18429/JACoW-IPAC2017-WEPVA033.

[199] B. Dalena, D. Boutin et al. Dynamic aperture at injection and 3.3 TeV energy choice. In *FCC Week 2018*, 2018. https://indico.cern.ch/event/656491/contributions/2923531.

[200] F. Burkart, W. Bartmann et al. FCC injectors. In *FCC-hh general design meeting*, 2017. https://indico.cern.ch/event/625274/contributions/2524839.

[201] B.Goddard, W. Bartmann et al. Possible reuse of the LHC as a 3.3 TeV high energy booster for hadron injection into the FCC-hh. In *Proc. 6th International Particle Accelerator Conference, Richmond, VA, USA*. JACoW, 2015. http://jacow.org/IPAC2015/papers/thpf094.pdf.

[202] F. Burkart. private communication, 2017.

[203] K. Fuchsberger. Turnaround - Analysis and possible improvements. In *7th Evian Workshop on LHC beam operation*, pages 33 – 39, 2017.

[204] J. M. Haight. Automated control systems do they reduce human error and incidents? *Professional Safety*, 52:20 – 27, 2007.

[205] T. Salmi, A. Stenvall et al. Quench protection analysis integrated in the design of dipoles for the future circular collider. *Phys. Rev. Accel. Beams*, 20:032401, 2017. https://doi.org/10.1103/PhysRevAccelBeams.20.032401.

[206] W. Bartmann E. Renner. private communication, 2018.

[207] J. M. Jimenez. Technical challenges and breakthroughs for the FCC-hh. In *Future Circular Collider Study Kickoff Meeting*, 2014.

[208] F. Perez, P. Chiggiato et al. Analysis of beam-induced vacuum effects. Technical Report CERN-ACC-2019-0019, 2019. https://cds.cern.ch/record/2655286.

[209] D. Delikaris. private communication, 2018.

[210] S. Baird and B. Bellesia. Summary of session v: Shutdown modifications 2008/9 and future shutdowns. In *Proceedings of Chamonix 2009 workshop on LHC Performance*, pages 12 – 13. 2009. https://cds.cern.ch/record/1161742.

[211] J. Bartram, R. Bentham et al. Approaches to risk management. In *Legionella and the prevention of legionellosis*, chapter 3, pages 39 – 56. World Health Organization, 2007.

[212] L. Tavian. Transient modes and their impact on the cryoplant size and operation margins. In *FCC week 2018*. 2018. https://indico.cern.ch/event/656491/contributions/2921137/.

[213] T. Nelson and F. Orton. A new approach to congestion pricing in electricity markets: Improving user pays pricing incentives. *Energy Econ.*, 40:1 – 7, 2013. https://doi.org/10.1016/j.eneco.2013.06.005.

[214] A. Ballarino and J.-P. Burnet. Powering the high-luminosity triplets. In *The High Luminosity Large Hadron Collider*, chapter 8, pages 157 – 164. World Scientific, 2015. https://doi.org/10.1142/9789814675475_0008.

[215] M. Di Castro. Robotics inspections, maintenance and early intervention for the FCC. In *FCC week 2018*. 2018. https://indico.cern.ch/event/656491/contributions/2938741/.

[216] M. Brugger, M. Calviani et al. R2e – Experience and outlook. In *4th Evian Workshop on LHC beam operation*, pages 13–18. 2012.

[217] V. Mertens. Civil engineering, infrastructure and operation – cdr status and plan. In *FCC week 2017*. 2017. https://indico.cern.ch/event/556692/contributions/2483412/.

[218] T. Hönh. Advanced power-quality technologies for Future Circular Collider (FCC). In *FCC week 2018*. 2018. https://indico.cern.ch/event/656491/contributions/2915665/.

[219] M. Riedewald, D. Mirek, et al. Dynamic multidimensional data cubes for interactive analysis of massive datasets. In M. Khosrow-Pour, editor, *Encyclopedia of Information Science and Technology*, volume 2, pages 924 – 929. Idea Group Reference, 2005.

[220] P. Kafka. Living PSA-risk monitoring - current use and developments. *Nucl. Eng. Des.*, 175(3):197 – 204, 1997. https://doi.org/10.1016/S0029-5493(97)00037-X.

[221] J. P. Drago, R. J. Borkowski et al. The in-plant reliability data base for nuclear power plant components: Data collection and methodology report. Technical Report NUREG/CR-2641, ORNL/TM-8271, Oak Ridge National laboratory, 1982.

[222] G. Sirin, C. J. J. Paredis et al. A model identity card to support simulation model development process in a collaborative multidisciplinary design environment. *IEEE Syst. J.*, 9(4):1151–1162, 2015. https://doi.org/10.1109/JSYST.2014.2371541.

[223] J. Celko. *Joe Celko's Trees and Hierarchies in SQL for Smarties*. Morgan Kaufmann, 2 edition, 2012. https://doi.org/10.1016/C2010-0-69241-8.

[224] International Electrotechnical Commission. Analysis techniques for dependability - Petri net techniques. International standard IEC 62551:2012, 2012.

[225] A. Rauzy. Guarded transition systems: A new states/events formalism for reliability studies. *Proc. Inst. Mech. Eng., Part O: J. Risk and Reliab.*, 222(4):495–505, 2008. https://doi.org/10.1243/1748006XJRR177.

[226] S. Holzner. *Java 2 black book*, chapter 22, pages 979 – 1022. Coriolis group, Scottsdale, Arizona, 2001.

[227] *LabVIEW^{TM} User Manual*, chapter 21, pages 21–1 – 21–6. National Instruments Corporation, 2003. http://www.ni.com/pdf/manuals/320999e.pdf.

[228] Reliability of nuclear power plants. *Int. At. Agency Bull.*, 17(4):46 – 48, 1975. https://www.iaea.org/publications/magazines/bulletin/17-4.

[229] B. Hahn, editor. *Wind farm data collection and reliability assessment for O&M optimization*. Number 17 in Expert group report on recommended practices. IEA Wind, 2017. https://community.ieawind.org/viewdocument/recommended-practice-17-wind-farm.

[230] WInD-Pool – Business case. Technical report, Fraunhofer IWES, 2016.

[231] T. Pinna, J. Izquierdo et al. Fusion component failure rate database (FCFR-DB). *Fusion Eng. Des.*, 81(8):1391 – 1395, 2006. https://doi.org/10.1016/j.fusengdes.2005.05.011.

[232] C. M. Piaszczyk. Accelerator reliability database. In *Proceedings of the 1999 Particle Accelerator Conference*, volume 2, pages 1465–1467, 1999. https://doi.org/10.1109/PAC.1999.795583.

[233] A. Drago, G. Franchetti et al. *Report on 1st Annual Workshops of all WP6 Tasks*, chapter 4, pages 12–13. ARIES Consortium, 2018. https://edms.cern.ch/ui/file/1817635/1.0/ARIES-Mil-MS26-Final.pdf.

[234] R. Visintini. Power converters for accelerators. In *CAS - CERN Accelerator School: Power Converters*, pages 415–444, 2016. https://doi.org/10.5170/CERN-2015-003.415.

[235] F. Zimmermann, A. Apollonio et al. FCC-ee operation model, availability and performance. In *62nd ICFA Advanced Beam Dynamics Workshop on High Luminosity Circular e+e- Colliders (eeFACT2018)*. Submitted.

[236] B. Desforges and A. Lasseur. 2000 SPS & LEP machine statistics. Technical Report SL-Note-2000-060-OP, CERN. http://cds.cern.ch/record/702628/.

[237] C. W. Allen, W. Colocho et al. PEP-II hardware reliability. In *IEEE Symposium Conference Record Nuclear Science 2004.*, volume 3, pages 1788–1792. https://doi.org/10.1109/NSSMIC.2004.1462590.

[238] J. T. Seeman. Last year of PEP-II b-factory operation. In *Proceedings of EPAC08*, pages 946 – 950. European Physical Society Accelerator Group.

[239] M. Tanaka, H. Asai et al. Operation statistics of KEKB in FY2009. In *Proceedings of the 7th Annual Meeting of Particle Accelerator Society of Japan*, pages 340 – 343, 2010. In Japanese.

[240] Y. Funakoshi. Operational experience with crab cavities at KEKB. In *ICFA Mini-Workshop on Beam-Beam Effects in Hadron Colliders*, CERN yellow report, pages 27 – 36, 2014. https://doi.org/10.5170/CERN-2014-004.27.

[241] V. Mertens. Civil engineering, infrastructure & operation overview. In *FCC Week 2018*. https://indico.cern.ch/event/656491/contributions/2915646/.

[242] The TLEP Design Study Working Group. First look at the physics case of tlep. *J. High Energy Phys.*, 2014(1):164, Jan 2014. https://doi.org/10.1007/JHEP01(2014)164.

[243] E. Bargalló Font. *IFMIF Accelerator Facility RAMI Analyses in the Engineering Design Phase*. PhD thesis, Polytechnic University of Catalonia, Barcelona, February 2014.

[244] J.-P. Signoret, Y. Dutuit et al. Make your Petri nets understandable: Reliability block diagrams driven Petri nets. *Reliab. Eng. Syst. Saf.*, 113:61–75, 2013. https://doi.org/10.1016/j.ress.2012.12.008.

[245] B. Bertsche P. Pozsgai. Conjoint modelling with extended coloured stochastic Petri net and reliability block diagram for system analysis. In *Probabilistic Safety Assessment and Management: PSAM 7 – ESREL '04 June 14–18, 2004, Berlin, Germany*, volume 6, page 1382–1387, London. Springer. https://doi.org/10.1007/978-0-85729-410-4_223.

[246] M. Bozzano, A. Cimatti et al. Safety assessment of AltaRica models via symbolic model checking. *Sci. Comput. Program.*, 98:464 – 483, 2015. https://doi.org/10.1016/j.scico.2014.06.003.

[247] I. Renault, M. Pilliere et al. KB3: computer program for automatic generation of fault trees. In *Annual Reliability and Maintainability. Symposium. 1999 Proceedings*, pages 389–395, 1999. https://doi.org/10.1109/RAMS.1999.744149.

[248] J. Penttilä and V. Ikonen. Dynamic simulation of dependability of nickel reduction plant. In *ASME 2013 International Mechanical Engineering Congress and Exposition*, volume 15, page V015T12A012. https://doi.org/10.1115/imece2013-64308.

[249] Y. Noh, K. Chang et al. Risk-based determination of design pressure of LNG fuel storage tanks based on dynamic process simulation combined with Monte Carlo method. *Reliab. Eng. Syst. Saf.*, 129:76–82, 2014. https://doi.org/10.1016/j.ress.2014.04.018.

[250] E. Zio, P. Baraldi et al. Assessment of the availability of an offshore installation by Monte Carlo simulation. *Int. J. Press. Vessel. Pip.*, 83:312–320, 2006. https://doi.org/10.1016/j.ijpvp.2006.02.010.

[251] B. Bouslah, A. Gharbi et al. Joint production, quality and maintenance control of a two-machine line subject to operation-dependent and quality-dependent failures. *Int. J. Prod. Econ.*, 195:210 – 226, 2018. https://doi.org/10.1016/j.ijpe.2017.10.016.

Appendix A: Gantt charts for maintenance study cases

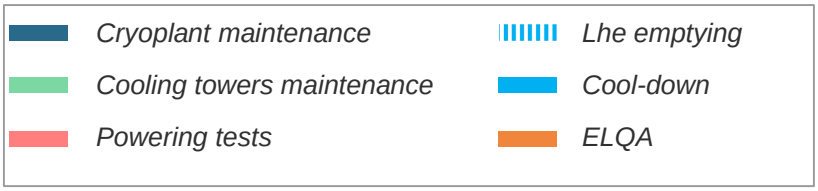

Figure 1: Colour notation for the Gantt charts. Lhe stands for liquid helium and ELQA for electrical quality assurance testing.

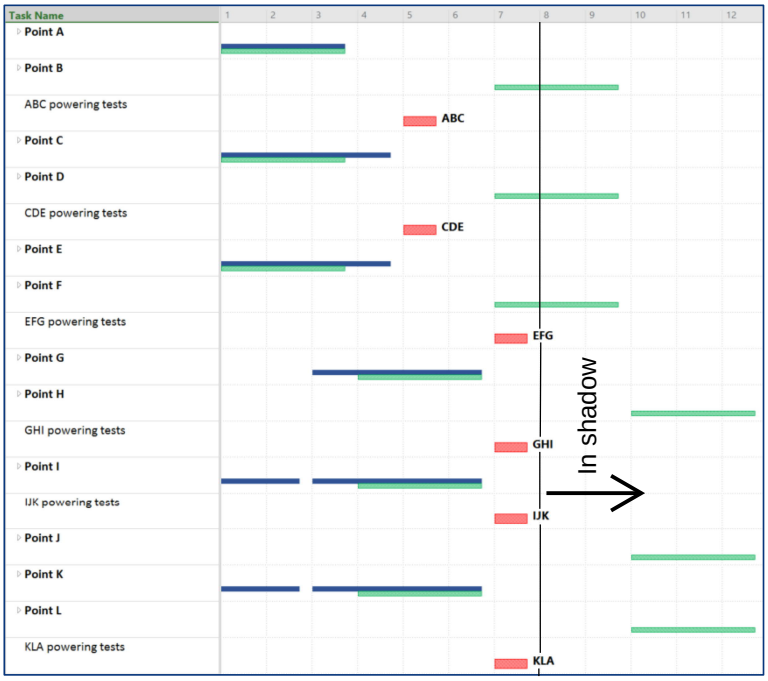

Figure 2: Gantt chart for case 1. After week 7 the maintenance activities do not affect the operations.

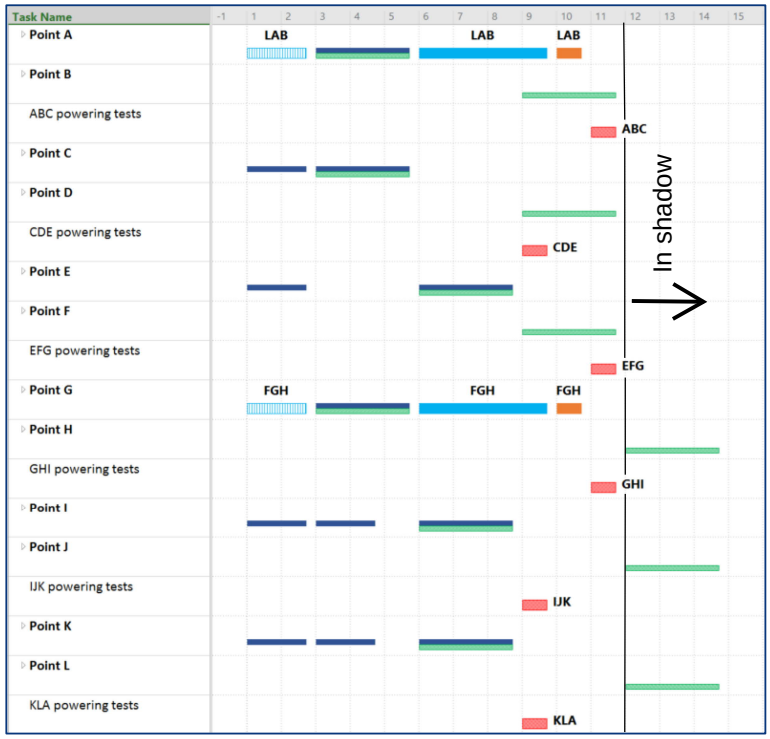

Figure 3: Gantt chart for case 2. Without the connection boxes, sectors at points A and G need to be warmed up during cryogenics maintenance. After week 11 the maintenance activities do not affect the operations.

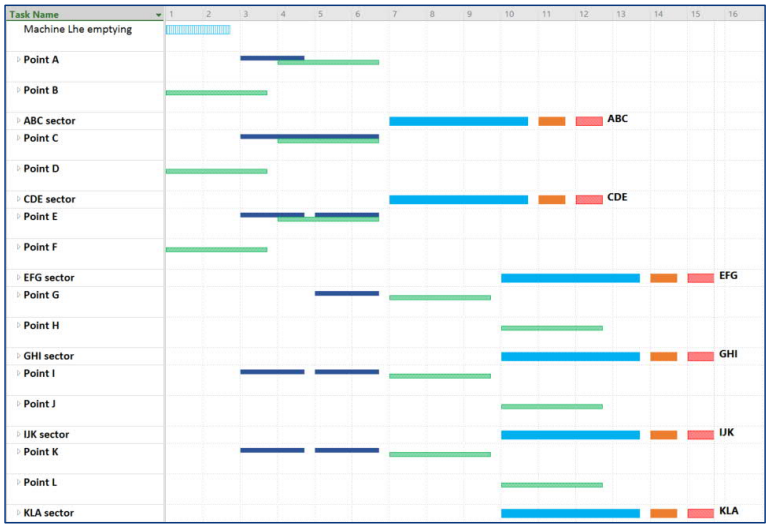

Figure 4: Gantt chart for case 3. All sectors need to be warmed up. Maintenance activities take 15 weeks.