Finding Faults: a scoping study of Fault Diagnostics for Industrial Cyber-Physical Systems

Barry Dowdeswell\textsuperscript{a,*}, Roopak Sinha\textsuperscript{a}, Stephen MacDonell\textsuperscript{a}

\textsuperscript{a}Auckland University of Technology, Auckland, New Zealand

Abstract

Context: As Industrial Cyber-Physical Systems (ICPS) become more connected and widely-distributed, often operating in safety-critical environments, we require innovative approaches to detect and diagnose the faults that occur in them.

Objective: We profile fault identification and diagnosis techniques employed in the aerospace, automotive, and industrial control domains. Each of these sectors has adopted particular methods to meet their differing diagnostic needs. By examining both theoretical presentations as well as case studies from production environments, we present a profile of the current approaches being employed and identify gaps.

Methodology: A scoping study was used to identify and compare fault detection and diagnosis methodologies that are presented in the current literature. We created categories for the different diagnostic approaches via a pilot study and present an analysis of the trends that emerged. We then compared the maturity of these approaches by adapting and using the NASA Technology Readiness Level (TRL) scale.

Results: Fault identification and analysis studies from 127 papers published from 2004 to 2019 reveal a wide diversity of promising techniques, both emerging and in-use. These range from traditional Physics-based Models to Data-Driven Artificial Intelligence (AI) and Knowledge-Based approaches. Hybrid techniques that blend aspects of these three broad categories were also encountered. Predictive diagnostics or prognostics featured prominently across all sectors, along with discussions of techniques including Fault trees, Petri nets and Markov approaches. We also profile some of the techniques that have reached the highest Technology Readiness Levels, showing how these methods are being applied in real-world environments beyond the laboratory.

Conclusions: Our results suggest that the continuing wide use of both Model-Based and Data-Driven AI techniques across all domains, especially when they are used together in hybrid configuration, reflects the complexity of the current ICPS application space. While creating sufficiently-complete models is labour intensive, Model-free AI techniques were evidenced as a viable way of addressing aspects of this challenge, demonstrating the increasing sophistication of current machine learning systems. Connecting ICPS together to share sufficient telemetry to diagnose and manage faults is difficult when the physical environment places demands on ICPS. Despite these challenges, the most mature papers present robust fault diagnosis and analysis techniques which have moved beyond the laboratory and are proving valuable in real-world environments.

Keywords: Industrial Cyber-Physical Systems, Faults, Automotive, Aerospace, Avionics, Industrial Control.
1. Introduction

Industrial Cyber-Physical Systems (ICPS) are mechanisms that augment their computing elements with sensors and electromechanical actuators that allow them to interact with the physical environment they operate in [1]. By evaluating feedback, both from other ICPS they are connected to and from their local industrial environment, they perform a wide range of valuable and often hazardous tasks [2]. Varying widely in complexity and scale, they are found controlling equipment in aircraft, automobiles and factories.

ICPS should be thought of as being more than just computing devices. They form entire systems, viewed as a collection of seamless entities, including their multiple electrical, mechanical and computing subsystems. This homogeneity makes them fundamentally different to the earlier embedded Programmable Logic Controllers (PLCs) that were first used on General Motors automotive assembly lines in the 1960’s [3, 4]. These devices controlled only the machinery they were installed in or embedded in. They were seldom connected to other plant equipment and the sensors they used were often simpler devices such as limit switches, weight sensors or strain gauges. In contrast, modern ICPS act with higher degrees of autonomy than these earlier embedded systems, relying on sensors and actuators that often incorporate their own local data processing and conditioning. ICPS are therefore able to make control decisions based on their perception of their environment, driven by much deeper interaction with the physical characteristics of the world they operate in [5, 6]. Earlier embedded systems seldom featured this degree of complexity and capability.

Contemporary ICPS continue to present intriguing challenges as they have become increasingly more complex. Widely-distributed and now often physically-separated, ICPS are being used to create the Industrial Internet of Things (IIoT), where collections of discrete devices cooperate intelligently to perform large-scale industrial tasks [7]. ICPS differ from Cyber-Physical Systems (CPS) used in consumer or medical devices primarily in terms of their scale [8, 9], security [10, 11] and safety-critically [12, 13]. ICPS used in Smart Grids rely on industry-standard interfaces and sophisticated communications. They manage reliable power distribution across wide geographical areas by cooperating and co-ordinating the operations of the devices that control each substation. Examples of advanced ICPS include NASA’s Mars rover Curiosity which operates semi-autonomously, controlled by one of the most remote ICPS ever deployed on another planet [14, 15].

Detecting and diagnosing ICPS faults quickly and correctly has become imperative to ensure they are fully-operational at all times. We have learnt how to rely on ICPS more and more to manage complicated and often safety-critical tasks. Today, undetected failures in ICPS are not just costly: in safety-critical or hazardous conditions they can be life-threatening [16, 17]. For example, ICPS in the aircraft and aerospace sector rely on accurate readings from sensors to inform guidance, vehicle health and maintain stable flight control. They do this with a degree of reliability, precision and repeatability that human pilots can no longer achieve alone [18, 19]. Similarly, in the automotive sector, vehicles have become increasingly reliant on large local networks of sophisticated subsystems such as anti-skid breaking and fuel-efficient engine controls [20, 21]. Within each subsystem, information is gathered using sensors designed to capture one or more physical characteristics of the local environment, both within and outside the vehicle. The overall operation of a typical ICPS is, therefore, reliant on the co-operative behavior of each of its specialized subsystems, each one dedicated to specific aspects of the vehicle’s safe operation and reliability [22, 23, 24].

1.1. The focus and contributions of this study

We identified, categorized and analyzed fault identification and diagnosis strategies for ICPS employed across the aerospace, automotive and industrial control domains. Our goal was to present a snapshot of fault diagnosis as it is practiced today. We surveyed the differences in the approaches that have emerged in each sector and how they address the needs they describe. Our survey provides a guide to applicable techniques for designers seeking to implement fault identification, diagnosis and management into their ICPS.

We chose the aerospace, automotive and industrial control domains primarily because the ICPS they rely on must operate faultlessly for extended periods of time, often in close proximity to humans [25]. These sectors also exhibit high levels of integration between their computational cyber elements and the sensors that provide the information that all operational decisions are made on. For
example, ICPS in automobiles now sense the position of highway lane markings accurately, extract information from signs and determine the relative positions of adjacent vehicles.

We were also interested in the similarities and differences in fault diagnostic approaches that have emerged in these three safety-critical sectors over the period we studied. The scope of our study was deliberately limited to representative domains that have become highly-dependent on ICPS to manage mission-critical tasks. It is in these sectors that we would expect to find that diagnostics are highly-advanced and widely-used. However, we chose not to include the medical sector in this study. Medical ICPS have distinctive biological characteristics, regulatory requirements and a scale that is worthy of a separate study later. We also excluded cyber-physical devices in the Consumer Electronics sector from our study. They are driving a large and expanding part of the market however they are often less complex than the ICPS in our chosen sectors and the tasks they manage are usually less safety-critical.

A scoping study was used to map the key approaches that underpin fault diagnosis in these sectors and the sources of both theory and case studies available from practitioners [26, 27]. We framed our study via three research questions:

**RQ1:** What are the most common and widely-used fault identification and diagnosis techniques employed in ICPS in the aerospace, automotive, and industrial control domains?

**RQ2:** What relative levels of maturity have the techniques identified in RQ1 achieved when assessed using a systematic scale that is applicable to these domains?

**RQ3:** What research gaps and challenges in ICPS fault identification and diagnosis are being highlighted in the literature surveyed to answer RQ1?

This scoping study seeks to provide a thorough and systematic overview of the fault identification and diagnosis techniques currently in use in our sectors of interest. It profiles the diagnostic approaches we encountered and the techniques that are being used in different situations. By applying a systematic classification to each technique encountered, we are able to estimate the relative level of maturity of each approach, highlighting those which are being applied successfully in real-world environments.

### 1.2. How this paper is organized

Section 2 explores briefly what a ICPS fault is and the terminology used to describe the various stages in a fault management methodology. Section 3 then details the survey data capture and analysis protocol our scoping study employed. While scoping studies do not usually include assessments of the quality of studies uncovered, we chose to adapt and employ the NASA Technology Readiness Level (TRL) as a qualitative scale to compare the relative maturity of the fault diagnosis techniques we encountered [28]. Section 4 presents the results of the scoping study, mapping the fault diagnosis methodologies described in the papers that were included in this study. Finally, Section 5 presents our conclusions, briefly examining those studies that demonstrated the highest TRL. These exemplars discuss fault diagnostic techniques that have moved beyond the laboratory and are being applied in the real world.

### 2. Background - what is fault diagnostics?

ICPS bridge the connection between their “cyber” software, sensor and actuator hardware parts and the “physical” world they inhabit. Figure 1 illustrates the two distinct classes of devices that mediate communication across this divide for a warehouse package-handling robot. A sensor is a device that can convert an environmental characteristic such as proximity, pressure, temperature or light levels into an electrical signal that can be processed by a computer [29]. In contrast, an actuator is a mechanical device that can receive an electrical signal from a computer and cause a change, often as a result of moving something in its environment [30]. Motors are special classes of actuators that create movement, such as the mechanism that moves the package off the parcel tray once the robot has arrived at its destination.

Normal behavior for an ICPS such as this warehouse robot is to pick up packages, navigate reliably and efficiently to another location, and then unload its cargo. The robot’s activities rely on receiving inputs from its sensors and being able to co-ordinate the movements of its actuators to complete tasks that achieve previously-defined goals. Our example
Sonar Sensor to detect walls and other nearby obstacles.

Movement Actuator arm, controlled by a motor, that loads and unloads packages.

Figure 1: Sensors and Actuators for a Warehouse Robotic Package Handler.

robotic package handler has pre-defined patterns of behavior that enable it to traverse warehouse aisles, locate shelves and deliver packages to specified locations. While it is working, it can detect both obstacles and humans, navigating safely around them.

The difficulty inherent in this interaction between the cyber and physical parts of an ICPS often results in faults occurring. Any change in the way that an ICPS operates that leads to unacceptable behavior or degraded performance is defined as a fault [31]. For example, the wheels of the robotic package handler might become entangled with warehouse rubbish from the floor and stop rotating. If the control program detects this problem, it can respond with an appropriate behavior, perhaps stopping and requesting a human for assistance. This sort of situation is not a fault: it is the ICPS managing its behavior in a way that is appropriate. In contrast, not detecting that it cannot move properly and carrying on regardless is a fault since the ICPS did not recognize the issue and change its behavior accordingly. Similarly, failing to detect the edge of stairs and falling down them is unacceptable behavior, possibly due to a faulty precipice sensor. Lee and Seshia comment that it is not enough to separately understand both the computational and electromechanical elements [30]. Rather, it is at the intersection of the cyber and the physical that the most challenging fault scenarios emerge.

2.1. Fault identification, diagnosis and management concepts

Detecting faults is the first stage of a Fault Management Strategy [35]. Detecting a fault should start a multi-step process that attempts to diagnose and potentially correct problems so that the ICPS can resume operating at optimal levels. This implies that the ICPS needs to be able to hold a dynamic representation of what normal behavior is so that it can recognize misbehavior.

Fault management strategies include Fault Isolation, which is the process of accurately identifying the location of the fault and its nature [34]. This can be difficult to determine reliably in large systems that contain many interconnected subsystems. Hence, fault isolation includes the anal-
ysis of multiple possible fault sites to determine the nature of the real, underlying fault. The fault symptoms presented, or Fault Evidence, often include secondary system misbehaviors that are the result of the primary fault but which are not the root cause. Bradatsch [36] defines fault latency as the time between the occurrence of the fault and its recognition by the device’s fault management system.

Once a list of possible fault candidate locations has been identified, the next step is Fault Assessment. This examines evidence and seeks to determine the most likely root locations of the fault, as the problem may include a compound failure located at multiple, distinct points [37, 38]. This leads to the final stage of the diagnosis, Fault Risk Assessment. Not all faults are important enough to require intervention if the system is able to operate satisfactorily in a degraded condition. Gadhab [39] discusses the use of “limp-home” strategies for automobiles that allow them to continue to operate safely in a degraded mode until they can be repaired.

Mortellec et al. [40] provide a wider perspective on what an ideal diagnostic system should provide. Besides being able to uniquely identify the true location and nature of a fault, diagnostic systems must be able to communicate effectively with other systems to help facilitate fault rectification. They must deliver their findings rapidly, especially in safety-critical situations. Finally, it is paramount that they must not report false information.

3. Research method

Scoping studies are one method of rapidly mapping the key concepts that appear within a research area [27, 41]. Often smaller in scope than full systematic reviews or mapping studies, scoping studies allow the breadth of coverage and the depth of the information extracted to be tailored to address research questions appropriately [42, 43]. Arksey and O’Malley [27] and Antman et al. [44] both explain how scoping studies are an appropriate way to quickly capture and present both the available information and the gaps. They can also be used to focus and inform later literature searches for practitioners when they do not have time to perform a thorough initial analysis themselves. Our scoping study protocol follows the four steps of framing research questions, identifying relevant studies, analysis and then presentation of the results as outlined by Arksey and O’Malley [27] and refined by Cacchione [26].

3.1. Step One: Framing our research questions

Scoping studies are also effective where the researchers do not have a single or highly-focused research question that they are seeking to answer [45]. The research questions detailed in Section 1.1 were designed to identify, highlight and categorize practical fault recognition and diagnosis techniques that have been found to be effective both in laboratory studies and in the field. Since this scoping study examines multiple yet similar sectors with potentially differing needs, understanding the focus and spread of the challenges and how they are being addressed should be of interest to practitioners who are designing their own ICPS.

3.2. Step Two: Identification of relevant studies

Scopus was used to search for papers that included the terms “cyber-physical”, “aerospace”, “aircraft”, “automotive”, “industrial” and “manufacturing” for the fifteen-year period from 2004 to 2019. This starting period for the search was chosen since it coincides with the emergence of the term Cyber-Physical System. The first use of the term can be traced to the National Science Foundation meetings in 2001 that discussed networked embedded control systems [46, 47]. In 2006, Lee [48] highlighted the implications of these discussions about connecting discrete embedded systems. Prior to this, Wiener’s earlier pioneering work on cybernetics informed much of the thinking on control systems theory, arguably setting the agenda for later CPS research [49].

3.3. Step Three: Study selection and classification

From an initial pool of 1,700 candidate papers returned by our queries, we performed a pilot study on thirty of these papers. Particular papers were chosen primarily because they contained well-written explanations of fault identification and diagnosis techniques that provided valuable background information. These were used to create an initial set of fault identification or diagnosis approach classifications that identified both broad conceptual differences and a list of specific techniques applicable to those approaches. Table 1 lists these categories. RQ1 asks what the nature of fault identification and diagnostics is within our chosen
domains. The broadest primary classifications that emerged divided the approaches into three high-
level categories that helped to delineate the research
activity. We encountered Physics-Based Model-
Driven diagnostics, Data-Driven Model-Free Arti-
ficial Intelligence (AI) techniques and Knowledge-
Based graph approaches. Hybrid techniques that
blend aspects of these approaches were also en-
countered. The similarities and differences between
these broad classes are profiled in more detail in
Section 4.

To examine the specific fault-finding methods
found within our three primary approaches, sub-
categories were created to identify the characteristics of each technique. Beyond these classifications,
trends such as Predictive diagnostics or _diagnosis_
became of particular interest to us since this ap-
proach featured more widely than we initially ex-
pected. The complete list of studies, classified ac-

Table 1: Fault Identification and Diagnosis Classification
Categories.

<table>
<thead>
<tr>
<th>Category</th>
<th>Example Sub-Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physics-Based Modeling</td>
<td>Kalman Filters, Markov Models, Fault Trees, Other</td>
</tr>
<tr>
<td>approaches</td>
<td>Stochiometric processes, Model Validation/Invalidation,</td>
</tr>
<tr>
<td></td>
<td>Monitor-based oracles</td>
</tr>
<tr>
<td>Data-Driven AI and</td>
<td>Artificial Neural Networks, Machine Learning, Fuzzy</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>Logic, k-Nearest Neighbour, Big</td>
</tr>
<tr>
<td>approaches</td>
<td>Data/Data Mining</td>
</tr>
<tr>
<td>Knowledge-based</td>
<td>Bayesian Decision Theory, Binary Trees, Petri nets,</td>
</tr>
<tr>
<td>approaches</td>
<td>Network Message Analysis, Expert Systems</td>
</tr>
</tbody>
</table>

ies. Arksey and O’Malley stress the importance of
including such “grey matter” in scoping studies.

Our minimum inclusion criteria for a study re-
quired it to present and explain the fault identifica-
tion or diagnosis approach that was being applied.
We also sought papers that included case studies
demonstrating the effectiveness of their techniques.
Many papers were excluded because they only men-
tioned “faults” or “diagnosis” as an aspect of the
nature of ICPS without presenting specific exam-

3.4. Step Four: Analyzing and Presenting the Data

During the first phase of the analysis, the classifi-
cation categories allowed us to perform a thematic
analysis [51, 52]. Each of our categories and sub-
categories represent a technique or approach used
or proposed by a practitioner as a way of identify-
ing, diagnosing or rectifying a fault [53]. The anal-
ysis also included examining where diagnostic re-
search is focused in each sector and is presented in
Figures 5, 6 and 7.

Scoping studies do not usually attempt to assess
the quality of the studies uncovered [26]. However,
we chose to adapt and apply a qualitative scale dur-
ing the classification phase to rank the relative level
of maturity of the diagnostic techniques we found.
Each study was evaluated using the NASA Tech-
nology Readiness Level scale [54, 55]. This is a sys-
tematic metric for assessing how mature a partic-
ular technology is that is now widely used in both
aerospace and defense for technology planning. The
TRL has been progressively refined since the 1980’s
through its use at both NASA, ESA and the US
military [28, 54, 56]. It is now embodied in the
standard ISO 16290 [57]. In 2014, the European As-

sociation of Research & Technology Organisations
(EARTO) identified an increased use of the TRL
amongst its members as a planning tool to manage
innovation [58].

RQ3 sought to identify research gaps, especially
those exhibited amongst the most promising ap-
proaches. The TRL provide criteria for assigning a
classification between TRL 1, representing basic
principles being observed or reported through to
TRL 9, characterized by technologies proven in real
environments that are ready for widespread adop-
tion. We calibrated our fault diagnostic TRL de-
scriptions using the approach of Terrile et al. [59].
They note that the relative TRL steps are not lin-
ear with the steepest steps being in the range TRL
6 to 8. Section 5 details the four divisions we chose
to classify studies into an appropriate range. The granularity of the resulting TRL categories allowed us to distinguish between studies that were purely theoretical and those that were profiling fault diagnostic techniques that are being applied in live environments. Table 5 illustrates the fault diagnostic level characterizations we adapted from the NASA categories.

By the end of the classification and analysis phases we had identified fourteen studies that could be ranked at the highest TRLs between 7 and 9. These report mature, field-proven fault-finding and diagnostic strategies that have been deployed in production environments. In those papers, we should expect to see state-of-the-art exemplars that detail how ICPS respond to and recover from fault situations they encounter.

3.5. Threats to validity

In scoping surveys such as ours, the primary threats to the validity are our choice of which papers to include and our thematic classifications. Surveys are by definition secondary studies that report broad, summarized characteristics of primary studies, the source papers published that present research about an area of interest [60]. As distinct from Systematic Literature Studies (SLS) that provide highly-detailed evaluations of a smaller set of papers [61], scoping studies show where research activity is concentrated and what aspects of a topic are attracting interest, often examining a larger number of papers in less depth.

Internal validity is concerned with the risks that might lead to an incorrect conclusion [62]. This was partially mitigated during the analysis phase by ensuring that each primary paper was initially scanned to determine if it did indeed contain one of our classification classes. For some classes, a list of appropriate synonyms was built iteratively. Our inclusion criteria for a paper included a check to see if groups of related terms were present. The classifications defined in Section 3.3 such as “Model-Based” were expected to show up where models were discussed. However, within the same paper, the classification of “Model-Free” was expected to be applicable when discussions featured AI, Neural Networks, Markov approaches or Data-Driven techniques. Intellectual property restrictions on what can or cannot be published may also be a contributing factor to the amount of detail that can be published about implementations. This was considered when evaluating the relative TRL across sectors.

4. Diagnostic Techniques in Industrial CPS

Examples of fault identification and diagnostic methods examined initially during the pilot study were described by authors as having evolved along three primary pathways: Physics-Based modeling and analysis frameworks, Data-driven or Model-free AI techniques, and Knowledge-Based graphical approaches [63]. While classifying our studies, we also identified hybrid approaches which blend aspects of these methods.

4.1. Physics-Based, Model-Driven Diagnostics

Modeling is used by designers to gain a deeper understanding of a system. By creating models that imitate the physical characteristics of the ICPS components, they can explore the interaction the sensors and other physical devices have with the cyber parts of the ICPS [30]. Physics-Based Modeling techniques for diagnostics rely on consistency checks against these models. These detect the differences between the telemetry captured from the live ICPS and the values predicted by the model. Table 2 summarizes physics-based modeling diagnostic techniques across our survey domains.

Consistency checks use data captured by observers who filter the individual readings to distinguish between noise caused by telemetry errors and values that indicate faulty behavior [118]. These differences will often be small but seldom non-zero when the ICPS is performing within acceptable tolerances [21]. Techniques for determining when an aspect of a model is invalidated were discussed in 48% of papers, especially in the industrial control domain. Both Kalman Filters and Markov Models were discussed as ways of recognizing model invalidation. These techniques implement observers that can process sequential measurements that vary over time. Kalman Filters are more applicable when the range of possible readings is highly-linear. They apply recursive algorithms where weighted-averages are used to estimate the next value. They work well in noisy environments that produce sequences of unreliable readings. Zolghadri et al. describe an implementation of a Kalman filter to detect jamming of a flight control surface by filtering the error signal before it is processed by the on-board avionics [68]. The authors explain how the number of sensors providing input to the model affects
both the design and worst-case performance. Tuning the model parameters requires trade-offs against the real-time capacities of the diagnostic systems that rely on the model. Shraim et al. discuss fault management for quadrotor unmanned vehicles to improve rotor positioning accuracy [23]. Unmanned Aerial Vehicles (UAV) require real-time fault tolerance since they now rely on autonomous, sensor-driven stability control that is no longer managed entirely by the pilot. The models used have to take into account the complex aerodynamic characteristics of the UAV. Dearden et al. discuss similar aspects of autonomous operation, describing fault diagnostics for Mars Rovers where Kalman filtering provides situational awareness to indicate fault conditions [75]. They contrast the number of sensors required to manage rover operations with the low computational power available to perform fault identification using multiple sub-system models.

In contrast, Markov models are used to model non-linear, randomly changing systems with discrete states. A dynamic model is Markov or has the Markov Property if the future state of a system depends only on a limited number of previous states. Markov Chain and Markov Decision processes rely on observing the full set of values or states for the aspect of the ICPS that is being diagnosed. In contrast Hidden Markov Models operate where the sequential state of a system is not fully observable. Kunst et al. profile damage propagation through ICPS using Hidden Markov models [19]. Similarly, Windmann and Niggeman [65] and Ribero et al. both apply Markov Models to monitor industrial processes and identify faults as they propagate.

Fault Trees are a way of modeling all reasonably-probable fault scenarios [22]. They are tree structures that facilitate a top-down, systematic approach to identify chains of possible faults. Logical operators can be applied to nodes to identify likely fault pathways. Fault trees are usually considered to be knowledge-based approaches but they were most often encountered in studies that employed hybrid approaches. Mohre et al. demonstrate correlations between fault tree nodes and compositional safety analysis models [12]. Kassmeyer et al. apply fault trees to track fault scenarios across multiple automotive feature variants [86].

Across all sectors, a wide range of specialized Model-Invalidation approaches were encountered, both theoretical and in-use. Provan [88] discusses how acceptable inputs can be modeled, an important pre-requisite to detecting misbehavior. Monitors [117] are code within a fault identification system that is responsible for detecting anomalous situations or behavior. Similarly, Monitor-Based Ora- cles provide ways of both capturing and evaluating possible fault occurrence [8, 113, 100].

Formal modeling languages including the Architecture Analysis & Design Language (AADL) [119] and Modeling and Analysis of Real-time and Embedded Systems (MARTE) [120] model ICPS during their design phases. AADL originated in the aerospace sector to model embedded systems and has now found wide use in the automotive domain. MARTE extends the UML to provide similar capabilities. Huang et al. [121] describe a simulation platform modeled in AADL that allows transient faults to be evaluated. Khelif and Shawky demonstrate how to use AADL to design co-simulations that are easier to diagnose later [102]. Shulte proposes a state machine architecture for fault detection based on SysML [22]. However, no papers in the survey discussed production ICPS implementations that employed either AADL or MARTE mod-
els from the design phases directly. Procter and Feiler present an introduction to the AADL EMV2 Error Library where they discuss the use of an error ontology during modeling [122]. We searched the literature for examples of the use of EMV2 in production fault diagnostic systems beyond the design phase but found few applicable examples. Lu et al. discuss redundancy approaches using AADL and EMV2 however their work does not demonstrate how to apply their fault trees in a production, real-world example [123]. Similarly, Zhang et al. discuss the design of fault tolerant systems using EMV2, but it is applicable only to early-stage modeling [124].

Creating and maintaining models is labor-intensive. Many of the techniques rely on detecting situations where a model is invalidated. However, Milis et al. [33] highlight the amount of effort needed to calibrate models. Provan [88] also discusses two practical impediments to effective model-based diagnosis: the failure to integrate diagnostic modeling early enough in the requirements process and ambiguities in the models themselves at run-time.

4.2. Data-Driven fault diagnostics

Data-Driven diagnostic techniques employ training and learning to forge a representation of the system’s behavior [21]. Unlike Physics-Based models, Data-Driven fault detection does not rely on the existence of pre-built models. This approach is preferred when the ICPS can provide telemetry that contains enough information to distinguish between either normal or degraded operations. AI fault diagnosters make sense of that information by using discriminating logic that copes with the changes seen in the ICPS as they occur. This ability to make intelligent decisions distinguishes AI from machine learning, which involves ICPS learning without being explicitly programmed. Milis [33] discusses cognitive agents that apply expert reasoning to mimic the behavior of human experts.

Artificial Neural Networks (ANN) [131, 11, 127] and pattern-recognition algorithms [144] are illustrative of data-driven techniques. Since they do not rely on static, pre-built models as reference points, they remove the need to keep the model up-to-date as the system evolves. Data-Driven diagnostic systems learn behaviors through training. Detection logic allows them to compare current values with previously learnt values [101]. Hence these Model-Free methods do not have to completely understand the underlying architecture of the system being examined [132].

Data-Driven approaches often scale better than Model-Based techniques [9, 8]. As long as sufficient computational resources are available, Data-Driven techniques work as effectively with a large number of sensors as they do with a few [132]. Since they construct knowledge representations dynamically, they are often easier to update than formal models [133].

Unlike Model-Based methods, Data-Driven approaches do not assume the probabilistic distributions of sampled values that Markov processes rely on [9]. Similarly, AI methods, including machine learning, do not rely on processes being stochastic or random. The trade-off is that while Physics-Based models are labor-intensive to create, model-free techniques require large example data sets to train the observers [125, 126, 127]. Iverson et al. [132] explain that for avionic ICPS, large volumes of archival sampled values are collected during routine operations that are can be used for training neural networks.

Fuzzy logic employs truth values that are real numbers between zero and one rather than being boolean [13, 38, 147]. This allows decisions to be made about non-numerical or imprecise data from ICPS, stored in structures called fuzzy sets. These sets represent partial truths and decisions are made by arriving at a consensus. Fuzzy logic algorithms are able to re-evaluate thresholds for situations where values are expected to change dynamically as the system is being observed. Song [148] discusses recognizing faults using threshold predictions. Each sampled value is checked to see if it falls within a range defined by the previous value read.

Condition monitoring allows Data-Driven fault observers to obtain real-time data about the ICPS they are monitoring. These data points replace the reference values that pre-built Model-Based solutions rely on since AI and machine-learning approaches are model-free [9]. Lee et al. [2] and Fleischmann et al. [152] describe these techniques in terms of system health monitoring. Where deviations from the norm are observed, the result is similar to the model-invalidation discussed earlier. Wang et al. discuss this in the context of cloud computing and predictive maintenance [89].

Wang et al. [153] caution against over-reliance on AI approaches. They suggest that given the complexity of some fault scenarios, the conclusions
Table 3: Data-driven A.I. Model-Free fault identification and diagnosis techniques across all sectors.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Aerospace</th>
<th>Automotive</th>
<th>Industrial</th>
<th>All</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Artificial Neural Networks</td>
<td>54%</td>
<td>50%</td>
<td>47%</td>
<td>50%</td>
<td>[33] [11] [125] [126] [127] [21] [128]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[129] [130] [131] [128] [132] [133] [134]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[135] [136] [137] [138] [139] [140] [141]</td>
</tr>
<tr>
<td>Machine Learning</td>
<td>38%</td>
<td>17%</td>
<td>18%</td>
<td>24%</td>
<td>[9] [132] [24] [21] [133] [134] [143]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>[144] [145] [146] [139]</td>
</tr>
<tr>
<td>Fuzzy Logic</td>
<td>8%</td>
<td>25%</td>
<td>18%</td>
<td>17%</td>
<td>[13] [38] [147] [148] [138] [135] [149]</td>
</tr>
<tr>
<td>Big Data</td>
<td>8%</td>
<td>0%</td>
<td>18%</td>
<td>10%</td>
<td>[150] [151]</td>
</tr>
<tr>
<td>Condition Monitoring</td>
<td>8%</td>
<td>8%</td>
<td>24%</td>
<td>14%</td>
<td>[9] [2] [152] [89]</td>
</tr>
</tbody>
</table>

drawn by data-driven systems may not be sufficiently robust enough to be free of false positives and negatives. However, Iverson et al. profile fault finding for the International Space Station (ISS), reporting that when a large amount of nominal data is available, Data-Driven systems can become highly effective at detecting anomalies [132].

4.3. Knowledge-Based approaches

Knowledge-Based approaches are applicable where large amounts of historical data are available. The underlying dependencies that define the system are derived from these sources using a range of techniques. All fault diagnosis systems need to observe real-time data, basing their evaluations on either qualitative or quantitative aspects of the telemetry. However, only knowledge-based approaches utilize significant amounts of historical data to inform their classifiers [2]. Unlike Data-Driven AI approaches, Knowledge-Based methods do not require pre-classified training sets. Rather, they mine the historical data using statistical methods. Chen et al. explain the value of historical information gathered from experts in building knowledge bases to inform current fault diagnoses [38].

The resultant dynamic models they construct are represented using dependency graphs. Petri nets are directed bipartite graphs where nodes represent discrete fault events that may occur. The graph arcs define possible transitions between states [158, 159].

Bayesian Belief Networks are knowledge-based directed graphs that model probabilities [38, 154]. Each node represents a step in a cause and effect chain with a conditional probability. While observing, the fault system updates the probability at a node when new information is available. Hence, Bayesian networks can provide both diagnostic and predictive evaluations.

Binary Decision Diagrams are directed acyclic graphs. Waszecki et al. [156] encode observation patterns extracted from messages exchanged by automotive ECUs to capture fault scenarios that can be evaluated during diagnosis. Network message analysis also complements other knowledge-based approaches, either as a carrier of fault messages or as an indicator of misbehavior [148]. Schweppe et al. [160] discuss the Automotive Keyword Protocol ISO 14230:2000 [161], a widely-accepted standard for analysing faults via network messages exchanged over a vehicles CAN bus. Pons et al. [157] outline a similar approach using Causal Graphs rather than Binary Decision Diagrams.

4.4. Hybrid fault diagnostic approaches

Hybrid approaches that blend techniques from any of the three broad approaches were encountered in 14% of the papers but featured in 19% of all industrial control studies. Hybrid techniques skewed the overall ratios of our three primary categories since practitioners can adopt any combination of methods to create their fault identification and diagnostic methodologies. Figure 3 illustrates the spread of Hybrid approaches across our three domains. Lee et al. [89] employs Model Validation from the Physics-Based Modeling category with Condition Monitoring from the Data-Driven AI category in an intelligent manufacturing scenario. This allows their system to analyze and predict faults from patterns shared via a cloud-based system. The system is implemented using intelligent agents. Chen et al. [38] combine Bayesian networks with Fuzzy Logic to diagnose faults in automotive braking systems while Banerjee et al. [135] profiles a system with an amalgam of Fuzzy Logic Data-Driven predictors and Model-Based statistical data.
Table 4: Knowledge-Based fault identification and diagnosis techniques across all sectors.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Aerospace</th>
<th>Automotive</th>
<th>Industrial</th>
<th>All</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian Networks</td>
<td>0%</td>
<td>17%</td>
<td>43%</td>
<td>31%</td>
<td>[91] [38] [154] [155]</td>
</tr>
<tr>
<td>Binary Decision Trees</td>
<td>0%</td>
<td>33%</td>
<td>0%</td>
<td>15%</td>
<td>[156] [157]</td>
</tr>
<tr>
<td>Petri nets</td>
<td>0%</td>
<td>0%</td>
<td>43%</td>
<td>23%</td>
<td>[158] [159]</td>
</tr>
<tr>
<td>Network Message Analysis</td>
<td>0%</td>
<td>10%</td>
<td>29%</td>
<td>62%</td>
<td>[156] [160] [148] [143]</td>
</tr>
</tbody>
</table>

Using multiple approaches in this way allows practitioners to apply the most appropriate technique to different aspects of an ICPS. Rizzoni et al [162] discuss how both model-based and neural network techniques facilitated the development of on-board diagnostics and fault monitoring to measure vehicle emissions in automobiles. They trace the motivation for continuously assessing emission compliance in each vehicle back to the California Air Resource Board (CARB) requirements that came into force in 1991. Each vehicle is required to monitor its own emissions to ensure compliance. That required neural network approaches to facilitate the tasks of data capture and sensor filtering followed by model invalidation to test compliance.

4.5. Predictive Diagnostic techniques

Predictive Diagnostics or Prognostics is the ability to detect the signs of an impending fault before a failure occurs and to estimate when it might happen [136]. Figure 3 suggests that the ability to predict ICPS faults in advance is of interest in all three domains. Predictive Diagnostics becomes feasible when it is possible to both capture and process large amounts of high-fidelity data about the operation of an ICPS and recognize the fault symptoms in advance. Janasak and Beshears [163] state that one aim of European air carriers is that by 2050, all flights should arrive within one minute of their scheduled time. Current delays and disruptions can be up to fifteen minutes due to undiagnosed faults, an issue that better predictive capabilities might alleviate.

4.6. Overall trends in the data

In each sector, there is an emphasis on the development of smart sensors and the conditioning of the sensor data using a range of techniques such as Kalman Filters or Markov models. Coupled with that, the representation of ideal values or behavior was described using either models or dynamically using AI data-mining techniques. Once a definition of what is normal can be determined, deviations from expected values or behaviors can be detected. Artificial Neural Networks and Machine Learning were evidenced as alternatives to Model Invalidation in the Data-Driven AI category. However, the widespread use of hybrid techniques in different parts of the ICPS reflects the complexity of the systems being profiled: no single technique for fault recognition and analysis predominates or is sufficient for all needs. The predominance of Data-Driven techniques in aerospace is in contrast to the lack of evidence for the use of Knowledge-based approaches in that sector while Network Message Analysis was a technique profiled in 29% of the industrial studies that employed Knowledge-based approaches. Those contrasts are explored more deeply in Section 5 where we examine the most mature techniques in more detail.

5. Investigating mature fault diagnostic techniques

RQ2 asked what levels of maturity the diagnostic techniques adopted in each sector have achieved.
Table 5: Adapting the NASA Technology Readiness Levels for Assessing Fault Diagnosis.

<table>
<thead>
<tr>
<th>TRL</th>
<th>NASA categorization</th>
<th>Proposed Fault categorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Actual system “flight proven” through successful mission operations.</td>
<td>Actual fault diagnostic system proven through successful identification and classification of real faults in a production environment.</td>
</tr>
<tr>
<td>8</td>
<td>Actual system completed and “flight qualified” through test and demonstration (ground or space).</td>
<td>Actual fault diagnostic system qualified through test and demonstration in a production environment.</td>
</tr>
<tr>
<td>7</td>
<td>System prototype demonstration in a space environment.</td>
<td>Functioning prototype demonstrated in a production environment.</td>
</tr>
<tr>
<td>6</td>
<td>System/subsystem model or prototype demonstration in a relevant environment (ground or space).</td>
<td>Functioning prototype demonstrated finding and/or diagnosing faults in a relevant environment beyond the laboratory.</td>
</tr>
<tr>
<td>5</td>
<td>Component and/or breadboard validation in a relevant environment.</td>
<td>Creation of a breadboard and/or software validation that can search for and/or identify faults in a relevant environment.</td>
</tr>
<tr>
<td>4</td>
<td>Component and/or breadboard validation in a laboratory environment.</td>
<td>Creation of a breadboard and/or software validation that can search for and/or identify faults in a laboratory environment.</td>
</tr>
<tr>
<td>3</td>
<td>Analytical and experimental critical function and/or characteristic proof-of-concept.</td>
<td>Proof-of-concept experiment with an appropriate simulation of the fault environment.</td>
</tr>
<tr>
<td>2</td>
<td>Technology concept and/or application formulated.</td>
<td>Concept and technology to perform detection and/or diagnosis proposed, including a mathematical formulation.</td>
</tr>
<tr>
<td>1</td>
<td>Basic principles observed and reported.</td>
<td>Basic fault detection or diagnosis principles observed and reported.</td>
</tr>
</tbody>
</table>

The TRL fault classifications we developed for our study are shown in Table 5 in parallel with the matching NASA descriptions.

Mankins [55] explains that each level in the TRL scale represents a different maturation of the technology or methodology. Heder [164] notes that the TRL has drawn criticism for its use outside of the environment it was originally designed for, explaining that in the European Union the approach has not always been tailored properly for specific disciplines. However in NASA the concept of “flight-readiness” was already deeply ingrained in their culture [165]. Adapting this concept to machinery to establish what stage of technological readiness it has reached was a natural step within their context. We considered this when designing our study, carefully crafting our adaptations of the individual level descriptions to ensure we stayed true to the intent of the TRL.

Assessing the maturity of a technological approach requires a careful evaluation of the context that it is being trialled or applied in. Our TRL categories are divided into four distinct groups. Studies classified as TRL 7 to 9 represent the most mature implementations. They provide a fascinating glimpse of techniques which are either close to or fully operational in live production environments.

Studies from TRL 5 and 6 provide evaluations from trials performed in highly-realistic environments beyond the laboratory. They often use case studies to illustrate how the diagnostics will work in particular situations. In contrast, studies at TRL 3 and 4 present functioning prototypes that are being evaluated in either a laboratory or simulated environment.

Levels 1 and 2 categorize fault identification and diagnosis techniques that are purely theoretical or are presented with a formal mathematical treatment. Papers at this level do not report concrete outcomes from case studies or field trials.

Figure 4 highlights the TRL maturity levels we observed across all our domains of interest. Amongst the survey papers are studies of fault identification and diagnostic techniques that have moved beyond the laboratory and are being applied in real-world environments. In these papers, we should expect to see state-of-the-art exemplars that detail how ICPS respond to and recover from fault situations they encounter. The papers we classified at TRL 7 and above present evaluations of how well
Adapted Technology Readiness Level (TRL)
Percentage of studies at that TRL
35%
30%
25%
20%
15%
10%
5%
0%
1 2 3 4 5 6 7 8 9
Aerospace
Automotive
Industrial

Figure 4: Technology Readiness Level by Sector

Benowitz [117] profiles the Fault Protection Engine currently used by the Mars Curiosity rover. Since the rover is too far away to rely on external systems for assistance, the fault protection engine has to proactively manage faults within a large number of interrelated subsystems autonomously.

Earlier rover designs implemented discrete fault management within each subsystem. On Curiosity, the architecture implements monitors, code within each module whose responsibility is to recognize anomalous behavior. Each module has specific knowledge of the subsystem they are operating within that informs their judgments while filtering sensor readings. Monitors signal problems by raising an error flag. As well as detecting faults, they maintain a count of the occurrences that is later used by the fault protection engine to ascertain how persistent or serious the fault is.

Benowitz explains that error flags are latched but never cleared by the ICPS module-level monitors. This allows the fault protection engine to manage the overall health of the rover by polling in its own time, making decisions without being flooded by messages from subsystems. The fault engine maintains a model that contains a response that is appropriate to each situation the monitors are signalling. Curiosity has over 1,000 monitors operating at any one time. Since the rover may be performing any number of different tasks at any time, ranging from landing to exploration, fault management has to be contextual.

Curiosity’s Model-Based approach is in contrast to the hybrid Model-Based and Data-Driven approach employed by Zolghadri et al. [68]. They profile the flight surface control systems they developed for the Airbus A380. Like Curiosity, their fault management is situation-aware. They note...
that fault signatures are often difficult to detect when an aircraft is parked or taxiing, or when the data rates from sensors are low. Their approach calculates residuals, the result obtained by comparing the current servo positions with the estimated position predicted by the model. They tune the sensitivity of Kalman filters to establish a trade-off between reliably detecting signals and robustness with respect to normal environmental variations.

Azam et al. [133] take a similar approach using neural networks to dynamically model and monitor fifty flight parameters. They discuss the difficulty of using model-based approaches that cannot manage the complexity of accommodating all reasonable parameters in all flight modes. Their data-driven approach also provides estimates of fault severity.

Iverson et al. [132] and Schwabacher et al. [126, 129, 166] provide a highly detailed treatment of the hybrid fault monitoring system certified by NASA for International Space Station (ISS) operations and for Ares I-X launch pre-diagnostics. The Inductive Monitoring System (IMS) is a ground-based ICPS that processes telemetry from the ISS in near real-time. It relies on rule-based, Model-Based and Data Driven algorithms in three distinct subsystems of the IMS. They employ a clustering approach from a fixed number of training points, an approach that allows them to rapidly tailor IMS for new situations. Schwabacher et al. note that there is a need for mission-critical systems such as these to be flight-certified since ground controllers rely on them to make go/no go decisions about launches. They note that many Space Shuttle launches were delayed due to unreliable fault diagnoses. When launch faults can be evaluated more rapidly, redundant or hot-swappable modules can be deployed to reactivate launch sequences to meet critical time windows.

Studies such as these help to explain the proliferation of hybrid techniques encountered. In aerospace, 54% used Artificial Neural Networks and 38% employed Machine Learning, coupled with a range of Model Invalidation methods that were discussed in 43% of all aerospace studies.

5.2. Studies from the Automotive sector

The automotive ecosystem is built up of millions of discrete, complex and mostly unconnected ICPS. Each vehicle operates as a self-contained network of co-operating subsystems. Stout’s Automotive Defect and Recall Report shows that in 2018, nearly eight million vehicles were recalled in the US to address software-based defects [167]. That total is higher than all the recalls for software issues in the previous five years. Figure 6 highlights where diagnostic research is focused in the automotive sector.

Modern vehicles feature up to 120 embedded ECUs, connected by five or more system buses [168, 169]. Sarecco highlights how large and complex the software is currently in vehicles, reporting that the 2017 Ford 150 pickup requires 150 million lines of code [170]. Charet [171] contrasts this with the F-35 Joint Strike Fighter that required only 5.7 million lines of code while the Boeing 787 Dreamliner uses only 6.5 million lines.

This complexity is reflected in the automotive survey papers at the highest TRL. Nasri et al [172] explain that the increasing sophistication of in-car electronics, including Adaptive Cruise Control, Lane Detection and Light Detection And Ranging (LIDAR) technologies, leads to more intricate fault scenarios. They detail the implementation of diagnostics that analyse messages flowing between subsystems on the vehicles Controller Area Network (CAN). Many of the current diagnostic tools rely on proprietary software from vendors that are not easy to integrate into system-wide diagnostic frameworks. They detail their implementation of a hierarchical chain of localised diagnosers that are monitored by a single global fault analyser. A Directed Graph approach is used to identify faults, capturing CAN messages via hardware-in-loop connections.

The scope of what is deemed a “safety-critical” component in the automotive sector is also changing. In May 2018, back-up cameras became mandatory on US vehicles, transforming an optional luxury item into something that required
much more rigorous quality control and deeper vehicle integration [167].

Over-the-Air (OTA) access to diagnostic data from automobiles is profiled as one route to addressing the difficulty of fault-finding in disconnected automotive ICPS. The global remote diagnostics market is forecasted to grow at 17% annually over the next five years, driven primarily by the potential operational cost savings to automakers [173]. Steinkamp et al. describe General Motors new OTA system which is capable of handling 4.5 TB of data per hour from vehicles [167].

However, Dragojevic et al. [174] identify remote access to diagnostic data from a vehicle as a significant technical challenge. Traditional automotive architectures featured highly-specialized ECUs that were optimized for minimal functionality to balance safety concerns. Full operating systems for vehicles emerged though middleware such as Adaptive AUTOSAR [175], leading to greater opportunities to aggregate diagnostic data that could be shared with remote fault analysis systems. Without functionality such as OTA, remote vehicle diagnostics cannot be performed in an IoT ecosystem. Dragojevic et al. profile their work on an OTA bridge solution that connects with the on-board vehicle network. However, they note that Adaptive AUTOSAR needs to encompass safety aspects to certifiable levels before it can be widely deployed.

Kane, Fuhrman and Koopman detail the use of runtime monitor-based oracles that mine the data used by OTA systems for fault finding [114]. Runtime monitors analyze system traces to see if they conform to acceptable behavior patterns. They tune their oracles using large amounts of previously captured telemetry and describe methods used during live vehicle trials. Since monitors operate as hardware-in-loop devices and often interact with safety-critical components, they have to be designed as high-integrity devices. They address this by creating isolated monitors with well-defined interfaces.

5.3. Studies from Manufacturing and Control

Unlike the automotive and aerospace sector, most industrial systems are stationary in one location and are therefore easier to connect into factory-wide monitoring systems. Industrial production machinery therefore offers numerous opportunities to perform local or remote diagnostics. Ramos et al. cite maintenance costs of up to 60% of the production costs as a key driver for factory diagnostics and prognostics [101]. Figure 7 highlights where diagnostic research is focused in the manufacturing and industrial control sector.

International, industry-wide initiatives foster standardization across this sector. Chen et al. [91] discuss trials of sensors for gearboxes in the context of manufacturing initiatives such as the Machinery Information Management Open Systems Alliance (MIMOSA) [176]. Lee, Jin and Bagheri [2] discuss Industry 4.0 and Big Data as similar driver of standardization. Their approach demonstrates end-to-end factory machinery feeding sensor data into multiple analytical systems for near real-time fault identification and prediction. They employ deep-learning for Data-Driven prognostics.

Ramos et al. [101] also profile Service Oriented Architectures to expose fault-finding services at multiple factory levels. Their case study focuses on self-recovering machinery that is supported by the factory infrastructure using hardware-in-loop techniques. Manufacturing is typically managed by multi-layer IT infrastructures that connect higher-level Enterprise Resource Planning (ERP) through layers down to factory automation systems such a SCADA. Ramos et al. profile their eSonia system which manages assets on multiple levels. Many production operations require assembly lines to be able to be re-configured dynamically to suit changes in demand. This requires a degree of self-awareness from plant equipment, which must be able to signal if it is available when changes are requested.
### Table 6: Recurring themes across sectors.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Aerospace</th>
<th>Automotive</th>
<th>Industrial Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Existing degree of connectivity?</td>
<td>Low.</td>
<td>Becoming more connected.</td>
<td>Already highly connected.</td>
</tr>
<tr>
<td>Difficulty of becoming more connected?</td>
<td>Hard due to distance and low bandwidth.</td>
<td>Hard due to large number of discrete vehicle instances.</td>
<td>Already highly-connected due to high degree of localization.</td>
</tr>
<tr>
<td>Amount of diagnostic data available?</td>
<td>High.</td>
<td>Becoming very high.</td>
<td>Already very high.</td>
</tr>
<tr>
<td>Need for autonomous operation?</td>
<td>Very high in remote planetary rovers and drones</td>
<td>Very high due to cost of local fault repair.</td>
<td>Already established via predictive diagnostics and self-management.</td>
</tr>
</tbody>
</table>

### 6. Conclusions, gaps and future work

This scoping study was written with a view to providing an overview of mature fault identification and diagnosis techniques for practitioners who are seeking to understand the state of the current practice and who are creating ICPS. The wide use of Model-Based (62%) and alternative Data-Driven AI (33%) techniques across the aerospace, automotive and industrial control domains reflects the complexity of the current ICPS application space.

As the number of interconnected ICPS increases along with the intricacy of the tasks they manage, the use of Model-Based approaches alone was often profiled as becoming intractable. Milis [33] discussed the difficulties of calibration to align with real systems. Scalability of models was also discussed in this context but only Yen et al [8] discussed partial models, a technique for segmenting models into sub-models. No studies profiled Digital Twins as a solution. Model-Based diagnosis remains a viable strategy, yet how we create complete-enough or partial models quickly and reliably remains a challenge. The AADL EV2 Error Annex has potential to be used beyond the early modeling stages however we found no evidence of its use in the field.

Model-Free AI approaches were evidenced as a viable way of addressing this challenge, demonstrating the increasing sophistication of current machine learning systems. However, there was no discussion of explainable AI, where the decisions made by algorithms could be justified.

The proliferation of hybrid fault systems that blend different aspects and techniques reached 19% in the industrial control sector, indicating the importance of further research into multiple-method solutions, where models are tuned by real-time data. Design-for-Certification was highlighted as a significant driver to ensure products could be deployed beyond the laboratory [174, 126, 151].

Predictive diagnostics is a promising area that was often discussed in-context with the ability to mine sensor data with enough granularity to allow faults to be predicted. Predictive techniques were prevalent in 30% of all industrial control studies, driven by the availability of large amounts of local data. Further research to develop remote connectivity in the aerospace and automotive sectors should lead to more powerful predictive capabilities. However, the potential volume of the data available from these ICPS also presents challenges of scale.

Statistical aspects of Knowledge-based diagnostic approaches were poorly represented across the aerospace sector. Most applications of the technique in the automotive and industrial control sectors discussed Bayesian approaches and various Petri net derivatives. This may be due to the increasing presence of hybrid approaches which employ Knowledge-Based methods in the midst of other techniques. There was little evidence of traditional Expert Systems.

Connectivity is a key characteristic of ICPS yet it has deeper implications in our sectors of interest. Table 6 illustrates how connectivity for facilitating diagnostics is made more challenging because of the different environments ICPS operate in. Brief discussions in the papers of emerging cloud technolo-
gies pointed towards ways of establishing connectivity in more achievable ways.

While the TRL analysis provided a way of identifying and profiling the most mature approaches, these results cannot always be extrapolated across all three sectors. Almost all the avionic and aerospace studies profiled originated from organizations who were partnering with agencies such as NASA and ESA. These do not face the same intellectual property restrictions that restrict what we might expect to find published in the automotive and industrial control sectors.

During our paper selection, promising papers from the medical device ICPS sector gave a tantalizing glimpse of the differences and challenges that sector presents. We look forward to exploring that domain in a later study, where complex, safety-critical devices and regulatory certification are the norm rather than the exception.

References


Health Management 7 (2016) 1–16.


[56] ESA, ESA Software Engineering and Standardization.
[83] W. Ahmad, O. Hasan, Formalization of fault trees

[69] T. Ruppert, J. Abonyi, Software Sensor for Activity-

[67] X. Chang, J. Huang, F. Lu, Robust in-flight sensor

[65] S. Windmann, O. Niggemann, Efficient fault detection


[61] C. Wohlin, P. Runeson, P. A. Br-

[59] R. J. Terrile, F. G. Doumani, G. Y. Ho, B. L. Jackson,

[57] ISO/BSS, Space systems. Definition of the Technology

[55] R. Dearden, T. Willeke, R. Simmons, V. Verma,

[53] G. Ducard, The SMAC fault detection and isolation

[51] A. Kodali, Y. Zhang, C. Sankavaram, K. Pattipati,

[49] D. Trawczynski, J. Sosnowski, P. Gawkowski, Analyz-


[45] P. Goupil, R. Dayre, A. Gheorghe, H. Leberre, Sig-


doi:10.1109/ICPHYS.2018.

[41] IEEE Industrial Cyber-Physical Systems (ICPS),


[37] M. Allen, M. Fournou, G. Karsavina, K. Pattipati,

genetics and health management of machines,
Journal of Industrial Information Integration 6 (2017)


[23] T. Ruppert, J. Abonyi, Software Sensor for Activity-


[1] T. Ruppert, J. Abonyi, Software Sensor for Activity-


- Survey of 127 papers that profiles fault in industrial cyber-physical systems (ICPS).
- Focuses on safety-critical aspects of the aerospace, automotive and industrial control sectors.
- Evaluation of the most mature techniques ranked using a Technology Readiness Scale (TRL).
The Research focus on Fault Diagnostics in the Safety-Critical Aerospace, Automotive and Industrial Control sectors.
Barry Dowdeswell is a PhD Candidate in The School of Engineering, Computer and Mathematical Sciences at the Auckland University of Technology, New Zealand. His research focuses on fault diagnosis methodologies for industrial-scale cyber-physical systems. His current work includes the development semi-autonomous, multi-agent based fault diagnostic engines for IoT systems, particularly those built with IEC 61499 Function Blocks.

Roopak Sinha (PhD ’09, MCE ’16, BE (Hons) ’03, SFHEA ’17) is an Associate Professor in The School of Engineering, Computer and Mathematical Sciences at The Auckland University of Technology, New Zealand. He has previously held academic positions at The University of Auckland, New Zealand and INRIA, France. His research interest is “Systematic, Standards-First Design of Complex, Next-Generation Embedded Software” applied to domains like Internet-of-Things, Edge Computing, Cyber-Physical Systems, Home and Industrial Automation, and Intelligent Transportation Systems. He has served on several international standardisation projects, and works with several New Zealand companies to systematically reduce standards-compliance costs in IoT/embedded products.

Stephen MacDonell is Professor of Software Engineering at Auckland University of Technology and Professor in Information Science at the University of Otago, both in New Zealand. Stephen was awarded BCom(Hons) and MCom degrees from the University of Otago and a PhD from the University of Cambridge. His research has been published in IEEE Transactions on Software Engineering, ACM Transactions on Software Engineering and Methodology, ACM Computing Surveys, Empirical Software Engineering, Information & Management, the Journal of Systems and Software, Information and Software Technology, and the Project Management Journal, and his research findings have been presented at more than 100 international conferences. He is a Fellow of IT Professionals NZ, Senior Member of the IEEE and the IEEE Computer Society, and Member of the ACM, and he serves on the Editorial Board of Information and Software Technology. Stephen is also Theme Leader for Data Science & Digital Technologies in New Zealand’s National Science Challenge Science for Technological Innovation, Technical Advisor to the Office of the Federation of Māori Authorities Pou Whakātāmore Hangarau - Chief Advisor Innovation & Research, and Deputy Chair of Software Innovation New Zealand (SI^NZ).
Author contributions to *Finding Faults: a scoping study of Fault Diagnostics for Industrial Cyber-Physical Systems.*

**Barry Dowdeswell.** Corresponding author. PhD candidate, Auckland University of Technology:
- Conceptualization of the survey paper, methodology and survey data curation.
- Investigation.
- Writing - original draft.
- Writing – review and editing.
- Project administration.

**Associate Professor Roopak Sinha, Auckland University of Technology**
- Conceptualization of the paper.
- Supervision of the lead author.
- Methodology – design and advice on the application of the coding and analysis methodology and how they were applied.
- Writing – review and editing.
- Validation of cyber-physical system concepts and fault diagnostic concepts.

**Professor Stephen MacDonell, Auckland University of Technology**
- Conceptualization of the paper.
- Supervision of the lead author.
- Methodology – design and advice on the application of the survey protocols and review of how they were applied.
Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: