Data-driven Modeling, State Classification, and Performance Monitoring of Complex Physical Systems

by

Yaqing Xu

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Mechanical Engineering)
in the University of Michigan
2022

Doctoral Committee:

Associate Professor Kira Barton, Co-Chair
Assistant Professor Alex Shorter, Co-Chair
Assistant Professor Stephen Cain
Professor Albert Shih

Yaqing Xu

yaqing@umich.edu

ORCID iD: 0000-0003-0378-5010

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

This work is focused on using data-driven methods to investigate state behavior of complex systems. A complex system consists of several interacting elements and system behavior cannot be trivially inferred from the collective behavior of the individual elements. Important research questions in the study of complex systems include how to characterize individual agent behavior, and how system behavior may emerge from the collective behavior of the multi agent system. The work in this dissertation focuses on understanding and addressing modeling and analysis challenges in complex systems at both an agent- and system-level for manufacturing and biological systems.

My work in the manufacturing space is focused on using event-based modelling and machine learning to understand the interaction between working stations in a cylinder head production line to enhance performance monitoring. Companies strive to increase efficiency, improve quality and reduce costs by reducing downtime and improving productivity. In a manufacturing system, performance monitoring must consider the behavior of both individual machines and the interactions between these machines at the system level. Additionally, environmental factors, including: machine health, maintenance schedule, supply of raw materials and customer demands must be considered. Therefore, to develop an intelligent performance monitoring system, it is essential to understand both machine- and system-level interactions. However, current methods are either focused on monitoring single machines or simplified systems. To address this gap, a state-based model was developed to describe the performance of an individual machine, and a data-driven modeling approach was proposed to monitor system-level interactions, enabling the analysis of local disruptions on the overall performance of the system.

Elements of the framework developed in the manufacturing space were then applied to a compelling biological system, humpback whales. It is of interest for biologists to understand humpback whales' behavior in the wild, especially underwater behavior when direct observation is not available. Human activities including commercial fishing and shipping are threats to humpback whales in the wild. For example, fishing gear entanglement is the leading cause of injury and deaths for humpback whales. To help understand whale behavior and mitigate these threats, it is important to gain a better understanding of the animals in their natural habitat. In this part of my work, data-driven methods were used to combine information from multiple data sources to investigate the behavior of humpback whales in the Gulf of Maine.

These results have improved our understanding of animal movement, foraging ecology, and the temporal and spatial distribution of behavioral states at day scale. Importantly this new knowledge will directly inform mitigation strategies that seek to reduce fishing gear entanglement and vessel collision.

The ability to provide enhanced system-level classification and understanding of complex systems has the potential to impact several fields. This research provides new frameworks for state classification and characterization that can be used to understand agent behavior and system-level interactions within a manufacturing system. This framework was also applied for biological characterization to identify behavior patterns across day-scale time intervals. While these fields are drastically different, there are interesting aspects including state transition maps from each system that can lead to new insights in how systems interact across time and spatial domains.

CHAPTER 1

Introduction

1.1 Introduction

A complex system is a system with interacting elements, often called agents, whose behavior cannot be trivially inferred from the behavior of individual agents [5]. An agent is a self-contained module with a clearly defined goal that will adapt its decisions based on updates from a dynamic environment or other agents [6, 7]. Because an agent's behavior is dependent on many other factors, an equilibrium state is often hard to reach and requires information from a variety of resources [8]. The behavior of an agent in a complex system is time-varying and often non-linear. While the agent itself can be a complicated entity, its behavior may be able to be described through behavioral states or system level interactions. System level interactions include exchange of energy, matter, or information.The evolution of a complex system may be sensitive to initial conditions and small disruptions due to either a large number of agents, or because the agent behaviors are highly dependent on each other and environmental factors [9]. In many instances, a more thorough understanding of system-level behavior is necessary to understand the interactions between agents. Methods to effectively identify and describe these interactions between agents remains an open area of interest in the field of complex systems.

An important research question in complex system study is how to derive abstractions of dynamic models to describe the most important features of a system. These abstractions can take on many forms, including a system of complex networks to describe the brain [1], dynamic or stochastic models that describe the evolution of behavior over time such as for economic or biological systems [10, 11], and discrete-event systems that describe the transition for one state to another through event-based triggers as in manufacturing systems [12, 13].

The work in this dissertation focuses on understanding and addressing modeling and analysis challenges in complex systems at both an agent and system-level perspective for manufacturing and biological systems.

Despite the wide range of applications that can be categorized as complex systems (e.g. traffic,

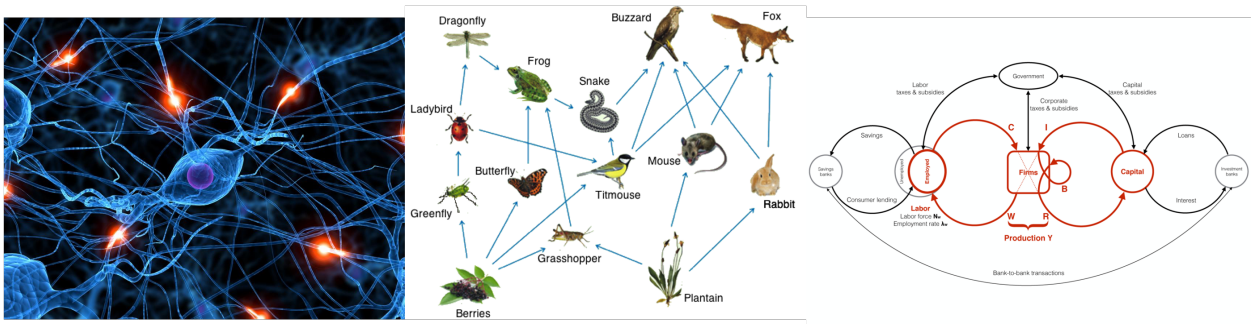

Figure 1.1: Three examples of complex systems. In the neural network, information is transmitted between neurons through chemical and electrical reactions. In the food web, species are linked through predator-prey relationships. In the economic market, the system-level behavior emerged from transactions between capital entities [1] [2] [3] [4]

.

energy, manufacturing, biological), there are common elements that can be used to describe these systems. In particular, the components in the system which are commonly referred to as agents are able to interact with each other and have the potential to form large-scale patterns. Within this description, agent-based characterizations and agent-to-agent interactions drive open research questions:

- What methods can be used to characterize an individual agent's behavior?
- How do agent architectures and the rules that govern their behavior vary both within and between systems?
- How does system behavior emerge from the collective behavior of all agents?
- How do environmental disruptions impact system-level behavior?

1.1.1 Motivation

The long-term goal of this research is to provide a framework for informed decision making and threat mitigation from the environment using data-driven methods. Towards this goal, this dissertation focuses on understanding agent behavior and the interactions between agents from a complex systems perspective. I focus my research on two specific application areas, manufacturing and biological systems, to investigate the extensibility of the proposed framework.

While dissimilar in many ways, these systems represent complex systems in which interactions between agents often drive new behaviors. In both systems, individual agents aim to accomplish certain goals, with collaboration playing an important role in both the individual and system-level

dynamics. The objective of this work is to develop a framework for understanding these interactions and using them to describe the behavior of the system. I apply data-driven methods to monitor, classify, and characterize these interactions towards an improved understanding of these complex systems.

Why study manufacturing system from a complex system point of view? Modern manufacturing systems collect large quantities of data about the production performances of the individual components, which provides a pathway for monitoring individual resources such as machines, material handling units, robots, etc. However, system-level interactions are often overlooked or ignored, resulting in an inability for the operator to make preemptive decisions regarding overall system performance. To better understand and make decisions about the performance of the entire manufacturing system, new methods for classifying and characterizing the interactions between elements within a manufacturing system must be developed.

Why use data-driven monitoring and classification methods to describe animal behavior? Wild humpback whales live together in pods, collaborating to enhance foraging efficiency and safety, among other reasons. Understanding the behavioral pattern of an individual whale including its temporal and spatial behavioral profiles in the wild is essential for understanding their collaboration strategies. Unfortunately, a significant portion of their behaviors is performed underwater, in an environment that is extremely hard to monitor. As such, novel methods for monitoring and characterizing the dynamics and movement patterns of these animals in their nature environments are key to better understanding animal behavior.

1.1.2 Target Research Areas

1.1.2.1 Manufacturing System

A manufacturing system can be viewed as a collection of interacting machines. A machine's working state is dependent on its own state and the states of other machines in the system. State transitions in a machine can be modeled as a finite state machine. Since a machine can only be in one state at a time, the state transition function can be modeled as a stochastic function of current and past states. Key performance indicators (KPIs) can be defined and used to determine if the production target is met. Additional metrics can be identified, monitored, and evaluated against expected values to determine the health of the system. Predictive health monitoring is an important area of research that has received considerable attention over the last several years [14]. In this area of research, historical data from individual agents and systems are used to assess a current health state and predict future behavior in the system, giving operators the opportunity to mitigate any potential threats. Understanding machine- and system-level interactions is an essential task towards the development of an effective performance monitoring system. To realize this objective,

one can start by classifying and modeling the range of behaviors from individual agents, such as machines and material handling devices, and then deriving methods for characterizing system-level interactions.

Windau and Itti presented a solution to upgrade an existing machine with an inertial machine monitoring system to detect and classify equipment failure or degraded states [15]. The proposed approach detects and classifies normal vs. abnormal equipment behaviors (loose belt, machine component failures) using a 3D printer as a use case. Liu et al. [16] introduced a distributed approach for plant-wide monitoring of a manufacturing system through the implementation of multiple models on different edge devices. A feature selection technique was proposed to classify the system into distributed modules for which an independent model was built based on feature subsets. Wang et al. [17] presented a health monitoring approach that leverages data analytics and subject matter expertise to detect and classify machine-part interactions useful for anomaly detection. A framework for manufacturing equipment health state monitoring and diagnostics using IoT big data in a real industrial setting has been proposed in [18]. Aspects of big data ingestion, big data management, data preparation, as well as predictive modelling, are taken into consideration for health state classification and monitoring. In [19], the authors introduced a digital twin-based performance monitoring approach that uses a data driven hierarchical digital twin (a combination of digital twins) for performance monitoring of different aspects of a manufacturing system such as equipment, controller, and process. Canizo et al. [20] presented a large-scale platform for real-time monitoring of cyber-physical systems based on big data technologies. The approach has been validated on an industrial use case that includes several industrial press machines and showed improvement on the OEE. Wan et al. [21] proposed a manufacturing big data solution for active preventative maintenance and fault diagnosis including big data collection and big data processing.

In [22, 23, 24], the authors introduced an approach to design and develop KPI-driven digital twin solutions for different manufacturing purposes including predictive maintenance and performance monitoring. A case study executing performance monitoring of manufacturing equipment (e.g., pump) using a hierarchical data-driven digital twin is presented. Another digital twin based approach for real-time monitoring of conventional machines is presented in [25]. The solution connects isolated machines to an interconnected system to monitor machine conditions in real-time. Other literature on the use of data-driven approaches to understand machine and system behaviors for performance monitoring and predictive maintenance in industry can be found in [26, 27, 28, 29, 30, 31, 32, 33].

While there are several data-driven performance monitoring frameworks in the literature, most of these approaches focus on monitoring individual machines. On the other hand, works that focus on system-level interactions demonstrate their frameworks on overly simplified systems that do not provide a validation of the scalability or generalizability of the proposed framework. Applications

to large complex manufacturing lines are still lacking.

1.1.2.2 Biological System

Complex system approaches have been applied to biological systems to study topics such as the structure of the brain [34], disease transmission [35], and evolutionary theory [36]. Depending on the topics, the agent in the system could represent a single neuron, a parasite and the host, or an entire species. System-level interactions that were studied within these systems include information transfer, parasite-host, and predator-prey relationship.

In this study, the focus is on the ecology of humpback whales.

Animal behavior monitoring has often been done in a controlled environment. Lauderdale et. al. studied how environmental enrichment can help increase activity diversity of dolphins in a zoo environment. There has also been a long history of direct observation approaches in wild animal monitoring including focal follow applied to whale behavior monitoring. Noren te. al. observed surface active behavior including spy hops, breaches, tail slaps, and pectoral fin slaps of killer whales to determine how the closeness of a vessel would influence whales' behavior [37]. Land-based visual observation, focal follow and passive acoustic monitoring are all available methods for animal behavior monitoring in the wild [38] [39] [40] [41]. These methods are usually labor-intensive and only work for observations of surface behaviors or have other limitations. Land-based observations were made to study the foraging habitat of humpback whales and how commercial whale watching affects the behavior of whales nearby. These methods provide a precise and detailed description of whale behavior and their interactions between each other and with other environmental factors in the wild. However, they are also inefficient, constrained due to the environment, require manual data collection, and can not observe underwater behavior.

Bio-logging tags enable the application of big data methods to the behavioral analysis of underwater biological systems [42] [43] [44]. Stimpert et. al. used bio-logging tags with sound recording devices to study the link between sound production and whale behavior through an analysis of received sound levels from different tagged whales [45]. Simon et. al. analyzed the behavior and kinematics of humpback whale lunge feeding using multi-sensor archival tags to study the energy cost of this behavior [46]. These studies are often only interested in foraging behavior, and an understanding of resting and traveling behavior in humpback whales is still lacking.

These methods require little human effort for data collection, can capture long periods of data, and can produce some kinematic features. However, detailed behavioral analysis and coordinated behavior within groups are still lacking.

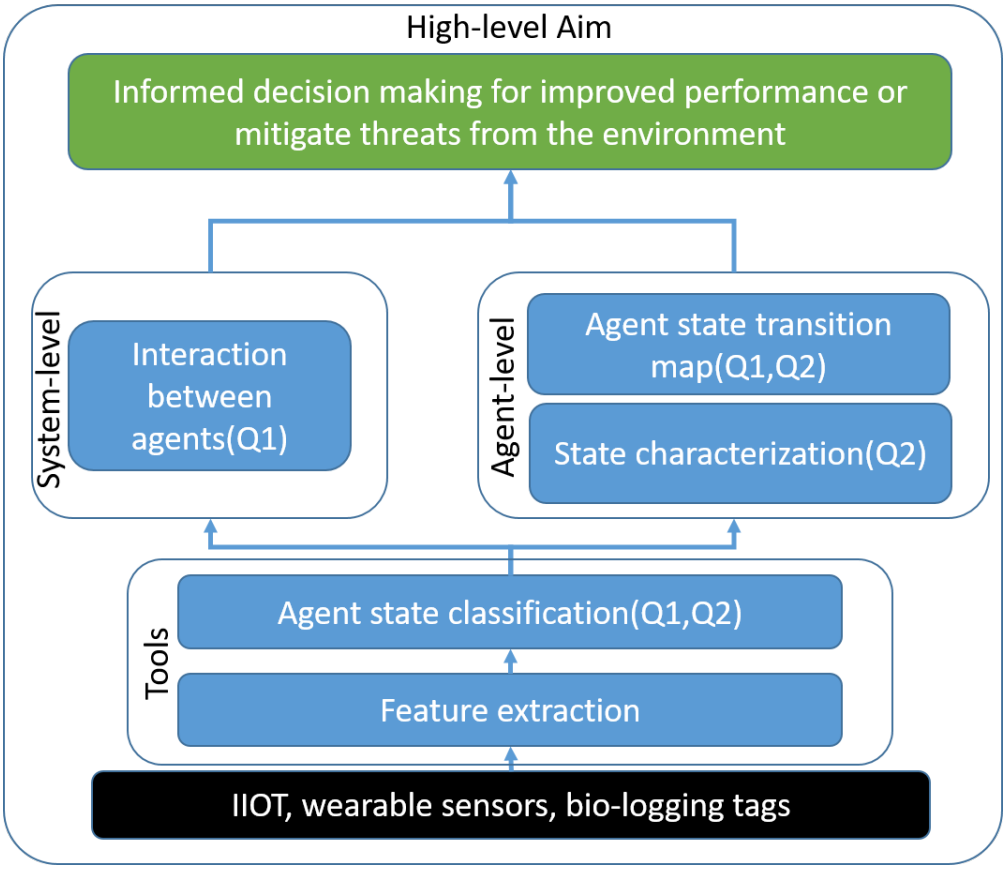

Figure 1.2: Framework for the dissertation.

1.1.3 Research Questions

Many of the examples presented in the previous application domains focused on single-agent characterization and analysis. While single-agent behaviors are critical to understanding complex systems, the system-level interactions, missing from most of the literature, must be investigated to develop descriptive relationships that can be used to understand the influence of individual agent behaviors on the system as a whole. Further, these relationships could then be used to develop predictive models for decision-making or to better understand the impact of environmental aspects on the complex system. To address this need, a framework for developing new knowledge and enabling tools to better understand complex systems must be developed. Figure 1.2 provides a high-level overview of the proposed framework used in this dissertation. Leveraging this framework, I proposed the following research questions:

RQ1 How can individual agent behaviors be used to characterize interactions between agents towards a system-level understanding of the collective behaviors?

Aim1: Derive a state-based description of agents leveraging classification methods and state transition models.

Aim2: Identify critical performance metrics used to quantify system-level interactions.

Aim3: Analyze the influence of local disruptions on the performance of a complex system.

RQ2 Can we develop robust, automated methods to classify and characterize agent behaviors in unstructured and uncertain environments?

Aim1: Identify key features in the data streams of agents in controlled environments that enable behavioral class identification.

Aim2: Identify the sensing modality, data analysis, and key features that enable behavioral class identification for agents in unstructured and uncertain environments.

1.1.4 Approach

To answer the proposed research questions, I focused my investigations on two key application domains: manufacturing and biological (e.g. humpback whales) systems. These domains were selected to highlight different aspects of complex systems. For example, manufacturing systems represent a controlled environment in which the behaviors of individual agents can be shown to impact the performance of the entire system. Additionally, while data collection is ubiquitous in manufacturing systems, methods for extracting specific features, characterizing the states of the agents from this data, and determining interactions between these agents remain open areas of interest. Biological systems, and in particular underwater biological systems, present an interesting data collection challenge that drives important questions in data analysis and feature selection. Additionally, given the importance of the environment in biological systems research, it is important to consider how the environment plays a role in the data collection and analysis capabilities.

1.1.5 Thesis Overview

This work is focused on using data-driven methods to investigate the state behavior of complex systems and understand the interaction between agents. These systems are often monitored using direct human observation, which limits the data that can be used to monitor and predict system states. As such, data-driven tools that leverage sensor measurements are needed for real-time characterization and assessment. Here I implement data-driven tools to investigate behavioral states and evaluate the performance of manufacturing (a cylinder head production line) and biological (humpback whales) systems.

This dissertation is organized in the following manner. Chapter 1 provides an introduction to complex systems along with methods and motivation for the studies presented in this work. Chapter 2 presents a background on the current state-of-the-art research in the fields of manufacturing

and biological systems. Chapter 3 presents the development of a data-driven framework for characterizing agents and system-level interactions in a manufacturing system. Chapter 4 presents a method to classify humpback whale behavior and study agent association with location and time. Concluding remarks and some future directions are provided in Chapter 5.

CHAPTER 2

Background

To understand the complex systems that are considered within the scope of this dissertation, this chapter provides a brief overview and state-of-the-art description of the research challenges in the core applications areas of Manufacturing System-level Interactions, and Monitoring Humpback Whales.

2.0.1 Manufacturing System Machine Interactions

One of the biggest challenges for manufacturers is understanding what is happening on the factory floor. To address this challenge, manufacturers define and track performance monitoring KPIs that can help them determine what is contributing to, or taking away from, their success. Tracking KPIs provides decision-makers with real-time information about shop-floor activities so they can make faster and more confident production decisions. A KPI is a number or value that can be compared against an internal target, or an external target "benchmarking" to give an indication of performance [47]. That number or value relates to data collected or calculated from a process or activity. The emergence of Industry 4.0 and the rapid growth in the use of various smart digital sensors and the Industrial Internet of Things (IIoT) have provided manufacturers with large amounts of data that could be used to derive KPI-based insights into manufacturing performance. A number of methods to derive such insights from big manufacturing data have been proposed in the literature. A framework for manufacturing equipment health state monitoring and diagnostics using IoT big data in a real industrial setting has been proposed in [18]. Aspects of big data ingestion, big data management, data preparation, as well as predictive modeling are taken into consideration. Canizo et al. [20] presented a large-scale platform for real-time monitoring of cyber-physical systems based on big data technologies. The approach has been validated on a real industrial use case that includes several industrial press machines and showed improvement in the overall equipment effectiveness. Liu et al. [16] introduced a distributed approach for manufacturing plant-wide monitoring through the implementation of multiple models on different edge devices. A feature selection technique is proposed to break down the system into distributed modules for which an independent model is

built based on feature subsets. Wan et al. [21] proposed a manufacturing big data solution for active preventive maintenance and fault diagnosis including big data collection and big data processing. In [25], the authors presented a digital-twin based approach for real-time monitoring of conventional machines that connect isolated machines to an interconnected system and monitor machine conditions in real-time. Theissler et al. provided a survey that reviews machine learning-enabled performance monitoring and predictive maintenance approaches in the automotive industry [33]. Other literature searches on the use of data-driven approaches for performance monitoring and predictive maintenance in the industry could be found in [26, 27, 28, 29, 30, 31, 32].

Despite the recent work in this area, there are several important gaps that need to be addressed. In particular, the understanding and quantifying of the interactions between machines within a manufacturing system is critical for decision-making with respect to overall system performance.

2.0.2 Humpback Whale Behavior Monitoring

Humpback whales are a species of baleen whales that live in all oceans around the world. While they have been hunted to almost extinction in the 1960s, the population growth of humpback whales is observed around the globe thanks to many conservation efforts [48]. The whales carry out long migration routes between feeding grounds in the high-altitude areas and breeding and calving grounds in the low-altitude areas [49]. Although social behaviors are less common in baleen whales than in toothed whales, the humpback whale is an exception. Research has shown that humpback whales form social groups based on reciprocal altruism [50]. During feeding season, humpback whales will form groups to perform bubble net feeding to increase the efficiency of foraging [51]. During the breeding season, male whales will come together and compete for mating opportunities [52]. Humpback whales are vulnerable to many environmental factors. During the breeding season, females with calves will devote most of their time to rest in shallow water, which makes them vulnerable to ship strikes [53]. Fishing gear entanglement is also a major source of injury and mortality for humpback whales in the wild [54]. Increasing sound pollution in the ocean has been shown to decrease the range of communication between humpback whales. Understanding humpback whale behavior in the wild is an important task for mitigating these threats.

There have been extensive studies of humpback whale behavior based on land-based and boat-based observations. [38] estimated the increase rate of humpback whale population on the east coast of Australia by land-based visual observations. [39] revealed the space use pattern of humpback whales by a combination of land-based observation and focal follow. [40] measures the vocal calling of blue whales by passive acoustic monitoring. These methods are usually labor-intensive and only work for the observation of surface behavior. [41] detected foraging behavior for toothed whales by focal following and observing the presence of prey or remains, but this only works for

toothed whales when the prey is large enough.

Developments in bio-logging tag technologies have enabled the collection of fine-scale data including location, acoustic, kinematic, and physiological data on wild animals. This technique provides crucial insights into animal behaviors in their natural habitat. Bio-logging tags are also used in marine mammal studies. [42] identified whale calling during a dive and analyzed the call rate for different positions and body postures. [43] identified lunge feeding which is indicative of foraging by estimating speed, body orientation, and dive profile from tag data. [44] designed a decision tree algorithm to detect foraging lounges with kinematic data. These studies were often only interested in foraging behavior, while knowledge about resting and traveling behaviors in humpback whales is still lacking.

There are several research gaps regarding long-term, detailed underwater motion monitoring for behavioral classification of humpback whales in the wild. Further, it is unclear how whales would adapt their behavior when subjected to environmental changes. This gap leads to the second research question: Can we identify and characterize agent behaviors in unstructured and uncertain environments? The goal is to classify and characterize whale behavior in the wild using tag data, and to understand the implication for mitigating threats from the environment.

CHAPTER 3

Data-Driven Modeling of Manufacturing Machines and System-Level Interactions for Enhanced Performance Monitoring

Performance monitoring is of great importance in manufacturing. It helps reduce downtime and increase efficiency. System performance is affected by both independent machines and interactions between machines. This chapter introduces a data-driven framework to assess the performance of discrete manufacturing systems and identify relationships between machines. Data-driven models for monitoring the behavior of both machine and system levels are developed to analyze the performance of a manufacturing system. Data are collected from an automotive production line using Industrial Internet of Things (IIoT) solutions and a sophisticated information system software that tracks the state of the entire plant. The proposed framework uses probabilistic automata and logistic regression modeling to estimate Key Performance Indicators (KPI) at machine and system levels and monitor the output variables. Model outputs are used as a reference to detect abnormal behaviors based on deviations from production targets. Results using data from an industrial case study demonstrate the ability to perform real-time performance monitoring, capture errors within confidence intervals, and identify predictive cause and effect relationships between machines in a production line to support well-informed decision making.

3.0.1 Introduction

Manufacturing is undergoing a transformation driven by digital technologies such as the Industrial Internet of Things (IIoT) and cloud computing platforms that allow the collection, storage, and exchange of huge amounts of data across a system's value chain [55]. Manufacturers are embracing digital technologies to enhance the performance of their production lines [56]. Manufacturing system performance is affected by the behavior of both individual machines and the interactions of these machines at the system level. Efficiency of manufacturing operations is often controlled

by monitoring Key Performance Indicators (KPIs) [29, 57] that allow decision makers to take corrective actions when the behavior at the machine- or system-level deviates from expectation.

Up-time, throughput, quality, and cost are important KPIs that manufacturers track and work to improve to enhance their capability and competitiveness. To achieve these objectives, several methods have been developed to evaluate the performance of manufacturing processes and systems. For instance, Overall Equipment Effectiveness (OEE) has been widely adopted to evaluate factory performance at the system level by assessing availability, productivity, and quality [58, 59, 60]. Equipment and process precision and health condition are highly related to OEE [58]. There has been wide interest in developing intelligent performance monitoring systems to maintain and improve OEE by assessing equipment and process health conditions as well as predicting and preventing unwanted degradation and failures.

Understanding system-level interactions is a complex problem that requires targeted data and data analytic frameworks that support the study of these interactions. A variety of interactions result from disruptions and disturbances such as rapid demand changes, rescheduling requirements, equipment failures, or system reconfiguration requests. Additionally, processing and analyzing manufacturing data from disparate sources and levels of the factory floor and deriving models to understand these data is not an easy task.

For factory floor operational efficiency, it is necessary to include performance monitoring models that provide a reference for the expected performance that can be used as a comparison to actual factory floor data. Since manufacturing requirements can change rapidly and equipment operations are subject to variability, performance monitoring and assessment tools must be able to adapt and run synchronously with the factory floor.

Understanding machine- and system-level interactions is an essential task towards the development of an effective performance monitoring system. To realize this objective, one can start by classifying and modeling the range of behaviors from individual components such as machines and material handling devices, then, derive methods for characterizing system-level interactions.

Windau and Itti presented a solution to upgrade an existing machine with an inertial machine monitoring system to detect and classify equipment failure or degraded states [15]. The proposed approach detects and classifies normal vs. abnormal equipment behaviors (loose belt, machine component failures) using a 3D printer as a use case. Liu et al. [16] introduced a distributed approach for plant-wide monitoring of a manufacturing system through the implementation of multiple models on different edge devices. A feature selection technique was proposed to classify the system into distributed modules for which an independent model was built based on feature subsets. Wang et al. [17] presented a health monitoring approach that leverages data analytics and subject matter expertise to detect and classify machine-part interactions useful for anomaly detection. A framework for manufacturing equipment health state monitoring and diagnostics using IoT

big data in a real industrial setting has been proposed in [18]. Aspects of big data ingestion, big data management, data preparation, as well as predictive modelling, are taken into consideration for health state classification and monitoring. In [19], the authors introduced a digital twin-based performance monitoring approach that uses a data driven hierarchical digital twin (a combination of digital twins) for performance monitoring of different aspects of a manufacturing system such as equipment, controller, and process. Canizo et al. [20] presented a large-scale platform for real-time monitoring of cyber-physical systems based on big data technologies. The approach has been validated on an industrial use case that includes several industrial press machines and showed improvement on the OEE. Wan et al. [21] proposed a manufacturing big data solution for active preventative maintenance and fault diagnosis including big data collection and big data processing. In [22, 23, 24], the authors introduced an approach to design and develop KPI-driven digital twin solutions for different manufacturing purposes including predictive maintenance and performance monitoring. A case study executing performance monitoring of manufacturing equipment (e.g., pump) using a hierarchical data-driven digital twin is presented. Another digital twin based approach for real-time monitoring of conventional machines is presented in [25]. The solution connects isolated machines to an interconnected system to monitor machine conditions in real-time. Other literature on the use of data-driven approaches to understand machine and system behaviors for performance monitoring and predictive maintenance in industry can be found in [26, 27, 28, 29, 30, 31, 32, 33].

While there are several data-driven performance monitoring frameworks in the literature, most of these approaches focus on monitoring individual machines. On the other hand, works that focus on system-level interactions demonstrate their frameworks on overly simplified systems that do not provide a validation of the scalability or generalizability of the proposed framework. Applications to large complex manufacturing lines are still lacking.

The performance monitoring approach presented in this chapter is based on industry data collected from a large automotive manufacturing plant. Our approach aims to develop modeling and simulation tools that provide scalable health monitoring solutions to support manufacturing operations during run-time. The specific contributions of this chapter include:

1. A method that integrates Subject Matter Expertise (SME) with data analytics to effectively identify critical operations or bottlenecks in the system;

2. A state-based description of the manufacturing resources (e.g. machine and material handling units) within a manufacturing system;

3. Identification of critical KPI metrics that can be used to quantify system-level interactions; A key KPI introduced in this work is a "$Flow$" metric, which is a system-level KPI that can be measured to predict the state of a manufacturing unit based on the system interactions.

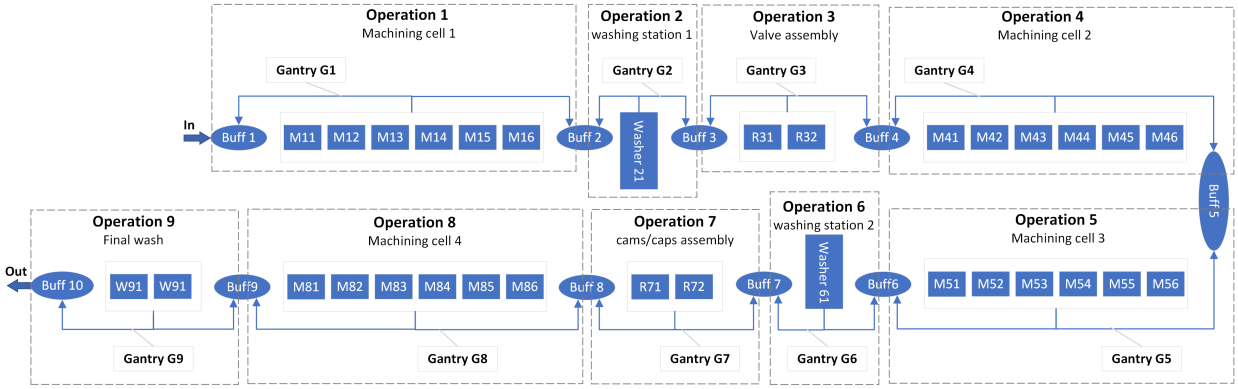

Figure 3.1: Layout of an example production line. "**Buff**" stands for **Buff**er, "**M**" for CNC **M**achine, "**R**" for **R**obot, and "**W**" for **W**asher.

4. A data-driven modeling approach for machine and system-level performance monitoring that provides insights into system-level interactions by analyzing the influence of local disruptions on the performance of the manufacturing system.

The rest of this chapter is structured as follows. Section 3.0.2 introduces an example manufacturing system used to illustrate the proposed work. Section 3.0.3 describes the overall methodology for data-driven performance monitoring. A system-level performance monitoring approach is given in Section 3.0.4. Section 3.0.5 demonstrates the validity of the approach through a case study using the system described in Section 3.0.2. The results and a discussion are provided in Section 3.0.6. Finally, Section 3.0.7 concludes the chapter and discusses other applications and future work directions.

3.0.2 Manufacturing System Example

The manufacturing system shown in Fig. 3.1 represents an example production line that could be found in many automotive manufacturing facilities. The system is comprised of several serial operations that consist of sets of manufacturing units, i.e., Computer Numerical Control (CNC) machines, washing machines, and assembly robots that work concurrently on the parts. Buffers and gantries exist between the operations to store and to move parts between the manufacturing units, respectively. The example system processes two types of parts. Both types pass through the same system operations in different batches; however, additional features are incorporated to one of the two types through additional machining and assembly processes.

The behavior of each manufacturing unit is described by a set of discrete states. The state transitions are event or time driven [61]. A manufacturing unit can be configured in an "Up" or "$Down$" superstate. The "Up" superstate defines the states when the manufacturing unit is

operating and could produce or move parts depending on the surrounding conditions. The "$Down$" superstate defines the states when the manufacturing unit is unable to move or produce parts. The "Up" and "$Down$" superstates each consist of a set of substates.

The "Up" superstate comprises the following substates:

- "$Busy$": state when a manufacturing unit is executing a task on a part. This state includes three additional substates: "$Loading$" state, which denotes when material is entering the work position, "$Unloading$" state, which denotes when material is exiting the work position, and the "$Working$" state, which refers to when the manufacturing unit is processing a part. A working state that falls within the predefined cycle time distribution is referred to as a "$Cyclic\ State$". Working states that do not fall within this distribution are referred to as "$Noncyclic\ States$".

- "$Blocked$": the state where a manufacturing unit has finished its assigned working task but the part cannot be cleared yet. Note that for a gantry there are three types of blocked states, namely, blocked primary, blocked secondary, and blocked tertiary. These three states depend on what is causing the gantry to be in a blocked state (e.g., downstream machine busy, upstream or downstream buffer empty, etc.)

- "$Starved$": the state where a manufacturing unit is ready to start its assigned task but it cannot because a part is not present yet.

- "$Waiting$": the state where a manufacturing unit is capable and ready to start executing a working task, and a part is available, but it cannot start because additional conditions are not fulfilled.

The "$Down$" superstate comprises the following sub-states:

- "$Setup$": the state initiated when a setup procedure (e.g., warm up, tool change, mastering, clean machine table and other surfaces, etc.) is launched.

- "$Bypass$": the state where the machine is not functioning correctly and it is simply passing entering parts without working on them.

- "$Break$": A state triggered manually by an operator to go for a break.

- "$Repair$": the state wherein the machine has a physical intervention such as opening a gate for maintenance.

- "$Tool\ change$": the state triggered when a tool change procedure is initiated.

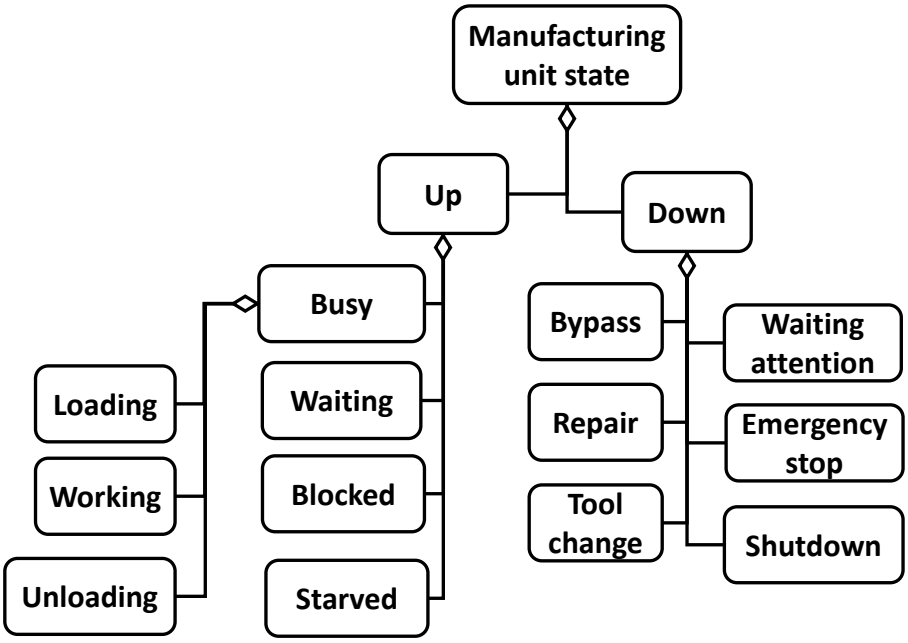

Figure 3.2: The different states of a manufacturing unit

- "$Waiting\ attention$": the state the manufacturing unit enters when a fault occurs while in an "Up" state.
- "$Emergency\ stop$": a safety state which is triggered by an emergency stop button or by another safety system. It is configured to abort the operation of the machine and place it in a safe condition as quickly as possible.
- "$Shutdown$": the state triggered manually by an operator, usually using a button on the machine control panel, to shut down the machine.

The system is connected to a sophisticated information system software that tracks daily up-time, downtime, and real-time bottlenecks. Figure 3.2 summarizes the observed and collected discrete states of each of the manufacturing units in the system. The time duration of each of these states is recorded by the system controller. Data collected from the system include the timestamp of the start and end of all the "Up" and "$Down$" states of each manufacturing unit. The duration of each state is provided by accumulating the time between the start and end of the state. Data were collected at a sample rate of 1 second. In addition to the state names and time stamps, data collected also include the manufacturing unit's name and part identification numbers, which are used to track parts throughout the production line.

It has been noticed by subject matter experts at the production line that the manufacturing

units operate under non-steady state conditions that cause a high cycle time variability, which is manifested by different state time-durations. This cycle time variability has a negative impact on throughput and subsequently negative impact on demand fulfilment. In the subsequent sections, we propose a methodology to further classify the machine states in order to understand this system variability. Machine- and system-level performance monitoring models are then derived to track deviations from normal behaviors and thus support decisions on how to react to these deviations in order to meet throughput targets.

3.0.3 Methodology

This section describes the proposed approach for monitoring system performance using manufacturing data. The goal is to identify predictive relationships between machine interactions and their effect on system-level performance, then incorporate this knowledge into models that provide insights into machine- and system-level stochastic conditions. Model predictions can be ultimately used to trigger warnings and recommend optimal operating parameters and critical responses to stoppages based on current system conditions.

Figure 3.3 gives an overview of the workflow of the proposed methodology, which consists of the following four operations:

1. *Identification of critical operations*: system data are first used to identify critical operations or bottlenecks in the system as these have a significant impact on the overall performance and the throughput of the system.

2. *Classification of manufacturing units productive states*: Data distribution fittings are used to further classify the manufacturing units states based on their time-duration and occurrence frequency. This state classification will be used to understand and model the machine-level and system-level behaviors for performance monitoring.

3. *Development of machine-level performance monitoring models*: after classifying the different behaviors of individual manufacturing units, probability distributions and event-based sequencing from the data are used to develop models that are used for performance monitoring of single manufacturing units.

4. *Development of system-level performance monitoring model*: The state classification of the different manufacturing units is also used here to develop a system-level model that predicts deviations of the throughput of the entire production line. The model identifies predictive relationships between machine interactions and their effect on the overall system performance. It is used as a monitoring tool for enhancing system performance.

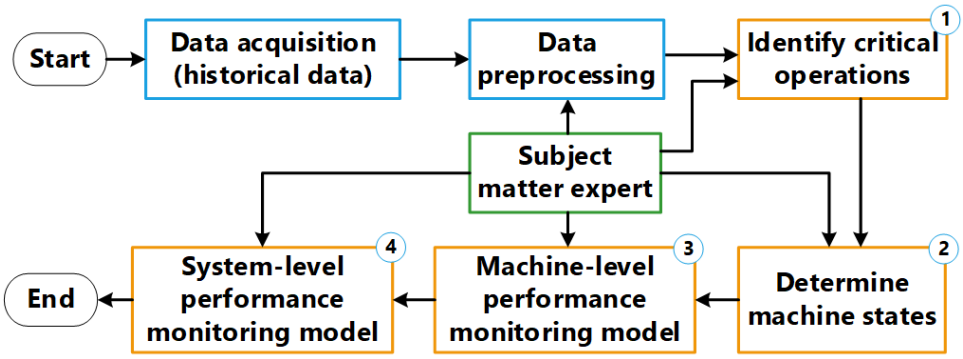

Figure 3.3: Workflow of the proposed performance monitoring system development

3.0.3.1 Identification of the Critical Operations

To initiate the monitoring process, it is important to identify the critical operations. A critical operation is defined as an operation in the production line that has a negative impact on the overall system performance. For instance, a bottleneck operation slows down production and limits the throughput could be a critical operation. Other system performance metrics that could be impacted by critical operations include energy consumption, capacity utilization, maintenance cost, etc. Thus, identifying critical operations could result in major performance improvements.

A critical operation can be identified using multiple metrics, including a bottleneck measure, downtime occurrence frequency at the machine level, total downtime at the machine and cell levels, and percentage of non-cyclic working states at the machine level. These metrics can be calculated as follows.

- Bottleneck is defined as the operation with the smallest number of parts produced per unit time, denoted as

$$I_{bottleneck} = \arg\min_i \frac{1}{d_i^w} \cdot n_i = \arg\max_i \frac{d_i^w}{n_i} \tag{3.1}$$

  where d_i^w is the average working state duration for the ith operation and n_i is the number of machines in this operation.

- Number of downtime occurrences o_{Mij}^d is defined as the number of occurrences of a down state in machine M_{ij} during a given data collection period.

- Downtime of machine i in cell j (M_{ij}) is denoted as

$$D_{Mij}^d = d_{Mij}^d \times o_{Mij}^d \tag{3.2}$$

  where d_{Mij}^d is the average duration of a down state for machine M_{ij} and o_{Mij}^d is the number of downtime occurrences in that machine.

- Downtime at the cell level is the accumulation of the downtime across all machines within that cell,

$$D^d_{Ci} = \sum_{j=1}^{n_i} D^d_{Mij} \tag{3.3}$$

where n_i is the number of machines in the ith operation.

- Non-cyclic percentage for machine M_{ij} is the percentage of non-cyclic working states within all machine working states. It is denoted as

$$P^{nc}_{Mij} = \frac{o^{ncw}_{Mij}}{o^{ncw}_{Mij} + o^{cw}_{Mij}} \tag{3.4}$$

where o^{ncw}_{Mij} and o^{cw}_{Mij} are the occurrence of non-cyclic working states and cyclic working states in machine M_{ij} during the defined time period, respectively.

A comparison of the operations in the manufacturing line with respect to the above metrics can then be carried out. The operation that demonstrates the *highest sum* (or other combined value) of the identified metrics will be defined as the critical operation. As a critical operation exhibits several unwanted behaviors (such as downtime and cycle time variability), it is beneficial to derive the performance monitoring models around such an operation, then generalize the models to the other operations. Details of this process are provided in the subsequent sections.

3.0.3.2 System State Classification

The goal of system state classification is to identify and define models of normative behaviors such that we can detect abnormal behaviors. Towards this goal, we propose an approach that considers KPIs, such as cycle time, and classifies different states based on this KPI in terms of normal vs. abnormal behaviors. We focus on the productive states, i.e. "$Working$", "$Starved$", and "$Blocked$" as their occurrence frequency is higher than the other states and their impact on the system throughput is significant. A qualitative statistical analysis of these states suggests that they can be separated into sub-states by duration range, and that within each range the data is often distributed normally. A Jarque-Bera test (JB test) [62] is used to provide an assessment for how well a distribution complies with a normal distribution structure by calculating the Jarque-Bera (JB) score. A JB score of 0 indicates that the data is normally distributed and a large JB score suggests that the data does not come from a normal distribution. If the JB score is equal to 1, it means that the null hypothesis that the data is normally distributed is rejected at a 5% significance level.

The JB score is calculated as follows:

$$JB(D) = \frac{n}{6}(S^2 + \frac{(K-3)^2}{4}) \tag{3.5}$$

where D is the set denoted as $D = \{d_1, d_2...d_n\}$, here d_i is the duration of a state type (e.g., "$Working$", "$Starved$", and "$Blocked$"), n is the data sample size, S is the skewness, and K is the kurtosis of D, defined as:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{d})^3}{(\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{d})^2)^{3/2}} \tag{3.6}$$

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{d})^4}{(\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{d})^2)^2} \tag{3.7}$$

Here $\bar{d}$ is the average duration of all the states in D.

Algorithm 1 is proposed to classify the system states based on their duration and frequency of occurrence. In Algorithm 1, sub-state duration ranges are defined by a list $L = \{(LB_0, UB_0), ..., (LB_n, UB_n)\}$, where LB_i is the lower bound and UB_i is the upper bound of sub-state i. LB_0 is initially set to 0 ($LB_0 = 0$) referring to the lowest cycle possible. As the algorithm iterates to cluster the state duration ranges, LB_{j+1} will be set to UB_j ($LB_{j+1} = UB_j$). Algorithm 1 consists of the following steps.

- **Step 1:** Let D' be a set that takes all the states in D, and j an index that iterates through the state duration ranges the algorithm finds until all the states in D' are clustered. In step 1, Algorithm 1 initializes j, D', LB_0, and L to $j = 0$, $D' = D$, $LB_0 = 0$, and $L = \emptyset$.

- **Step 2:** In step 2, Algorithm 1 iterates over all states d_i in D' and puts every state whose duration is shorter than d_i in a new (intermediate) set named D''. Algorithm 1 then calculates the JB score of D'', denoted as JB_score_i using Equation 5. The value of UB_j is set to be equal to the d_i that corresponds to the minimum JB_score_i.

- **Step 3:** In step 3, Algorithm 1 updates the list of substate clusters L with the range of substates j, then removes this cluster from the set D'. The lower bound of the next substate cluster will start at the upper bound of the previous cluster.

- **Step 4:** In step 4, steps 2 and 3 are repeated for $j = 1, 2...$ until all the states are clustered and D' is empty.

For a given LB_i, Algorithm 1 searches for a UB_i that makes the data distribution range from LB_i to UB_i have the lowest JB score. Distribution fittings are applied to the set of productive states of a manufacturing unit through Algorithm 1. After the duration range of each sub-state is found, the following process is used to classify the type of distribution for each sub-state.

Algorithm 1: Cyclic and non-cyclic sub-state classification

Input : $D = \{d_1, d_2...d_n\}$, set of a manufacturing unit state durations
Output: $L = \{(LB_0, UB_0), ..., (LB_i, UB_i)\}$, classification of manufacturing unit states based on their duration
Initialization: LB_0 = 0, j = 0, $D' = D$, $L = \emptyset$;
while $D' \neq \emptyset$ **do**
    JB_ref = 1;
    for $d_i \in D'$ **do**
        $D'' = \{d \in D' \mid LB_j \leq d < d_i\}$;
        JB_score_i = JB(D'');
        if $JB\_score_i < JB\_ref$ **then**
            $JB\_ref = JB\_score_i$;
            $UB_j = d_i$;
        end
    end
    $LB_{j+1} = UB_j$;
    $D' = D' - \{d \mid d < UB_j\}$;
    $L = L \cup \{(LB_j, UB_j)\}$;
    $j = j + 1$;
end

- A p-value test is conducted for each common distribution. The distribution with the highest p-value greater than 0.05 is selected. If all of the p-values are lower than 0.05, the next step is carried out.
- Bootstrapped Skewness-Kurtosis plot of the current dataset is compared to that of common distributions. The distributions that are close to the bootstrapped distribution are selected as candidates.
- If there is more than one candidate, a list of goodness of fit test statistics are calculated to find the best fitting distribution.

As a result, each productive state is further divided into sub-states. One of these sub-states is the "$Regular$" cycle time that is predefined by design (the most frequent in the data) and the rest are classified based on their duration, namely, "$Short$", "$Medium$", and "$Long$" cycle times.

3.0.3.3 Machine-level performance monitoring

A machine-level performance monitoring model can be used to track the activities of a single manufacturing unit by providing insights into the frequency of occurrence and duration of its states.

Such insights help with understanding the correlation between manufacturing unit states and consequently what drives their cycle time variability.

A single manufacturing unit can be monitored using a Probabilistic Automaton (PA) model G defined by the tuple

$$G = (Q, F, P, \Delta) \tag{3.8}$$

- $Q = \{Up \cup Down\}$ is a finite set of states, with Up = $\{Busy, Waiting, Blocked, Starved\}$ as the set of Up states and $Down$ = $\{Shutdown, Emergency\ stop, Waiting\ attention, Tool\ change, Repair, Bypass\}$ as the set of Down states.
- F_i is the frequency of occurrence of each sub-state $q_i \in Q$ in a defined time window.
- P_{ij} is the probability of transition to sub-state j when in sub-state i.
- $\Delta = Q \times P \rightarrow Q$ is a transition function.

The automaton G is constructed by recording the sequence of states from the input data, and calculating the transition probability between these states.

Figure 3.4 shows an example probabilistic automaton model. During a given data collection time period, occurrences of busy, waiting, and down states are 100 times, 9 times, and 2 times, respectively. Labels in the transition arrows correspond to the probabilities of state transitions. In this example, when the manufacturing unit is in a busy state, there is a 90% chance that it will stay in the busy state, 8% probability it will transition to a waiting state and 2% probability it will transition to a down state. For a down state, there is a 50% chance it will return to the busy state and a 50% chance it will transition to a waiting state.

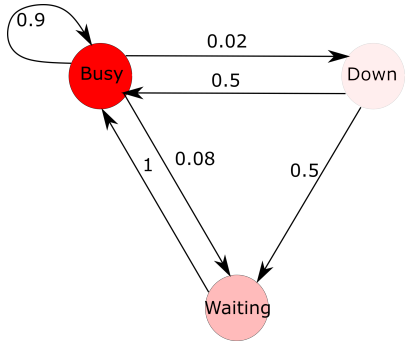

Figure 3.4: Example Probabilistic Automaton.

3.0.4 System-level performance monitoring

To effectively evaluate and monitor manufacturing system-level performance, a data-driven modeling approach is developed. System-level interactions are studied by capturing relationships be-

tween multiple manufacturing units such as machines and gantries, and evaluating their effects on system throughput. Logistic regression analysis is used with event-based sequencing of data to measure and predict system-level KPIs for performance monitoring and assessment. In this section, logistic regression analysis is briefly recalled, then, the steps of the proposed methodology for performance monitoring are presented.

3.0.4.1 Logistic Regression

Many problems faced in manufacturing are concerned with establishing a relationship between a set of system variables. For instance, we might be interested in the relationship between the system throughput and machine and material handling device cycle times, as well as the times that the parts spend in queues. Knowledge of such a relationship would allow us to predict the output of the system, and then adjust the parameters that would give the best outcome.

In many scenarios, there is a single response variable Y, also called the dependent variable, that depends on the value of a set of input variables, also called independent variables. A basic and commonly used type of relationship between the dependent variable Y and the input variables $x_i = x_1, \ldots, x_n$ is a linear relationship that for some constants b and $\theta = a_1, ..., a_n$ is given by

$$Y = \theta * x_i + b \tag{3.9}$$

The linear equation (3.9) is a tool commonly used in machine learning to learn a model that estimates values of the coefficient θ used in the representation of the available data. If the outcome variable Y is continuous i.e., can be measured and ordered, and has an infinite number of values between any two values, equation (3.9) is called a linear regression model. If Y is discrete i.e., taking on two or more possible values, equation (3.9) is called a logistic regression model. The difference between logistic and linear regression is reflected both in the choice of a parametric model and in the assumptions. Once this difference is accounted for, the methods applied in an analysis using logistic regression follow the same general principles used in linear regression [63]. We chose to use a logistic regression model for system-level performance monitoring because, compared to other machine learning algorithms that try to find a relationship between independent and dependent variables, logistic regression is less complex, easier to implement, easier to interpret, and very efficient to train. Logistic regression has been widely studied and applied in many applications due to its simplicity of use, efficiency, and effectiveness [64, 65].

3.0.4.2 System Performance Monitoring Using Logistic Regression

As mentioned before, the performance of a manufacturing system is often controlled by monitoring KPIs such as throughput, processing time, production downtime, and maintenance costs. KPI

analysis allows one to take corrective actions when the system deviates from its expected behavior. Logistic regression could be used to predict such deviations and determine predictive actions rather than corrective ones. We propose a methodology to develop models that predict future states of the system based on logistic regression and event-based sequencing of data. The prediction models are used as a system performance monitoring mechanism that supports decision makers in understanding unexpected system behavior and determining appropriate decisions. The development of a system-level performance monitoring model consists of the following steps.

- **Step 1. Identify a KPI of interest:** Manufacturing KPIs are performance measurements that help understand how a machine, system, or organization is performing. Effective KPIs should be well-defined, quantifiable, crucial to achieving a goal, and applicable to the system of interest. However, there are many KPIs to choose from. The first step in the methodology presented here is to choose the right KPI that measures something that aligns with the intended goals.

  One example of a KPI used in this work is flow state; a system-level KPI that can be measured to predict the state of a manufacturing unit based on known (modeled) system-level interactions. We define flow as a binary state that describes the productive state of a manufacturing unit. The flow classification ("Good Flow" and "Turbulent Flow") depends on the flowrate over a pre-defined window. Flowrate of a manufacturing unit is a continuous metric defined as the ratio of time the manufacturing unit spends in a "$Working$" state over the duration of the pre-defined time window. A threshold limit that describes when the flow transitions between a good flow state and a turbulent flow state is defined by analyzing historical data of the flowrate.

  The flow state is formally defined as:

$$\text{Flow} = \begin{cases} Good\ Flow & \text{if flowrate} > \text{threshold} \\ Turbulent\ Flow & \text{otherwise} \end{cases} \tag{3.10}$$

- **Step 2. Set up the logistic regression problem:** Logistic regression analysis is used to predict future values of the chosen KPI. Here, the future flow state (good or turbulent) is the dependent variable Y to be predicted over a given time horizon (H) using the states $x_i = x_1, \ldots, x_n$ from the previous time horizon (D) as independent variables. The variable $\theta = a_1, a_2, ..., a_n$ is the set of regression coefficients to be estimated from the data. The output of the dependent variable Y over a time horizon H is equal to 1 if the predicted flow state is a good flow and equal to 0 otherwise.

The time horizon H is determined by a subject matter expert. It should be accurately determined to better support the decision making actions. For instance, if H is too short, the user would not have enough time to make a decision that could potentially avoid a turbulent flow state. The independent variables $x_i = x_1, \ldots, x_n$ are the manufacturing unit's state set $Q = \{\text{“}Busy\text{”}, \text{“}Blocked\text{”}, \text{“}Starved\text{”}, \text{“}Waiting\text{”}, \text{“}Down\text{”}\}$, with their further classified sub-states. The x variables are defined as the dominant states in the manufacturing units over the time horizon D.

A range of plausible H and D time periods is first defined. The time period D is chosen on a trial and error basis. Manufacturing unit states are sampled using a sample rate of d during the time period D. One-hot encoding [66] is used to transform categorical data into numerical form.

The states from all of the machines during the previous time horizon (D) are stacked into a single input vector with the flow state over a given time horizon (H) provided as the output.

- **Step 3. Determine algorithm performance metric:**

  The logistic regression model outputs a number between 0 and 1 that represents the probability of the flow being good. A threshold is selected to distinguish between the predicted probability. If the probability is greater than this threshold, the flow is predicted to be in a good flow state, otherwise it is predicted as turbulent flow. The Area Under the Curve (AUC) - Receiver Operating Characteristic (ROC), also referred to as AUROC (Area Under the Receiver Operating Characteristics), is used to determine the threshold value. AUROC gives the rate of successful classification of the logistic regression model. AUROC is generated by plotting the true positive rate against the false positive rate. An ideal threshold is chosen to maximize the difference between these two rates.

  The logistic regression model is then trained and tested using training and testing data sets to validate and check the accuracy of the model. The minimum sample size required for training depends on the number of features and the complexity of the problem. The training, testing and validation set are split in a 5:2:3 ratio initially. If the performance is not optimal, the training set sample size should be increased.

3.0.5 Case Study

This section describes the application of the proposed methodology to the real-world production line presented in Section 3.0.2. Results and a discussion of the output of the case study will be presented in the following section.

3.0.5.1 Data Processing

Manufacturing data was collected from an industrial production line in a US manufacturing company over a five month period between 2016 and 2017. The manufacturing data was collected using a sophisticated information system that provides data from the manufacturing equipment and parts. The data was used to track part movements through the system and provide information about the amount of time a resource required to move or process a part (i.e., the time duration of the states described in Fig. 3.2). Each state was matched with a specific part identification number. Loading and unloading states were disregarded as they occurred before and after each working state and were assumed to be consistent.

3.0.5.2 Identification of Critical Operations

In this case study, we identified critical operations that impacted the throughput and maintenance requirements for the system. The KPIs chosen for this purpose included critical bottleneck identification, total downtime, downtime occurrence frequency, and percentage of non-cyclic working states. As given in Section 3.0.3.1, these metrics were calculated for each operation in the production line, and a summary table was developed to identify the operation that exhibited the highest value of these KPIs.

Table 3.1 summarizes the machines and operations that ranked in the top three for the KPIs defined in section 3.0.3.1. Operations 7 and 8 (including individual machines within these operations) appeared in the top three rankings of both operation and machine-level metrics. For example, R71 (Robot 1 in Operation 7) had the third largest number of downtime occurrences and the largest percentage of non-cycle working states, while Operation 8 showed the largest overall downtime, among other undesirable metrics. Since Operation 7 required manual operation, which might introduce additional disturbances to the system, Operation 8 was selected as the critical operation in this case study with machine M82 identified as the most critical machine due to its high percentage of non-cyclic working states.

Table 3.1: Top 3 Ranking for Critical Operation Identification

Metric	First	Second	Third
Bottleneck (operation level)	OP1	OP7	OP8
Total Downtime (operation level)	OP8	OP5	OP1
Total Downtime (machine level)	M81	M41	M82
# of Downtime Occurrence (machine level)	W92	W91	R71
Non-cyclic Percentage (machine level)	R71	R72	M82

Table 3.2: Cutoff Percentiles for Blocked and Starved States

Classification	Duration(s)
Instant	$<$50th percentile
Short	50th to 70th percentile
Medium	70th to 90th percentile
Long	$\geq$ 90th percentile

3.0.5.3 System State Classification

As described in Section 3.0.3.2, system state classification is a key step towards the development of machine and system performance monitoring models. The state classification took into consideration the variability in cycle times observed in the collected manufacturing data. Only the states that had higher time variability and higher occurrence frequency were further segmented into substates. We focused this work on the following states:

- "$Busy$": broken down into "$Loading$", "$Working$", and "$Unloading$" substates.
- "$Blocked$": broken down into blocked primary, blocked secondary, and blocked tertiary substates for gantry systems (see Section 3.0.2).
- "$Starved$": broken into instant, short, medium, and long substates.

The "$Down$" states did not exhibit significant variability and thus were lumped into one state class.
The "$Blocked\ Primary$" and "$Starved\ Primary$" states occurred in both the machines and gantries, and had similar cycle time distributions; thus, the descriptions of their substates was grouped together. "$Blocked\ Secondary$" and "$Blocked\ Tertiary$" states only occurred in the gantries and were processed separately. The sub-classification for each of these state groups included "$Instant$", "$Short$", "$Medium$", and "$Long$" substates, for which the duration was determined using cutoff percentiles. A subject matter expert determined the cutoff percentiles for each sub-state as shown in Table 3.2.

The description of a "$Working$" state was based on the CNC machines, since the gantry systems demonstrated relatively consistent cycle times. To fully classify the working states, we divided these states into several substates and then fit each substate to a known distribution by applying Algorithm **??**. To determine the distribution range for each substate, we applied Algorithm **??** to the machine data from Operation 8 using the following steps.

- Set the lower bound of the first substate to 0.
- Using the approach from Section 3.0.3.2, find the upper bound that minimizes the JB statistic.

- Set the lower bound for the second substate equal to the upper bound from the first substate. Repeat these steps until the remaining data points represent less than 5% of all data points.

Once the time range for a substate was determined, we found the distribution that best fit the data. There were several methods that were used to determine the distribution.

- P-value test: A list of commonly known distributions (e.g., Normal, Weibull, Chi-Square, etc.) was evaluated. Iterating through the distributions, we determined that none of these choices resulted in a p-value greater than 0.05. Bootstrapping [67] and skewness-kurtosis methods were then employed.

- Applying a bootstrapping and skewness-kurtosis metho ds, we evaluated the S-K fit for a set of common distributions. Figure 3.5 shows a skewness-kurtosis (S-K) plot, where both Gamma and Weibull distributions were good candidates for a fit.

- Goodness of fit test: Additional goodness of fit tests were used to evaluate the potential distributions. Applying Kolmogorov–Smirnov, Anderson–Darling, and Cramer-von Mises tests, we determined that Weibull was the best fit for this data set.

This process was repeated to find the distributions of all of the substates. After the substate durations were determined, the substate with the largest percentage of data points was the nominal cycle time state, denoted as substate R (Regular cycle time). The regular cycle time falls between T^R_{min} and T^R_{max}, which respectively refer to LB_1 and UB_1 derived from Algorithm 1. All the other substates were labeled as non-cyclic substates and classified according to their time duration. Substates with cycle time durations shorter than R ($< T^R_{min}$) were denoted as L. Substates with cycle time durations between T^R_{max} and $T^R_{max} + 10$ were denoted as A, substates with cycle time durations between $T^R_{max} + 10$ and $T^R_{max} + 20$ were denoted as B, substates with cycle time durations between $T^R_{max}+20$ and $T^R_{max}+30$ were denoted as A, and substates with cycle time durations greater than $T^R_{max} + 30$ were denoted as D.
The causes of cycle time variability exhibited in the different substates is assessed by subject matter experts. For instance, substate C, which took 20 - 30 seconds longer than the regular cycle time, was recognized as an event that included a sensor reading after the completion of the action associated with the part, see Table 3.3. This approach was repeated for all of the machines in operation 8. Further analysis could be used to apply this method across the entire manufacturing system to generate a substate distribution as illustrated in Table 3.3 and Figure 3.6.

3.0.5.4 Machine Performance Monitoring Model

To monitor the state behavior of the impacted machines, Probabilistic Automaton (PA) models of the gantries and machines were developed. Figure 3.7 shows the data-driven PA model

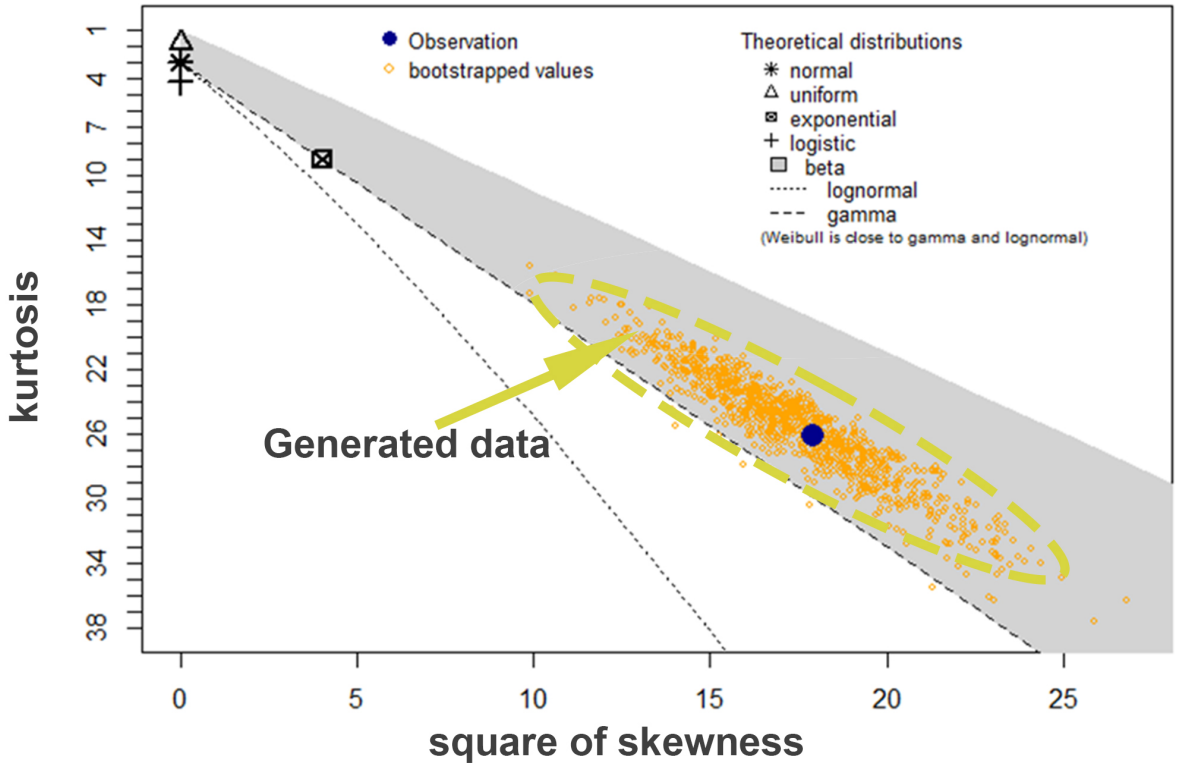

Figure 3.5: Skewness-Kurtosis plot of a sub-state. The black dot represents the observation and the yellow dots correspond to data generated using bootstrapping method. Beta distribution indicated by grey shades and Gamma distribution indicated by dashed line are the closest to the distribution of the sub-state

$G^{G8} = (Q^{G8}, F^{G8}, P^{G8}, \Delta^{G8})$ developed for gantry G8 in Operation 8. Each node represents a substate $q_i \in Q^{G8} = \{Working, Blocked, Starved, Down\}$. Note that "$Working$", "$Blocked$", and "$Starved$" were further classified into sub-substates in the PA model. In Fig. 3.7, "BP", "BS" and "BT" represent blocked primary, blocked secondary, blocked tertiary, respectively. The intensity of the nodes color is proportional to the frequency of occurrence $f_i \in F^{G8}$ for each state. Each arrow in the model represents a transition, and the label associated with the arrow represents the state transition probability $p_i \in P^{GG8}$. As an example, G8 in Fig. 3.7 shows dominant transitions from the working state as a 23% probability of transitioning back to the working state to start a new cycle, a 34% probability of transitioning to a blocked tertiary state, and a 32% probability of transitioning to blocked secondary state.

Similarly, Fig. 3.8 represents a PA model $G^{M42} = (Q^{M42}, F^{M42}, P^{M42}, \Delta^{M42})$ developed for machine 2 in Operation 4 (M_{42}). The states included $Q^{M42} = \{Working \cup Blocked \cup Waiting \cup Down\}$, where the working substate were $Working = \{L, A, R, B, C\}$, and the blocked states,

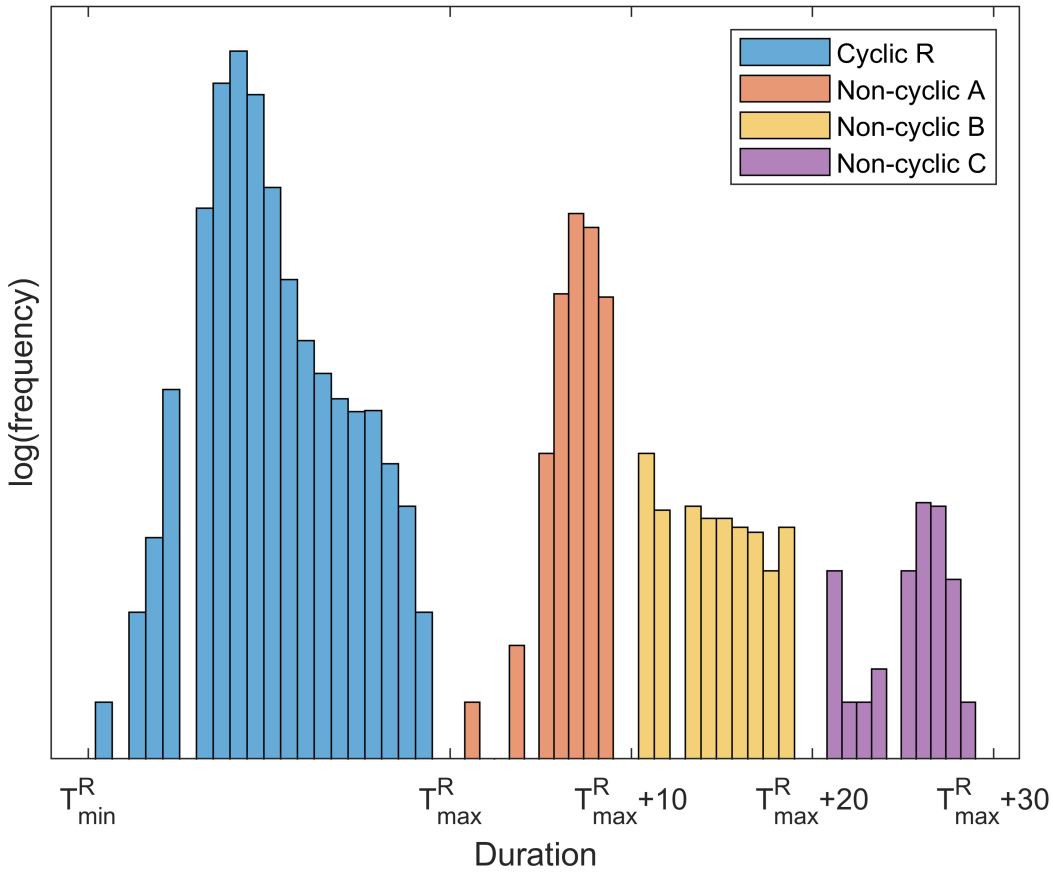

Figure 3.6: Example substate classification of the "$Working$" state in machine M82.

Table 3.3: Substates classification and their SME labeled causes.

Substate Classification	Time Range (s)	SEM Labeled Cause
Cyclic R	$[T^R_{min}, T^R_{max})$	Nominal distribution range
Non-cyclic A	$[T^R_{max}, T^R_{max} + 10)$	Unknown
Non-cyclic B	$[T^R_{max} + 10, T^R_{max} + 20)$	Quality checking
Non-cyclic C	$[T^R_{max} + 20, T^R_{max} + 30)$	Sensor reading
Non-cyclic D	$[T^R_{max} + 30, \infty)$	Unknown
Non-cyclic L	$[11, T^R_{min})$	Task not performed

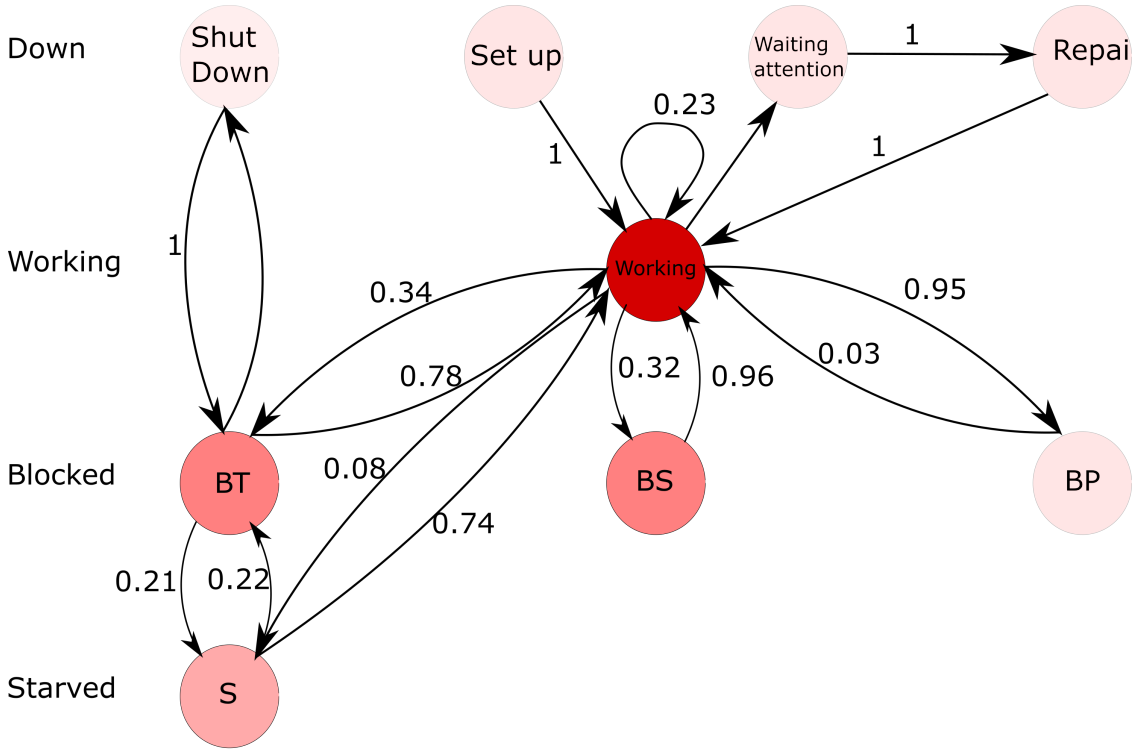

Figure 3.7: Machine-level performance monitoring model for gantry G8 (one day). Each node represents a substate with the frequency of the node illustrated by the color intensity. Each arrow represents a transition where the label represents the state transition probability.

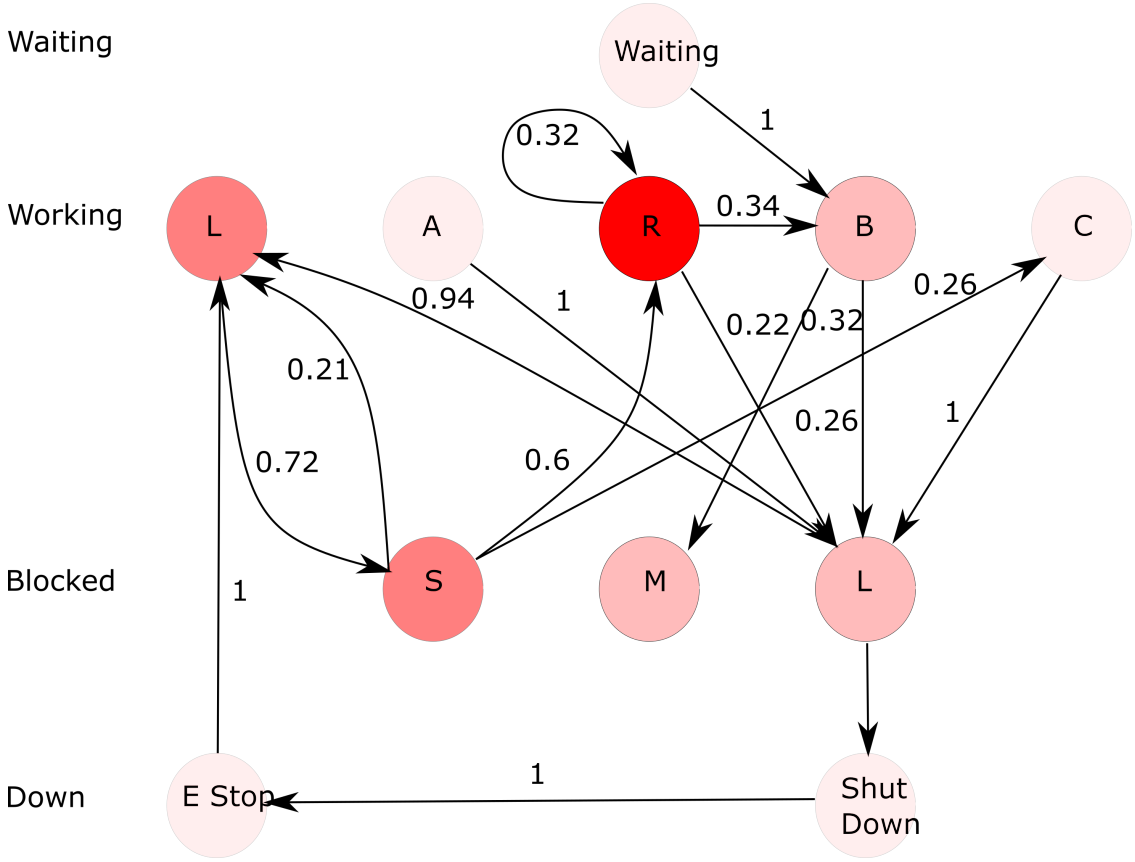

Figure 3.8: Machine-level performance monitoring model for a CNC machine

$Blocked = \{S, M, L\}$, included S for Short blocked state, M for Medium blocked state, and L for Long blocked state. The down state were given as $Down = \{Shut\ down, E - Stop\}$.

Figure 3.8 shows substate R (nominal working cycle time) as the most frequent working state with the highest probabilistic transitions for starting a new cycle and transitioning to the non-cyclic substate B. One can also note that after a medium or long blocked state, the machine primarily transitioned to non-cyclic substate L, where the assigned task was not performed. These model provided insights into the performance of the manufacturing system and helped to highlight the potential causes of abnormal behaviors. Given this information, SMEs (e.g., shop floor operator, maintenance personnel) could determine appropriate measures to avoid a loss in performance.

Once the machine-level models were derived, a system-level model was developed to investigate the impact of machine-level disturbances on the overall system performance.

3.0.5.5 System Performance Monitoring Model

As described in Section 3.0.4.2, a logistic regression model was developed to monitor the system-level KPI defined as flow. The model was used to identify performance trends in the data that could serve as indicators of the system transitioning to a turbulent flow state, which would affect overall system throughput. The following steps were used to derive the system-level performance monitoring model.

- **Step 1. Define the KPI of interest**: The KPI used to study system-level interaction was flow state. Using M82 as an example, data were segmented into windows of length D_0. Flow state was determined using the flowrate of the manufacturing units, defined as the ratio of time spent in a "$Working$" state during D_0. A threshold th value to determine the flow state classification was selected using a k-means clustering method (or SME determination). The flow state for this case study was defined as:

$$\text{Flow} = \begin{cases} Good\ Flow & \text{if flowrate} > 0.5 \\ Turbulent\ Flow & \text{otherwise} \end{cases} \tag{3.11}$$

- **Step 2. Set up the logistic regression problem**: The sample rate (d) was chosen as the lowest working duration in the system. For a given time t, state information from $t - D$ to t for all machines and gantries within the system was used to predict the flow state for machine M_{82} between t and $t + H$, where H was the prediction horizon and D was the trend time horizon.

The regression algorithm was given as:

$$Y' = \theta \cdot x + b = \begin{bmatrix} \theta & b \end{bmatrix} \cdot \begin{bmatrix} x \\ 1 \end{bmatrix} = \Theta x' \tag{3.12}$$

$$Y = f(Y') = \frac{1}{1 + e^{-Y'}}. \tag{3.13}$$

The input sampled at time t was defined as:

$$x^t = \{x_1^t, x_2^t, ...x_n^t\} \tag{3.14}$$

where x_i^t was the production state, defined as a one-hot machine state encoder, of the ith machine in the system from $t - D$ to t.

$$x_i^t = \begin{cases} [1,0,0] & \text{for a productive state} \\ [0,1,0] & \text{for blocked, starved, or waiting states} \\ [0,0,1] & \text{for a down state} \end{cases} \tag{3.15}$$

The output at time t defined the flow of the critical machine from t to $t + H$

$$Y^t = \begin{cases} 1 & \text{if good flow} \\ 0 & \text{if turbulent flow} \end{cases} \tag{3.16}$$

The outcome of the logistic regression was in the form of a matrix of weights, θ, that was used to identify which machine states could be used to predict the flow state of a specific machine. θ was derived by solving,

$$\Theta = \arg\min_{\Theta} \frac{1}{N} \sum_{t=0}^{T} (Y^t - f(\Theta \cdot \begin{bmatrix} x^t \\ 1 \end{bmatrix}))^2 + \frac{1}{2C}|\Theta|^2, \tag{3.17}$$

where N denoted the number of samples and C represented the regularization parameter that needed to be determined heuristically.

- **Step 3. Determine algorithm performance metrics**: The dataset was randomly partitioned into training, testing, and evaluation sets with a 5:2:3 ratio. The training and testing sets were used to optimize the parameters for the logistic regression. The evaluation set was used to evaluate the performance of the algorithm. The range of input (H) and output (D) durations

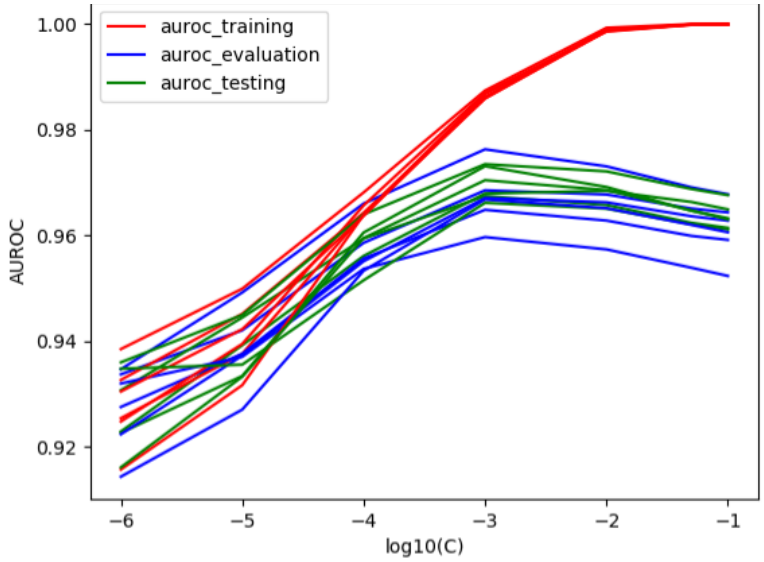

Figure 3.9: AUROC as a function of parameter C. AUROC of training data set increases as C increases, but the AUROC of the evaluation data set has a concave shape. C = 0.001 is chosen when the evaluation AUROC hits its local peak. AUROC of testing data set has similar trend to that on evaluation data set, which provides validation of this choice.

were selected to span below and above the working duration value d. The evaluation range for parameter C was 0.000001 to 0.1.

Parameter C was selected based on the highest AUCROC from the testing set, shown in Fig. 3.9 as $C = 0.001$.

Table 3.4 shows the prediction results for machine 2, operation 8. True positive represents the probability that a good flow output was classified as good flow, while false positive denotes the probability that a turbulent flow output was mis-classified as good flow.

3.0.6 Results and Discussion

This section presents analysis results for the industrial case study and discusses how the methodology proposed in this chapter can be used for a more effective decision making process on the factory floor.

Table 3.4: Flow prediction results for M82

True positive rate	91%
False positive rate	29%

3.0.6.1 Machine-Level Analysis

Machine-level analysis provides not only important insights into individual machine performance, but also enables an understanding of the impact of machine-level disturbances on downstream resources.

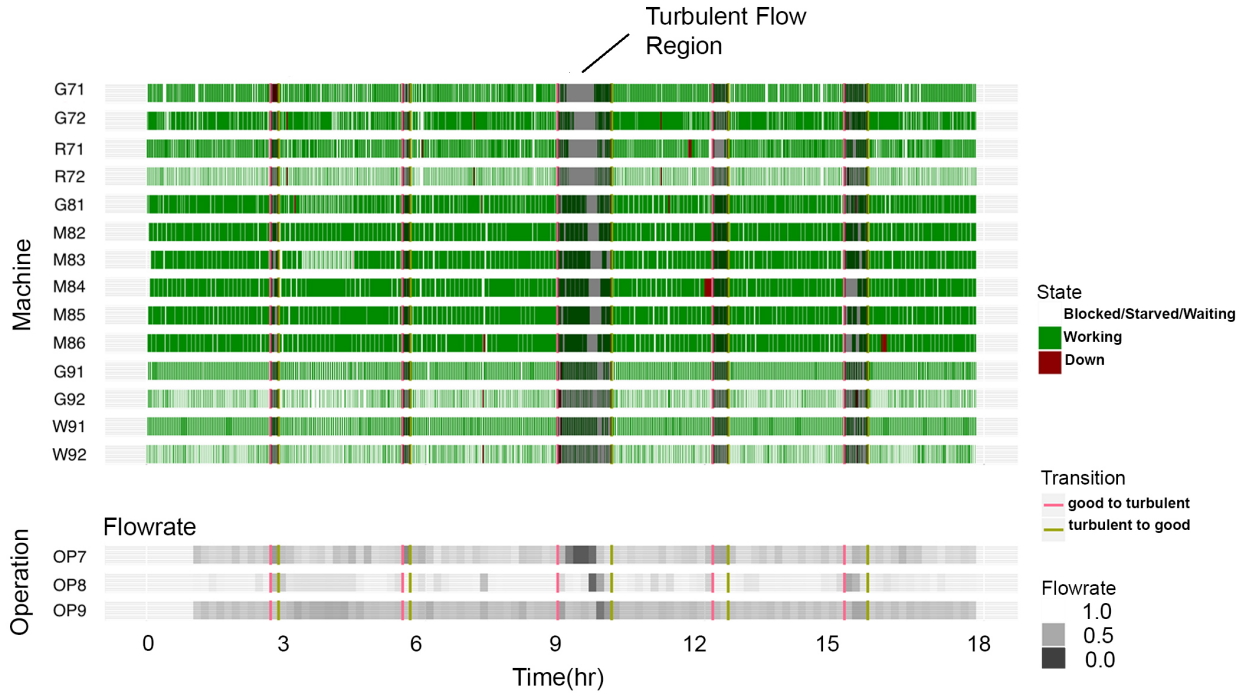

Figure 3.10: Performance monitoring for the machines in operations OP7, OP8, and OP9. In the upper graph, manufacturing states are indicated using green, red, and white colors, where the shaded region denotes a turbulent flow section. In the bottom graph, shading indicates the flowrate. Darker colors represent turbulent flow, with transitions signaled with red and green lines.

Figure 3.10 illustrates a one day example of the production states and flow classifications for the resources (e.g., CNC machines (M), gantries (G), robots (R), and washers (W)) in operations OP7, OP8, and OP9. The figure shows that the longest period of turbulent flow in these operations happens in the middle of the day. Figure 3.10 also shows that stacking the individual machine-level states provides a system-level view, where the impact of turbulent flow in one machine is observable on the downstream machines. With this type of analysis, one can identify when and where a turbulent flow starts / ends, investigate how it affects upstream and downstream machines / operations, and determine which operations require human actions to avoid performance degradation.

Figure 3.11 shows a zoomed-in view of the machines state in operation OP8. This figure provides information about the machine states when a turbulent flow starts, which can be used to identify trends and predict transitions to turbulent flow. For example, several machines in operation

8 begin to exhibit blocked states. As a result of this, gantry G8 starts to show a starved state, which then triggers more of the machines in operation 8 to show blocked states. Mitigation actions, such as additional buffer space proceeding operation 8 might provide sufficient capacity to remove parts from the machines and keep the production flow moving.

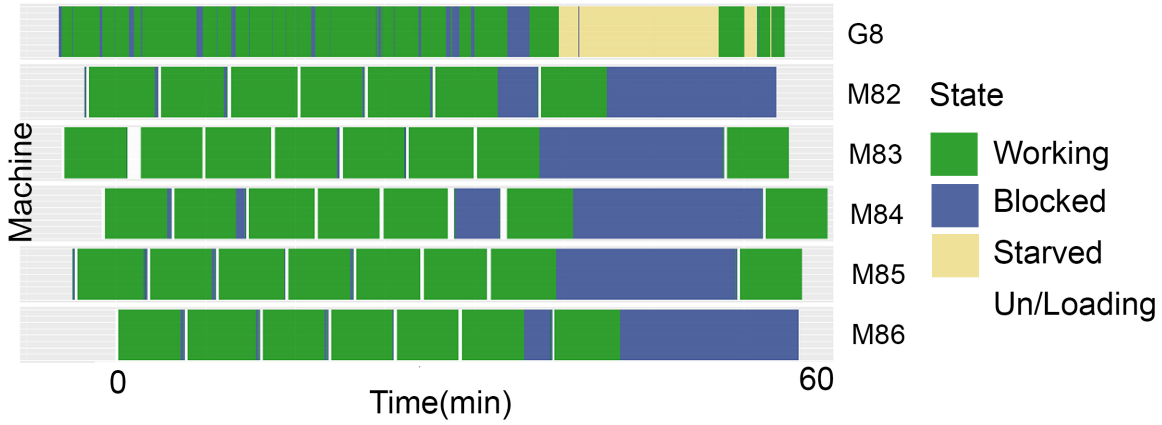

Figure 3.11: Performance monitoring of the machines and gantry in operation OP8. Green, blue, and yellow indicate working, blocked, and starved states, respectively. Loading and unloading states are labeled in white.

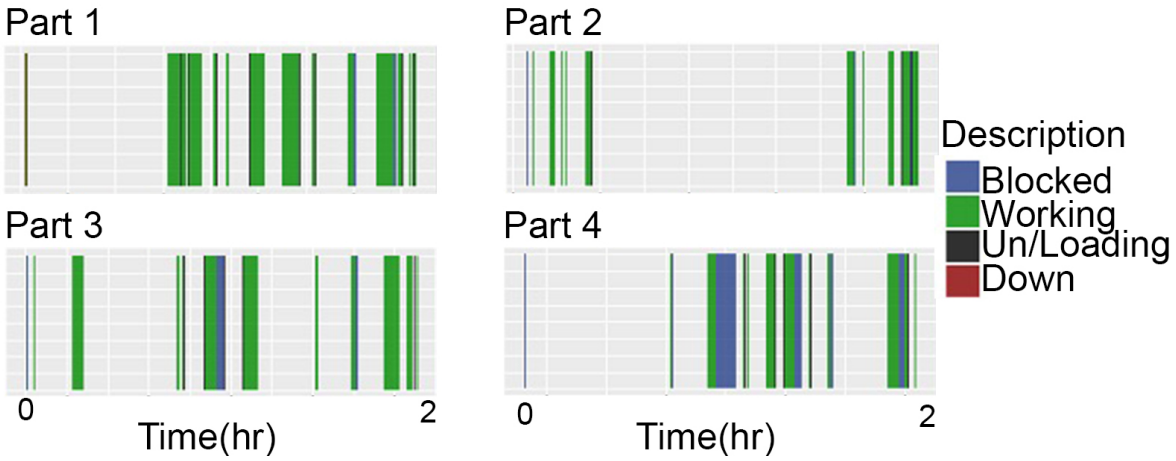

Figure 3.12: State sequencing with duration and timing for four example parts.

Another way to evaluate system performance is by tracking each processed part throughout the production line. Figures 3.12 and 3.13 track the production operation time and state distributions for four example parts. Using synchronized data from multiple data sources, the time spent in each state and operation for individual parts is shown. Figure 3.10 shows the flowrate progression throughout the system, it shows operators when and where a turbulent flow starts. And Figure 3.11

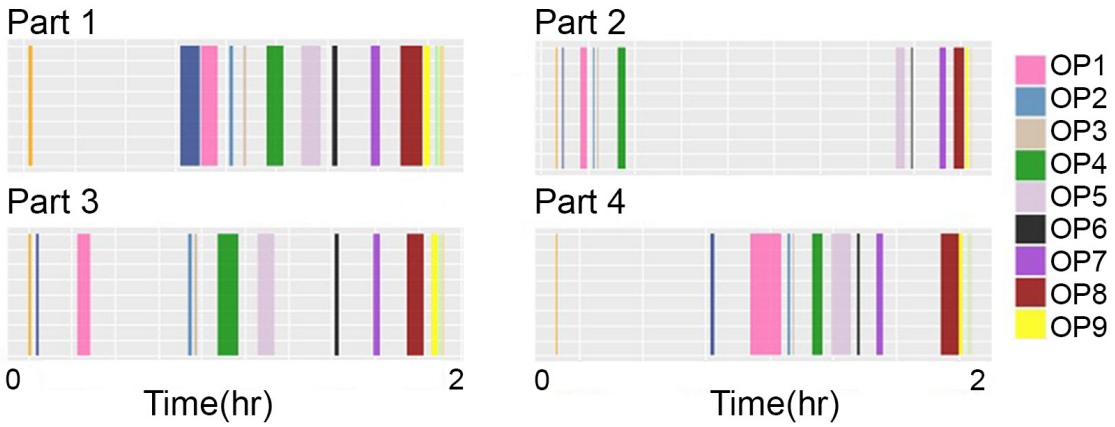

Figure 3.13: Operation sequencing, including duration and timing, for four example parts.

allows the operator to investigate further into the operation of interest. Figure 3.12 and Figure 3.13 provides a more detailed view of when a specific part is in a operation and the state associates with it. When a problem arises, this information helps to identify if the anomalous behavior is associated with a part, an operation, or a certain time of the day.

3.0.6.2 System Level Analysis

System-level analysis is used to determine the interactions between machines / gantries to better understand the propagation of poor behavior and derive methods to predict and eventually mitigate these behaviors. Here we use system-level analysis to quantify the influence of different states and operations on the flow state of a specific machine by evaluating the weighting matrix θ from the logistic regression algorithm. This analysis can be applied to any machine of interest to predict turbulent flow. If turbulent flow is predicted, an operator has a window of opportunity to prepare for a mitigation action. For instance, the operator could schedule a routine maintenance task when an operation is not productive or add or remove parts from a buffer to prevent a starved or blocked state in the neighbouring operations. Figures 3.14 and 3.15 provide visualizations of the matrix values of θ from the logistic regression model. Each element in θ corresponds to one of three states for a machine or gantry: "$Busy$", "$Idle$", or "$Down$" where, $Busy = \{Working \cup Loading \cup Unloading\}$ and $Idle = \{Blocked \cup Starved \cup Waiting\}$.

Analysis of the elements in θ reveals correlations between machine-level states and future flow state. For example, Fig. 3.14 shows that working states from operations OP6, OP7, and OP8 up to 20 minutes prior to the current time are associated with good flow in M82. Idle states from OP6, OP7, and OP8 up to 15 minutes prior and down states from OP9 up to 40 minutes prior are associated with turbulent flow in M82. Note that Fig. 3.14 shows the system-level interaction in

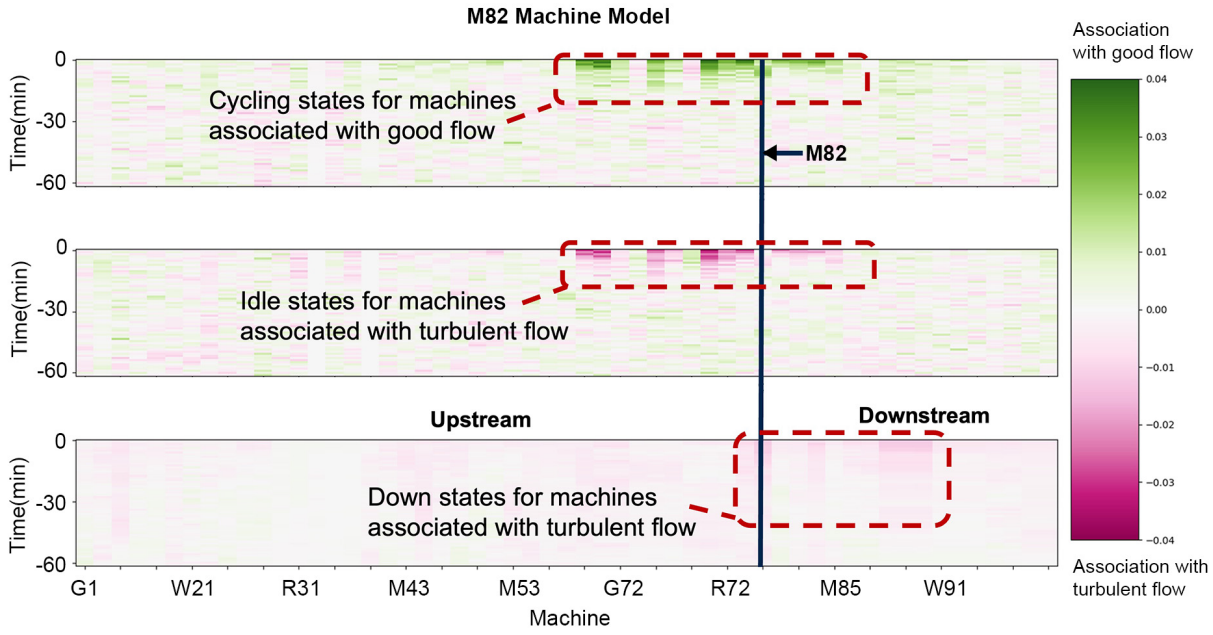

Figure 3.14: Flow association as a function of time and machine. The horizontal axis represents the location of each machine or gantry within the system; up-stream to the left, down-stream to the right. The vertical axis represents time before the current time. Intensity of a grid location is correlated to the value of the weighting gain corresponding to the time and machine. Darker colors indicate a higher association to good (green) or turbulent (red) flow for M82.

terms of upstream and downstream operations.

Figure 3.15 investigates the impact of system-level interactions on the flow state of a given machine. This figure illustrates that the working states in OP7 and OP8 are the best indicators of good flow in M_{82}, while idle states in OP7 and down states in OP9 can be used to predict turbulent flow. Although Figs. 3.14 and 3.15 present similar information, Fig. 3.14 includes information about the time horizon of the state influence, while Fig. 3.15 provides a more visually intuitive representation of the system-level interactions.

3.0.7 Conclusion

In this chapter, a data-driven framework for manufacturing machine- and system-level performance monitoring was developed and tested using data collected from an industrial manufacturing production line. This approach proposed four key steps that included identification of critical operations in the system that have a significant impact on the overall system performance, state classification to partition machine / gantry states towards understanding cycle time variability, the development of machine-level performance monitoring models that provide insight into a single

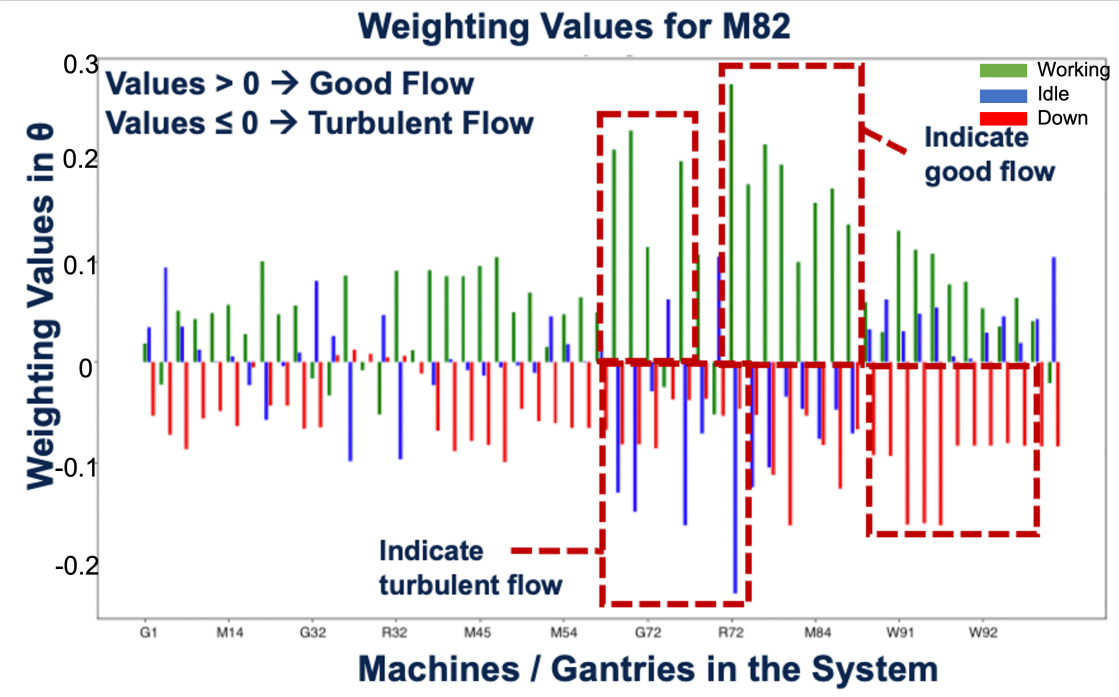

Figure 3.15: Weighted flow as a function of state. The horizontal axis describes the location of the machine or gantry within the system, while the vertical axis denotes the gain value from θ that corresponds to the influence of a given machine or gantry state on the flow of M82. A positive value indicates good flow and a negative value indicates turbulent flow. Larger values indicate stronger association with the particular flow classification for M82.

manufacturing unit state and state transition process, and the development of system-level performance monitoring methods that illustrate system-level interactions by capturing the relationships between manufacturing unit and the impact of their states on the flow value of a given machine. These interactions were demonstrated using different visualization techniques. This work provides a framework for investigating machine and system-level behaviors. Future work will explore methods for harnessing this information to derive mitigation strategies that prevent the loss in production performance.

CHAPTER 4

Movement and Behavior of Foraging Humpback Whales

Behavioral monitoring for whales in the wild is of great importance in whale population conservation and management. Such monitoring helps develop an understanding of whales foraging ecology and aid in the design of strategies to mitigate threats from the environment. This chapter introduces a ethogram based behavior classification algorithm to analyze the behavior of humpback whales in their natural habitats and to identify potential risks associated with fishing gear entanglement and vessel collision. Biologging tags enable observations of underwater animal behavior when direct visual surface observations are not possible. Since 2004, the Stellwagen Bank National Marine Sanctuary whale tagging project has deployed over 250+ suction-cup tags on humpback whales in the Gulf of Maine. This unique, long-term dataset contains focal follow data for most deployments. The proposed ethogram examines depth, dive shape, pitch, roll, heading, fluke rate, speed and jerk during each dive to classify the whale's behavior and cross examined the results with focal follow observations. Activity budgets and temporal and spatial distributions of different behavior are produced to better understand long-term behavior pattern of humpback whales in the wild. The ability to repeatably and reliably classify behaviors from tag data has implications for the study of activity budgets and energy expenditures, as well as future tagging work. The kinematic feature analysis provides insights for identifying environmental impact on whale behavior in the wild.

4.0.1 Introduction

It is crucial to understand when and where humpback whales spend their time, and to investigate how their behavioral state may affect risk. This information supports regulatory decisions made to reduce our impact on these animals. This regulation can include redirecting shipping lanes away from areas where the animals aggregate, and helping to create guidelines for ships when whales are present, like speed reductions. Humpback whale behavior has been studied using land-based

and boat-based observations. For example, Noad et al. used shore based observations to estimate humpback whale population size during migrations along the east coast of Australia [38], and Derville et al. investigated spatial movement patterns of humpback whales using visual observations from land and small boats during focal follows [39]. Parameters like average swimming speed and dive duration can be estimated using these approaches. But these methods are limited to daylight hours, and only collect information about behavioral states when the animals are at the surface.

Developments in bio-logging tag technology have enabled the collection of fine-scale data including location, acoustic, kinematic and physiological information from free swimming wild animals. These tag data have provided key insights into animal behaviors in their natural habitat. For example, Stimpert et al. used acoustic tag data to quantify calling rates of fin whales, a parameter required for accurate estimates of population densities from passive acoustic monitoring data [42]. Features from tag data have also been used to identify foraging. Goldbogen et al. identified lunge feeding events using features like changing rates of acceleration (jerk), estimated speed, body orientation, and the depth of the animal [43]. Researchers have then used these features with data driven approaches to automate behavioral classification. Allen et al. designed a decision tree algorithm to detect foraging lunges from tag kinematic data [44]. These studies have investigated behavior using tag data, but have tended to focus on a particular behavior (e.g. foraging), and have not presented behavioral budgets for the entire tag record.

Here we address this gap by using bio-logging data to establish baseline information on movement and behavior of Stellwagen bank humpback whales. The specific contributions of this chapter include:

1. A behavioral ethogram for humpback whales;
2. Identified features from tag data streams for behavioral state classification (e.g. foraging, traveling, resting...). These features were used with a supervised approach to analyze tag data for behavior classification;
3. Behavioral budgets and state transition map;
4. Identified potential risks associated with different behaviors.

4.0.2 Biological System Example

The waters in and around the Stellwagen Bank National Marine Sanctuary are an important habitat for many animals. Humpback, fin, sei, minki, and right whales use this area at different times throughout the year [68] These animals are a key part of the marine ecosystem, and biologging tags

have regularly been used to investigate questions related to behavior, biomechanics, and foraging ecology. Data from tags are also important for understanding how human activity (e.g. boat traffic and commercial fishing) impact these animals [69] In this work we present the classification and analysis of day-scale behavior of humpback whales using biologging tag data to improve our understanding of how these animals use their environment.

Stellwagen bank and the surrounding waters are essential feeding grounds for humpback whales [70]. In the summer months, animals regularly forage for krill and sand lance. The animals use both individual and cooperative foraging strategies. During these foraging events the animals use the entire water column, with bubble net and lunge feeding observed directly at the surface, and individual and cooperative feeding behaviors identified at the sea floor using tag data [51][71] These foraging areas are also hot spots for commercial fishing and whale watching, with boats and fishing gear regularly observed in close proximity to feeding animals. The impact of these interactions ranges from entanglement in fishing gear to disturbance during foraging. Fatal collision between ships and whales has been reported since the 1800s and historical data shows that humpback whales are among the most vulnerable great whales to ship strikes [72] [73][74]. Entanglement has also frequently been observed in the area, with 56 incidents of fishnet entanglement caused injuries recorded for humpback whales in the Gulf of Maine from 2013 to 2017 [75]. These adverse events occur in specific areas of the animal's environment. Ships strikes occur when the animals are at or near the surface, while entanglements can happen throughout the water column.

Figure 4.1 shows the layout for group interactions, and Figure 4.2 lists the states that describe humpback whale behavior.

4.0.3 Methods

The goal is to classify humpback whale behavior in the wild with bio-logging tag data, Figure 4.3 describes an overview of our method, which consists of the following three operations:

1. *Feature extraction*: Tag data were first calibrated and then converted to whale frame, features relevant to whale movement were extracted.

2. *State classification*: An ethogram based approach was developed and whale behaviors were classified into foraging, swimming, resting and other states.

3. *State characterization*: Activity budgets and mapping of behaviors were developed to understand whales' behavior at day scale. Distribution of depth and rolling events was analysed to determine the risk of ship strike and fishing gear entanglement accidents associated with each behavior

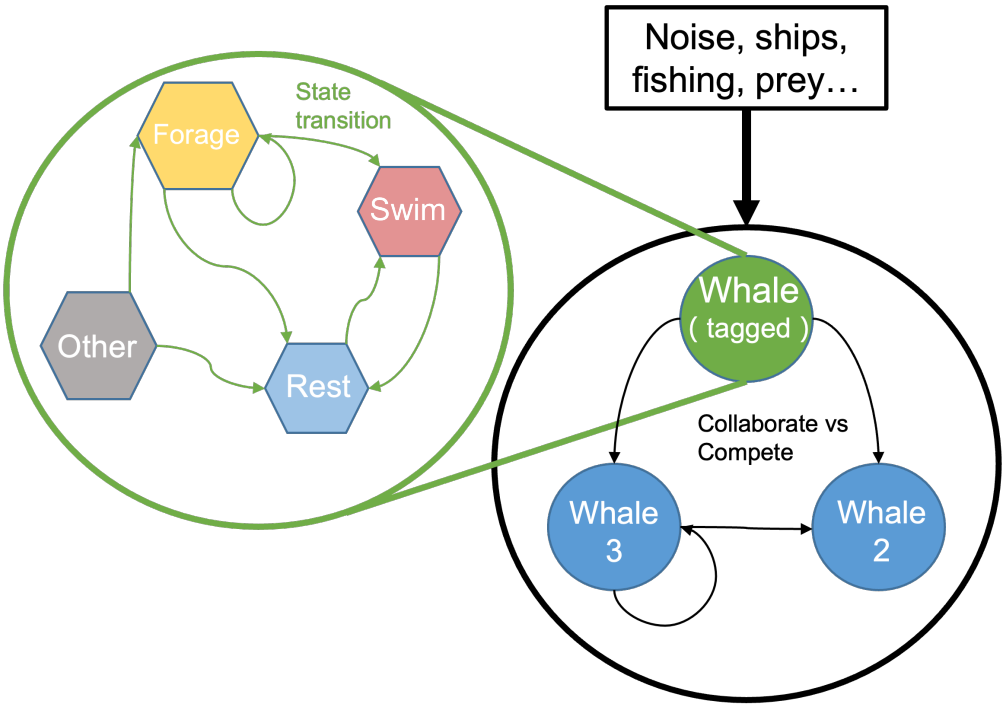

Figure 4.1: Structure of a biological system, behavior states of one whale can be affected by other whales' behavior and environmental conditions. The aim of this work is to develop methods to study the behavior of individually tagged whales.

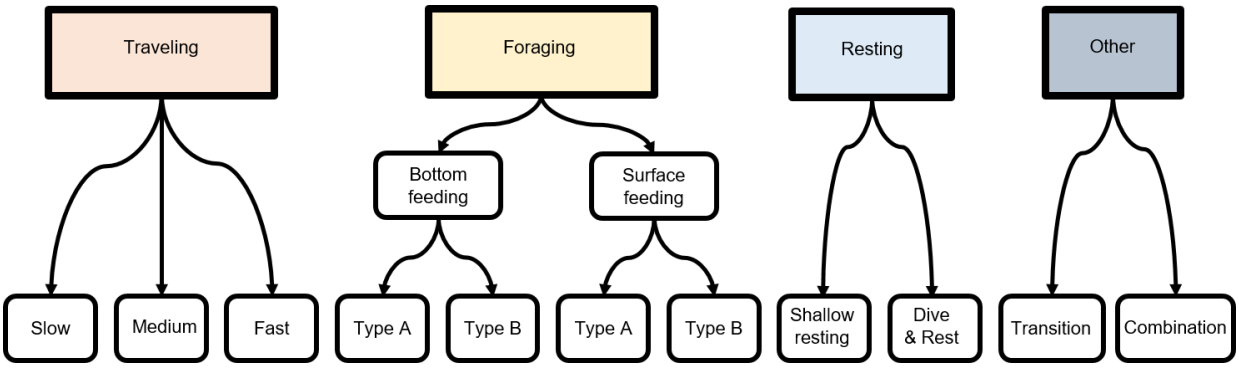

Figure 4.2: Humpback whale behavior states and substates.

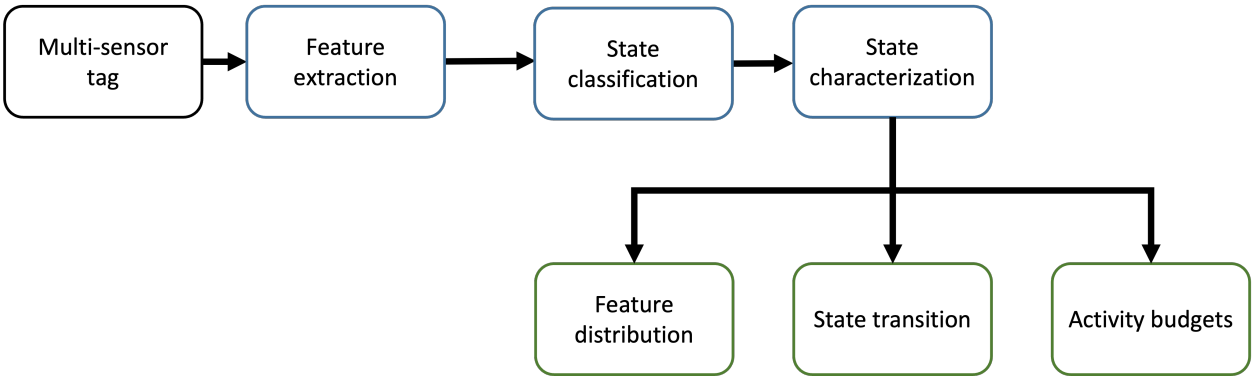

Figure 4.3: Workflow of proposed movement and behavior analysis framework

4.0.3.1 Data Collection

Data for analysis was collected in June 2018 and 2019 around Setllwagen bank (42.4539° N, -70.3361° W) and the Great South Channel off the coast of Chatham MA (41.6575° N, and -69.8685° W). Water depths range from 20-30m deep on the bank with depths of around 100 m in the surrounding water. DTAGs were secured to the animals using suction cups to measure acoustics and kinematics [76]. On board sensors included GPS, hydrophones, a pressure sensor, a temperature sensor, accelerometers and magnetomoeters. Kinematic sensors were sampled at rates of up to 250 Hz, acoustic sensors at rates up to 240 kHz, and the GPS was sampled when the animal was at the surface.

This study included 10 humpback whales, 4 females, 2 males and 4 with unknown gender. Whales were tagged using hand poles from small boats between 10 am and 2 pm. Tags were placed on the back of the animals near the dorsal fin, but orientation and location tended to shift during a deployment. Visual observations of animal behavior were made before the tagging event, and a 1 hour focal follow was conducted after the tag was attached. During the focal follow, photos for identification were taken and behavioral states of both tagged, and conspeific animals were recorded. Weather, sea state and the number and range of vessels in the area were also logged. Four trials were conducted in 2018 and the other six trials were conducted in 2019. Two trials had complete kinematic data but no GPS recording. The average duration of the 10 tag deployments was 25.2 hrs.

4.0.3.2 Feature Extraction

Tag sensor data were first converted to engineering units. Temperature data from the deployment were used to identify parameters for a second order polynomial to correct pressure sensor measurements for temperature affects. The relative orientation of the tag was aligned virtually' with the body of the animal (whale frame), and orientation (pitch, roll and heading) was calculated using

the approach proposed by Zhang et al. In the whale frame, the x-axis was aligned along the antero-posterior length of the animal, y-axis laterally across the body, and z-axis along the dorsoventral length of the body.

In addition to pose, orientation rates, jerk (time rate of change of the accelerometer measurements), speed and relative pitch were calculated for behavioral classification. Relative pitch was calculated using a high pass filter with pitch data to identify bouts of fluking and calculate fluking frequency. Orientation rate was calculated numerically from pose data. Jerk was calculated to capture changes in acceleration created by sudden changes in animal movement that can be indicative of foraging.

$$J = \sqrt{\frac{\partial a_x}{\partial t}^2 + \frac{\partial a_y}{\partial t}^2 + \frac{\partial a_z}{\partial t}^2} \tag{4.1}$$

Speed was also estimated using GPS data and flow noise. Speed and heading were calculated from consecutive GPS positions.

$$v_{GPS} = \frac{distance}{time} \tag{4.2}$$

Flow noise and data from the pressure sensor were used to estimate the speed of the animal underwater [43]. During dives with a pitch greater than 30 deg, a second order polynomial was used to estimate the relationship between the energy in the low frequency flow noise and the speed estimated using the animal orientation and vertical component of the animal's velocity calculated from the pressure sensor.

$$v_{flow} = \theta_1 * E_{flow}^2 + \theta_2 * E_{flow} + \theta_3 \tag{4.3}$$

A dive was identified when the animal was underwater for more than a minute at a depth greater than three meters. Periods at the surface where identified when the pressure sensor data was less than one meter. Dives were segmented into three phases: descent, bottom and ascent [77]. Dive shape was classified as U-shaped or V-shaped based the percentage of time at the bottom, and shape was a feature used to classify foraging. Dives with a bottom percentage higher than 50% were labeled as U-shaped, otherwise they were labeled as V-shaped, Figure 4.4. .

4.0.3.3 Segmentation and Classification

Tag data were used to identify three main behaviors: traveling, foraging and resting, Figure 4.2. Unclassified data was grouped together as 'other.' Supervised classification using features described in Table 4.1 were used to label the data. Travelling was characterized by a consistent heading, shallow dives, and a neutral roll angle. A fluke and glide gait pattern was also characteristic of this behavior. Average speed was estimated from GPS data, and heuristic thresholds were used to further subdivide traveling into periods of Slow, Medium and Fast travel. Resting behavior

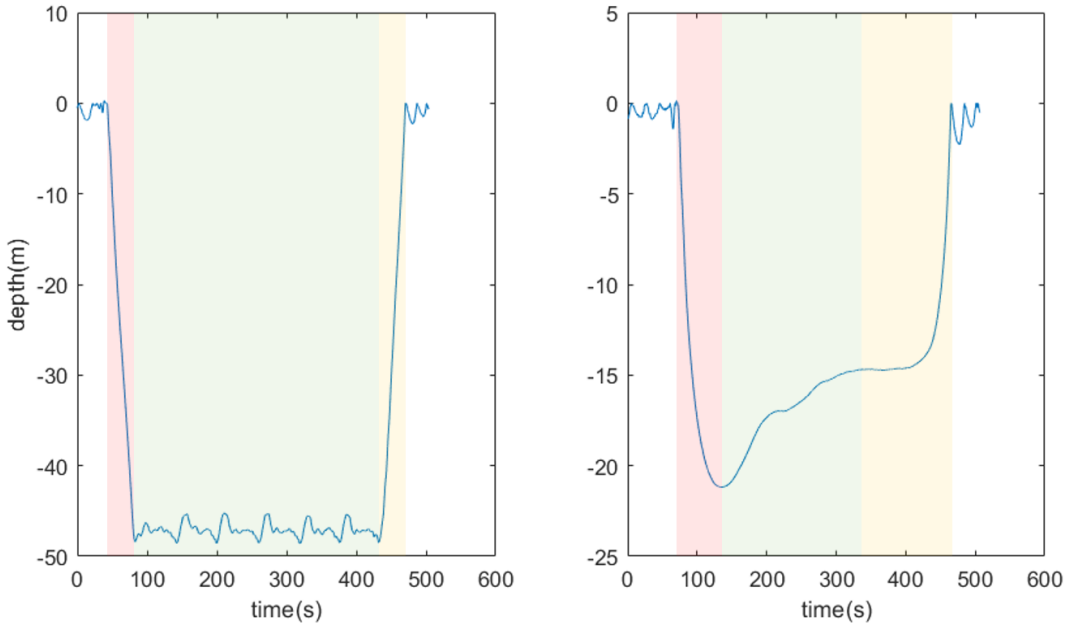

Figure 4.4: A u-shaped dive (Left) and a V-shaped (Right). The red, green and yellow boxes indicate descent, bottom and ascent phases respectively. In this example, the animal spent 81% of the dive at the bottom during the u-shaped dive and 48% of the time during the V-shaped dive.

was observed during both shallow dives($maximum\ depth < 5m$) and during a mix of shallow and medium dives ($5m \leq maximum\ depth < 20m$). During these dives the animals held a consistent pose with little to no fluking during bottom of the dive.

Foraging was characterized by high jerk and roll events, and was segmented by a subject matter expert using features in the dive profile, orientation and jerk data. High jerk events occur during impacts when the animal closes their mouth, or during large changes in speed when the animal mouth opens to engulf prey. Large changes body pose, like large roll angles, also occur during foraging. A threshold of 80 g/s was used to identify jerk events. Foraging was then divided into surface feeding and bottom feeding using pressure sensor measurements. Figure 4.6 shows the difference in kinematic features for bottom feeding (a) and bottom feeding (b). Both types of feeding were characterized by deep U-shaped dives and similar average speeds. Roll events were observed during the during type (b), but were not present during bottom feeding (a). The average pitch rate was also higher for type (a) bottom feeding.

The animals were frequently observed bubble net feeding at the surface. In the tag data this foraging strategy was characterized by shallow dives with large roll angles and circular swimming patterns. Dives with these features were classified as Type A shallow foraging dives. Lunge feeding was also observed in the data. Lunges were characterized by a relatively constant heading, large changes in depth during the bottom of the dive, increased fluking frequency during the change

Table 4.1: An ethogram listing features associated with different behavior classes and sub behaviors shown in Fig 4.2

Behavior Class	Sub Behavior	Sub Behavior Type	Feature Streams	Class Features	Sub Behavior Features
Travelling	Slow swimming		Heading, pitch and speed	Steady heading, and fluke and glide gait pattern	Average speed < 1.5m/s
	Medium swimming				1.5m/s < Average speed < 2m/s
	Fast swimming				Average speed>2m/s
Foraging	Bottom foraging	Type A	Depth, rolling, jerk and dive shape	Medium to deep dives, rolling and high jerk events happening at bottom of a dive	Rolling throughout dive
		Type B			U-shaped dive, continuously rolling at the bottom of a dive
	Surface foraging	Type A	Depth, rolling, jerk, heading and acceleration	Medium to deep dives, rolling and high jerk events happening at bottom of a dive	Swims in circles
		Type B			Accelerates during descent and decelerates during ascent
Resting	Shallow resting		Orientation rates, depth, acceleration, jerk, speed and pitch	Low angular rates, small acceleration and jerk, low speed, and lack of fluking	Shallow dives
	Dive and rest				Mix of shallow and deep dives

in depth, and a deceleration during engulfment. Bottom foraging occurred during the longer U-shaped dives, and was characterized by a series of rolling and jerk events. Figure 4.5 presents representative data from a series of U-shaped dives from whale 175i with the the roll and jerk events used to classify bottom foraging.

4.0.4 Results

4.0.4.1 Behavior classification - Representative Results

Fig 4.7 presents behavioral classification mapped on to tag data for whale 175i. The deployment lasted 21.4hrs, and the GPS track presented in 4.8 indicates an overground travel distance of 70 km. During the deployment the whale had three long bouts of bottom feeding(A, I, and K). Jerk events were present during the three periods of bottom foraging, but the rolling events were only observed during periods A and K. The Maximum dive depth was 45 ± 4m and ranged from 37 to 56 m with an average dive duration of 4.9 min. The bottom foraging in section I took place at night and had an average of 4.7 roll events per dive. Bottom foraging during the day in sections A and K had an average of 2.5 roll events and 7.3 jerk events per dive.

Periods of travelling were identified during both day and night time. Traveling dives had an average depth of 17 ± 13 m and average time of 173 ± 87 s. Heading tended to be relatively consistent during traveling and the animal used both continuous and fluke and glide gait. GPS estimated speed during the observed 'slow' swimming was 1.3 m/s.Traveling tended to occur between bouts of foraging (both bottom and surface), as well as before and after the resting periods. As the animal moved from a foraging state to traveling there tended to be short periods of transition. During these transitions dive duration, depth, bottom time percentage, and GPS location were comparable to the bottom feeding dives, but the roll and jerk event rates were significantly

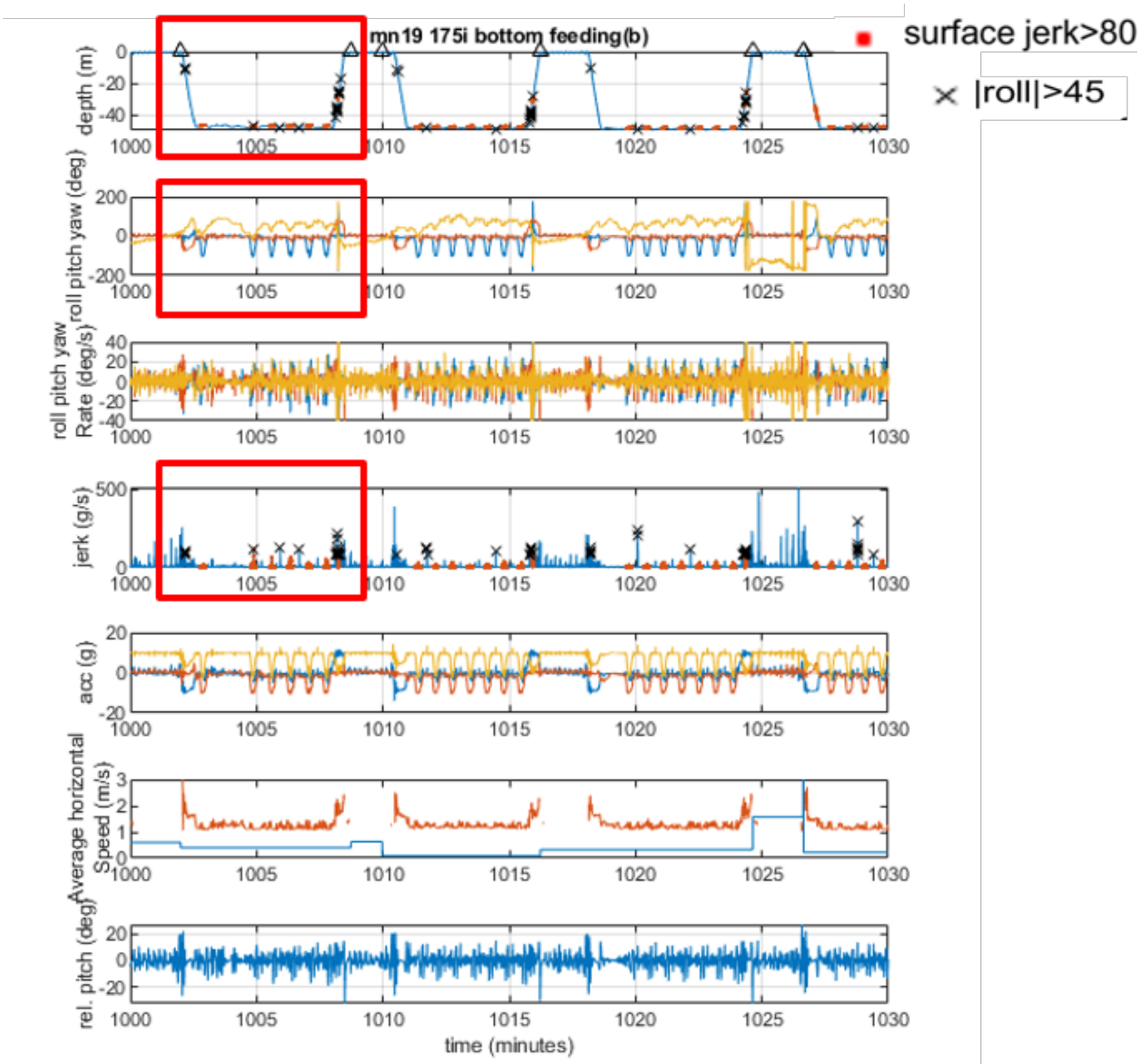

Figure 4.5: Depth, orientation, orientation rates, jerk, acceleration, speed and relative pitch from a tagged humpback whale during bottom feeding(b) sub behavior. Circled area shows an example dive, U-shaped deep dive, continuous rolling and high jerk bursts are features of a bottom feeding(b) dive.

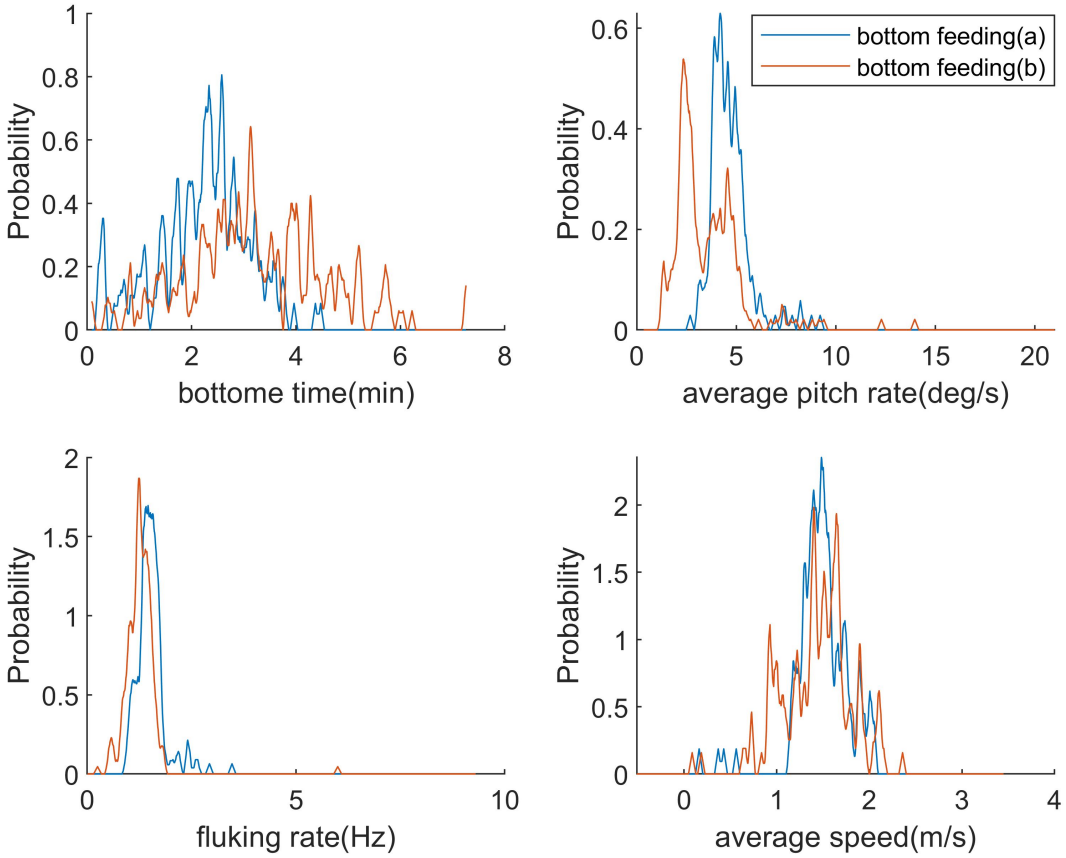

Figure 4.6: Distribution of bottom time, average pitch rate, fluking rate and speed during bottom feeding(a) and bottom feeding(b). Bottom feeding(b) has longer bottom time, smaller pitch rate and fluking rate, and slightly slower speed.

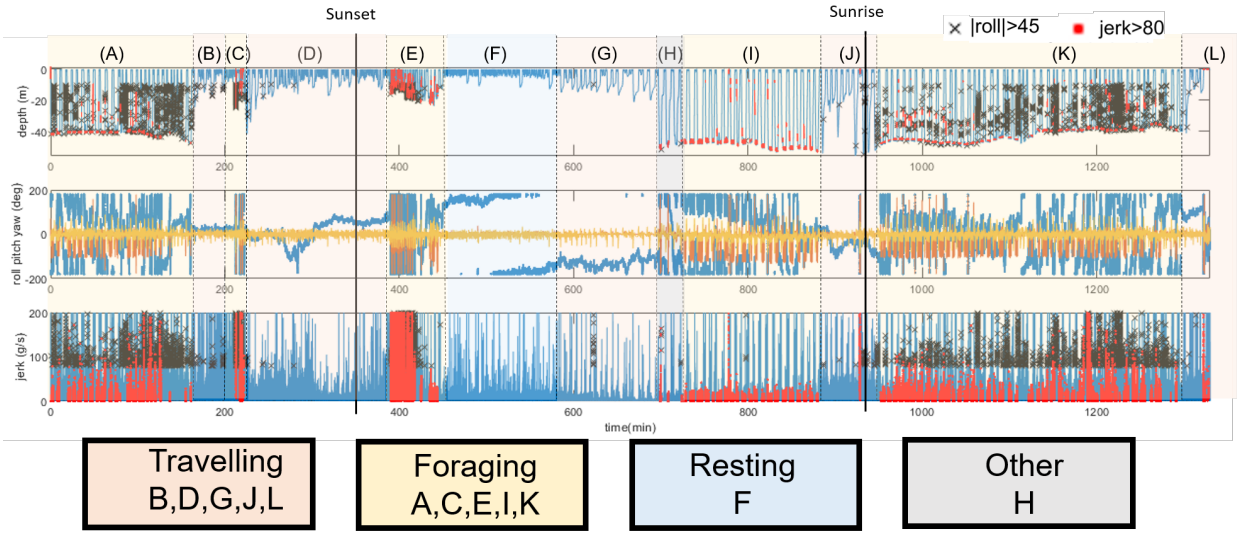

Figure 4.7: Classification results for a whale during a 21hr trial. Data streams include depth, orientation and jerk, red dots indicate jerk events higher than $80m/s^3$. Black crosses indicates roll events greater than 45^o. Periods of foraging have high jerk and high roll events. Travelling was classified using depth, yaw rate, and resting was classified using features in the depth and jerk data.

reduced. For example, as the animal transitions between foraging (k) to traveling (L) the roll rate per dive goes from 5.0 to 0 and the jerk rate also decreases from 2.3 to 0.

Resting (F) was identified during the overnight portion of the data. The resting occurred well after sunset, following a period of surface foraging (E). The period lasted 2.1 hrs (27% of he night), and was characterized by shallow dives (9 $\pm$ 6 m) that lasted an average of 2.0 minutes. While most dives (N) in this section were around 5m deep, there were a M dives to around 10m deep. After the period of resting, the animal spent 119 min traveling 5.6 km to a bottom foraging site.

GPS data provided important context for the tag sensor data and behavioral state. Figure 4.8 presents the GPS track of whale 175i. The animal was tagged at 13:54 pm, and the tag detached from the animal around 10 am then next morning. The animal spent the next three hours bottom foraging in a six square kilometer area near were the tag was attached. Bottom foraging was followed by a period of traveling where the animal swam 5 km north and spent a short time foraging at the surface. The rest of the night was spent resting, surface feeding, traveling, and bottom feeding. Just before sunrise, the whale arrived back to the area where the tag was deployed the day before, and began bottom feeding again.

4.0.4.2 Behavior classification - Summary Results

This study includes analysis of 253 hrs data from 10 animals. Tag deployments lasted an average of 25.3 hrs, with 69% of the data recorded during the day and 31% at night. As with the representative

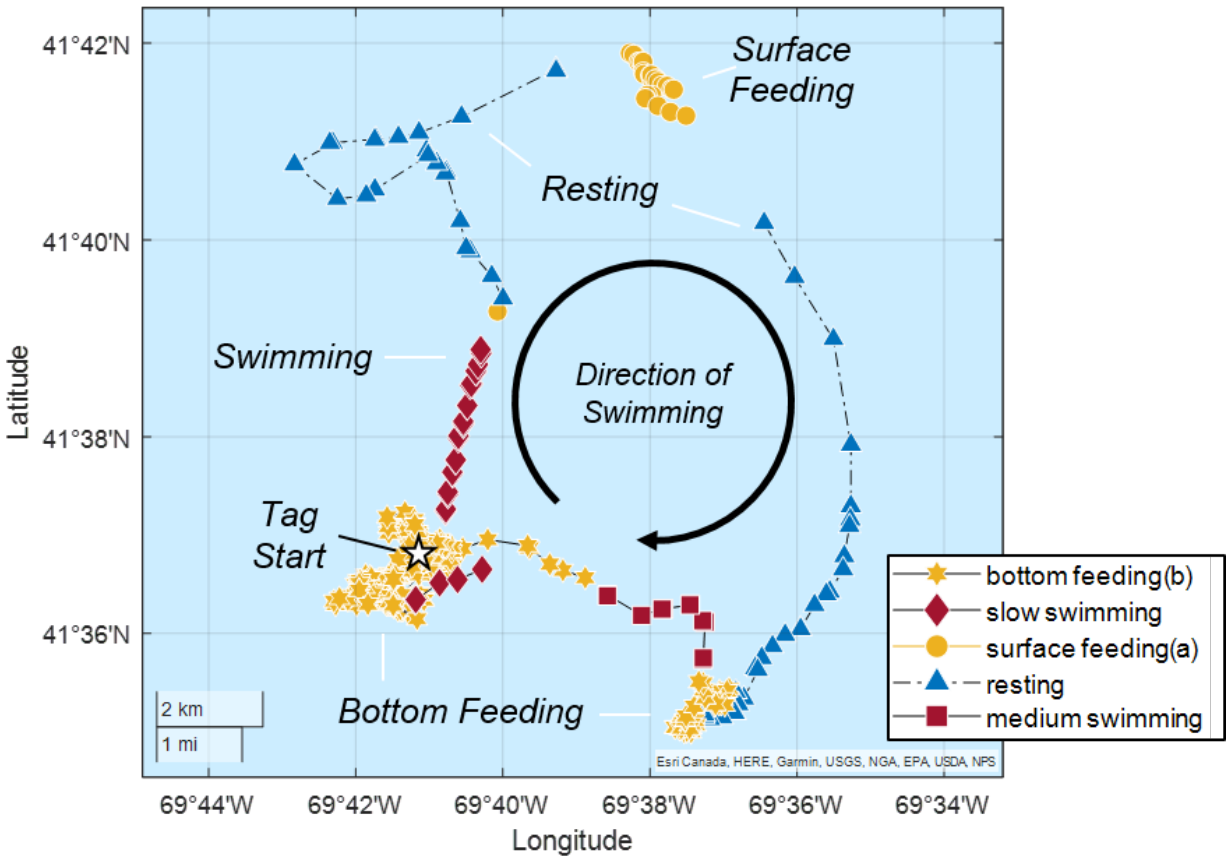

Figure 4.8: Representative GPS track of an individual during a 21 hr deployment. The color and the shape of the markers indicates behavior state. The swimming direction was clockwise.

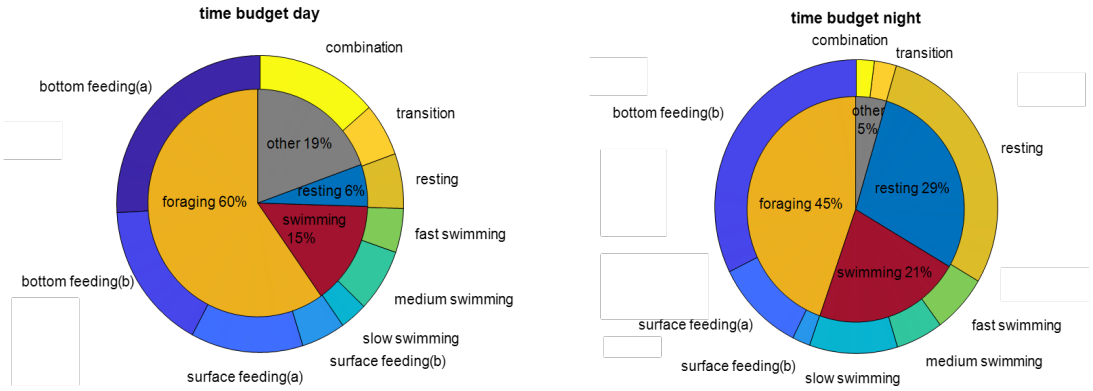

Figure 4.9: Time budget for 10 tagged humpback whales during day and night, where night is defined as 9 pm to 4:30 am local time. Night time accounts for about 38% of total time recorded.

data, behavior were classified into travelling, foraging, resting and other behavior, Table 2 and Figure 6. The whales spent 55% (139 hrs) of their time foraging (surface and bottom foraging), with 60% of the time (83 hrs) spent foraging during the day and 40% (56 hrs) at night. Bottom feeding (type A) accounted for 26% of the foraging but was not observed at night. The percentage of time spent traveling was slightly higher at night than during the day (22% vs 15%), but the animals tended to swim slowly at night. In addition to the slow night swimming, the whales spent more time resting at night than during the day, 51%, compared to 21%. Swimming speeds ranged from 0-3m/s, with an average speed around $1.5m/s$.

GPS data provided important context for tag data. GPS tracks for eight of the animals are presented in Figure 4.11, two deployments in 2018 had no GPS log. The 8 tagged whales tended to stay in the vicinity of the location where the tags were deployed (15km radius circle). In three of the data sets, the whales made a relatively long excursions away from and then back to the tag deployment location. The longest of which covered a distance of 118 km.

Where the animals tend to spend time in the water column is important information for mitigating threats of entanglement and collision. The deepest dives occurred during bottom feeding behavior and the shallowest dives happened during resting and slow swimming. Figure 4.10 shows that the most common depth for both type A and type B bottom foraging is around 40m. The peak less than 10m suggested that there were shallower dives during a bottom feeding section between deep dives. During these dives the animals are going to the sea floor and rolling during foraging events, increasing the risk of entanglement in bottom mounted fishing gear. The animals tended to dive to two depths, 15m and 4m, during resting. During this behavioral state the animals may be

Behavior	Dive depth (m)	Dive Duration (min)	Bottom time (min)	High roll time (min)	High jerk time (min)	Fluking rate (Hz)	Mean speed (m/s)
Bottom feeding(a)	40	3.0	2.2	0.6	0.1	1.5	1.5
Bottom feeding(b)	45	4.3	3.2	0.9	0.1	1.2	1.4
Surface feeding(a)	18	1.2	0.7	0.2	0.1	1.5	1.8
Surface feeding(b)	11	0.9	0.7	0.1	0	1.3	2
Resting	10	2.3	1.9	0	0	1.2	1.5
Slow swimming	9	1.3	0.9	0	0	0.4	1.4
Medium swimming	12	2	1.4	0	0	0.3	1.4
Fast swimming	12	1.3	0.9	0	0.2	2.6	2.2
Transition	25	2.1	0.9	0.1	0	1.2	1.5
Combination	15	1.1	0.6	0.1	0.1	1.6	1.9

Table 4.2: Kinematic features of each classified behavior. High roll time is duration of when roll angle is greater than 45 and high jerk time duration of sections when jerk is greater than 80.

less aware of their environment, and the shallow dives are not deep enough to avoid collisions with large ships.

4.0.4.3 State transition map

The labeled behavioral states were used to construct a state transition map, Figure 4.12. The swimming was an important transition between the observed behaviors, as a change in behavior was was often associated with a change in location. Bottom feeding was the most common state (39 percent), and the animals were most likely to be in a swimming state as they transitioned into and out of the bottom foraging state. Average duration of a bottom feeding bout was 4.7hrs, the longest among all behavior states. In contrast, the shortest state, swimming, averaged 1.4hrs. Resting was the least common observed behavior accounting for only 9% of the total recorded time.

4.0.5 Discussion

4.0.5.1 Activity budgets and site fidelity

This study includes data from 10 humpback whales in the Gulf of Maine during June in both 2018 and 2019. As tag data were recorded during the summer, when feeding is routinely observed, it is not surprising that feeding was the most common behavior. Bottom feeding was more frequent than surface feeding, possibly due to sand lance abundance in this area. These whales spent the 60% of their time foraging during the day, and 45% of their time foraging at night. A majority of he

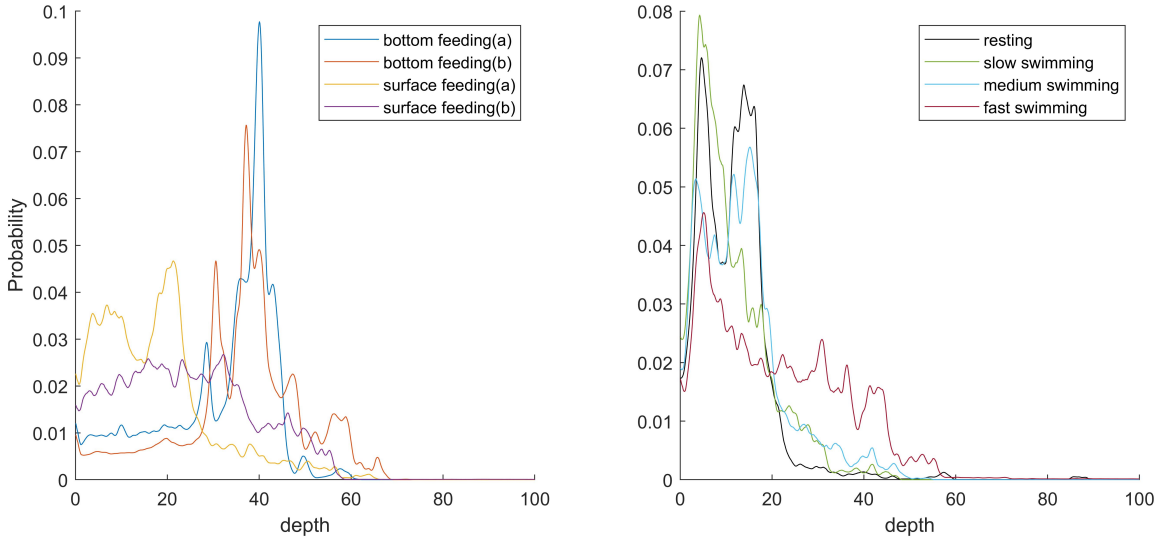

Figure 4.10: Distribution of dive depth for different behavioral states. This figure shows the most frequently visited depth for each behavior. The left figure shows that bottom feeding mostly happens around 40m, and surface feeding(a) and surface feeding mostly happens around 25m and 5m respectively. The right figure shows that the most swimming and resting behavior happens at depth shallower than 5m, while medium swimming and resting present a small peak around 15m.

feeding behavior for all whales tagged in this project took place within a radius of 15km at depths that ranged from 40m to 60m. The continuity of the animal locations between years indicates that this location has important environmental features that may be related to prey abundance. At this site, over 60% of the foraging is bottom foraging due to burrowing sand lance being their major food source. At night, time spent in bottom feeding(b) increases while time spent in bottom feeding(a) decreases to zero. This indicates that humpback whales adopt different feeding strategies depending on the time of day. As bottom feeding time is when the animals are most vulnerable to fishnet entanglement, actions should be taken to mitigate this threat. Whale-safe fishing gear and virtually-marked fishnet attachments were proposed to reduce entanglement caused injuries [78].

The state transition map shows that the most frequent transition was between bottom feeding and swimming, while the second most frequent transition was between surface feeding and swimming. During transitions between behaviors, observed features were often a combination of those observed in the two states. These transitions were more common during surface feeding than during bottom feeding, suggesting that surface feeding has more ambiguous kinematic features while the movement patterns during bottom feeding are more consist. Transition periods were labeled as 'other' and accounted for 19% of day time and 5% of night time data. This lower transition percentage at night is indicative of a reduction of activity. the larger percentage of resting, 29% of the time spent at night compared to only 6% during daytime, also supports this observation. The

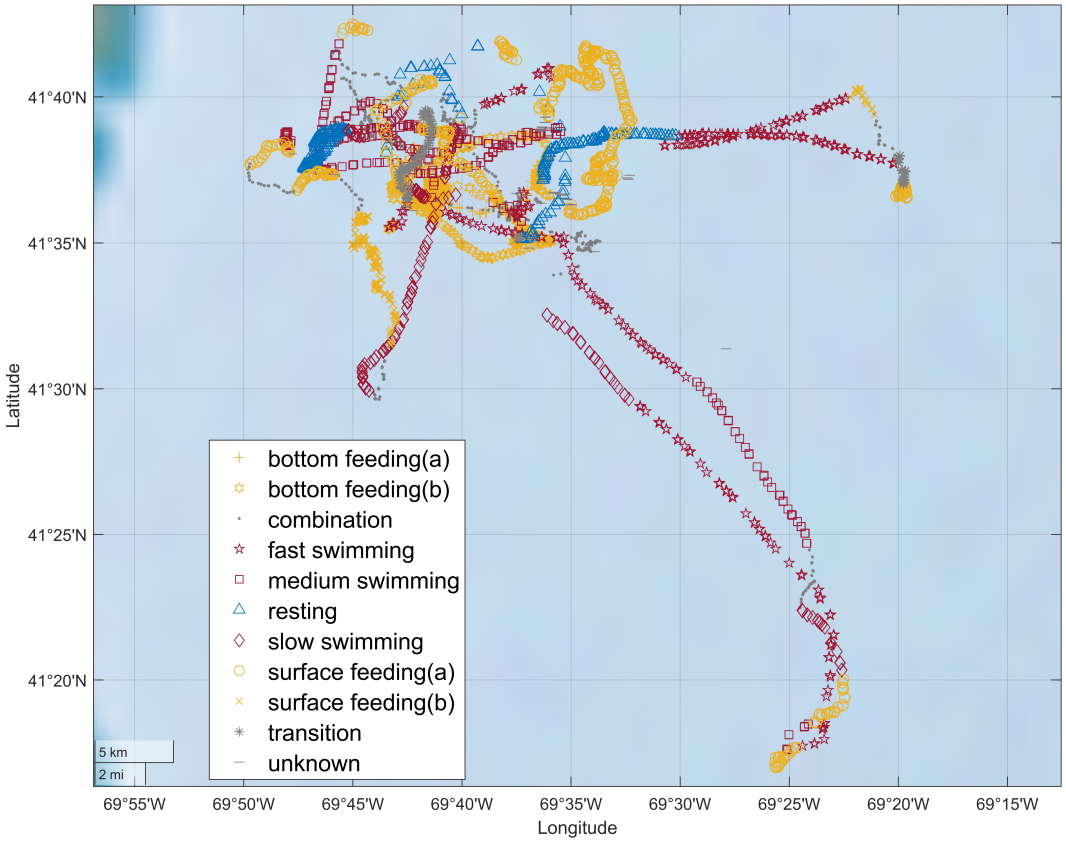

Figure 4.11: GPS tracks of all 10 whales, color coded with behavior states. Blue, red, yellow and grey indicates resting, swimming, foraging and other behavior respectively.

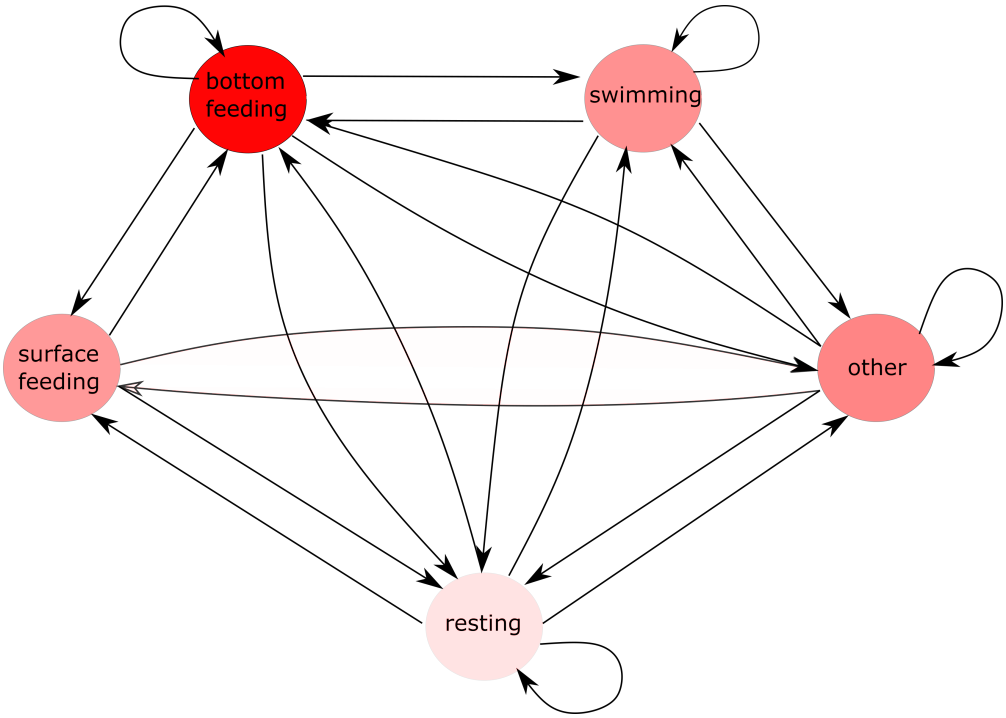

Figure 4.12: State transition map of all whales' data. A state is one continuous section of the same behavior, the size of the node represents the total length of that behavior. Each arrow represents a transition, and thickness of the line indicates the frequency.

state transition map and the GPS tracks indicate that resting states occurs around surface feeding states. Resting states were observed away from bottom foraging locations, and the whales did not appear to have a preference for resting location, posing a challenge for policies to mitigate ship collisions.

4.0.5.2 Kinematic characteristics

Dive profiles along with roll and jerk events were key features for the classification of foraging behavior. During foraging, whales swim rapidly towards their prey (increased fluking frequency, amplitude and speed), and then engulf large amount of prey-laden water. The increase in speed as he animals move towards prey, and the rapid deceleration once the mouth is opened result in large jerk events. Maximum dive depth and depth of the rolling events were important features that were used to separate surface feeding and bottom feeding. Maximum dive depth for most bottom foraging dives was over 40m, while most depths during surface foraging were lower than 20m. Tag location on varied from animal to animal and affected the magnitude of the observed jerk signal. In this work we used a single threshold value to identify events, but future work could explore thresholds that take this affect into account.

4.0.6 Conclusion

In this paper, an ethogram based behavioral classification scheme was developed and used to analyze tag data collected from 10 humpback whales in the Gulf of Maine. Kinematic features of each behavior were analysed to enhance understanding of underwater behavior. Time and location of resting and foraging behaviors were investigated to monitor the risk associated with different behaviors. This work focused on understanding humpback whales' behavior on a day-scale and identified potential threats from fishing gear entanglement and vessel collision. Future work will explore actions that could be taken to mitigate these threats.

CHAPTER 5

Dissertation Impact and Future Work

In complex systems, the challenges associated with characterizing agent and system level behaviors depend on different properties of the system. In a manufacturing system, the agent state are generally predefined, the control logic of each agent is predetermined, and the industrial internet of things(IIOT) and cloud computing technologies enable large-scale, real-time data collection from the production floor. This data provides information about the complete manufacturing process and makes testing and validation of our models possible. From a complex system perspective, where machines, robots, and gantries in the manufacturing system are modelled as agents, it is clear that agents within one production line are dependent on each other, through transferring of parts and materials, and on environmental factors including the customer demand, part availability and operator actions. This inherent connectivity leads to the potential for down states to propagate through the system, initiating unexpected downtime in other agents within the system. In order to characterize the behavior patterns of the manufacturing units at the system level, these interactions between agents must be identified, classified, and analyzed during production.

In a biological system, the agents are autonomous individuals with their own principles, which make it harder to obtain their dynamic models. Antoniak et. al. have been able to estimate dolphin whole body dynamics with just acceleration data [79]. A Fusion approach incorporating video data from surveillance cameras instrumented around the pool and tag data from the animal has been developed to investigate the movement track of dolphins [80]. Studies have been done to show the characteristics of dolphin kinematics when responding to different environmental stimuli [81]. These studies show the potential for single agent monitoring in a controlled environment for a biological system. However, when it comes to studying animals in an uncontrolled environment, observations are hard to obtain due to noise and natural environmental blockers. With limited information resources, the challenges of agent-level analysis include determining appropriate feature extraction and feature selection methods that will enable the classification of agent behaviors in unstructured and uncertain environments. Furthermore, once the behavioral states are classified, there are opportunities to leverage this information to evaluate how environmental disruptions impact behavioral states, and determine the implications for animal conservation and management.

To address these gaps, the work in this dissertation presented a unifying framework that could be applied to both manufacturing and biological systems, and leveraged tools from IIoT, biologging tags, data-driven modelling and machine learning to analyze agent-level behavior and interactions between agents. This work will enable more informed decision making with respect to improving system performance and mitigating environmental challenges.

In Chapter 3, a data-driven framework was proposed for improving manufacturing productivity by monitoring machine and system level interactions. This work characterized agent-and-system-level behavior through the classification of machine states and the introduction of a novel key performance indicator termed a "Flow" metric, which is a system-level KPI. The classification of machine states provides an agent-level analysis tool, while the flow metric tracks the performance of critical machines and reflects the system's overall performance. Agent-level performance monitoring was achieved through a probabilistic automaton modelling of machines in a production line. This showed the frequency and duration of each behavioral state and the transition frequency between them. Tracking the deviation of these parameters indicates when the machine's performance is changing. Furthermore, a framework for quantifying the interactions between agents in a complex system was demonstrated through a case study based on a real-world large-scale automotive production plant. A logistic regression model was used to model and simulate system-level behavior and investigate how machine behaviors impact each other. This work demonstrates that the framework can be applied to a manufacturing system to characterize agent to agent interactions.

In Chapter 4, the proposed framework was applied to a biological system involving humpback whales to study their behavior in the wild and to investigate how their behavior is associated with risk factors including vessel strikes and fishing gear entanglements. This work proposed an ethogram based approach to classify humpback whale behaviors using features derived from tag measurements. Kinematic analysis was carried out to distinguish the features of different behaviors. State characterization showed the impact to behavioral states associated with environmental factors. This outcomes of this work include a framework for feature selection in agent state classification when direct observation is unavailable, and a method for quantifying how environmental factors impact agent behaviors. This work provides insights on which behavioral states of humpback whales are associated with increased risks and when and where humpback whales are most susceptible to threats including fishing gear entanglement and vessel strikes. This work contributes to long-term agent behavior monitoring and risk analysis.

The application in different domains shows the extensibility of the proposed framework. While the manufacturing system and biological system studied in this dissertation had many different attributes, similar tools and techniques could be applied within this framework. In the classification of behavioral states of a single agent, statistical methods were applied to both systems to find the range of a features for different states. And during single-agent analysis, a Markov chain model

was applied to develop the state transition maps for both manufacturing and biological systems.

5.0.1 Future Directions

The ability to transition the knowledge from a purely observational modality to decision making is an open research questions in many complex system domains. This is particularly true for the manufacturing environment from Chapter 3. While this framework allows the identification of agent interactions, an understanding of how these insights can be used to improve system performance is less clear. On the production floor, decision variables including maintenance schedules, buffer sizes, and operator actions can all affect the behavior of the individual agents as well as the system. Exploring how to choose these variables would help improve the efficiency of the system. Future work in this domain should consider: (1) Investigating how to arrange maintenance schedules to minimize unexpected downtime as a function of the behavioral states of interconnected machines. (2) Studying how the system would perform under varying disruptions such as changing customer demands, and new product developments. (3) Exploring the operators' action and how it can affect system performance. (4) Developing a control protocol that considers risk and adaptability.

In Chapter 4, features for classification of whale behavioral states were selected on a trial and error basis. Correlation tests were conducted for each feature based on focal follow observations. In order to expand the framework to archival data sets, an automated feature selection method needs to be developed. Additional testing and validation is needed to make sure that this method works for whales not tagged before. Results from this research provide initial insights into the differences between the dynamics of different behavioral states. However, there are additional details that need to be addressed. For example, one might ask "Why do animals select specific swimming gaits?" A common belief is that gait is chosen to minimize energy consumption and increase efficiency in hunting. In order to test this theory, future work must includes dynamic modelling of humpback whales, investigating drag, energy cost and other gait parameters, to determine how these metrics are related to survival.

Another next step is to investigate the collaboration and competition between humpback whales within a pod. This work was limited to the study of one whale due to limited tagging data availability, in order to investigate multiple whale interactions, extra experiments need to be conducted. This would provide more understanding of the social structure among humpback whales. Pod size of humpback whale can range from 2 to 15 individuals. Tagging multiple whales is challenging due to their unpredictable behavior. Even just a close pair or a small group could provide valuable information about their interactions. As the tag does not directly measure the 3D location of the animal, a first step toward this goal is to develop an automated approach for whale localization with tag data. With this, and the measured starting position of whales, we will be able to reconstruct

the whole trajectory and the relative locations of other whales. The results of these interactions between whales can be used to investigate how humpback whales adapt their behavior in response to environmental threats.

Bibliography

[1] Qawi K Telesford, Sean L Simpson, Jonathan H Burdette, Satoru Hayasaka, and Paul J Laurienti. The brain as a complex system: using network science as a tool for understanding the brain. *Brain connectivity*, 1(4):295–308, 2011.
(Back reference pages: iv, 1, and 2)

[2] Melanie Mitchell and Mark Newman. Complex systems theory and evolution. *Encyclopedia of evolution*, 1:1–5, 2002.
(Back reference pages: iv, 2)

[3] Andrew D Taylor. Metapopulations, dispersal, and predator-prey dynamics: an overview. *Ecology*, 71(2):429–433, 1990.
(Back reference pages: iv, 2)

[4] W Brian Arthur. *The economy as an evolving complex system II*. CRC Press, 2018.
(Back reference pages: iv, 2)

[5] Mark EJ Newman. Complex systems: A survey. *arXiv preprint arXiv:1112.1440*, 2011.
(Back reference page: 1)

[6] Volker Grimm, Eloy Revilla, Uta Berger, Florian Jeltsch, Wolf M Mooij, Steven F Railsback, Hans-Hermann Thulke, Jacob Weiner, Thorsten Wiegand, and Donald L DeAngelis. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *science*, 310(5750):987–991, 2005.
(Back reference page: 1)

[7] James Ladyman, James Lambert, and Karoline Wiesner. What is a complex system? *European Journal for Philosophy of Science*, 3(1):33–67, 2013.
(Back reference page: 1)

[8] Nigel Goldenfeld and Leo P Kadanoff. Simple lessons from complexity. *science*, 284(5411):87–89, 1999.
(Back reference page: 1)

[9] George M Whitesides and Rustem F Ismagilov. Complexity in chemistry. *science*, 284(5411):89–92, 1999.
(Back reference page: 1)

[10] J Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460(7256):685–686, 2009.
(Back reference page: 1)

[11] Seyedali Mirjalili. Genetic algorithm. In *Evolutionary algorithms and neural networks*, pages 43–55. Springer, 2019.
(Back reference page: 1)

[12] G Maione and D Naso. Adaptation of multi-agent manufacturing control by means of genetic algorithms and discrete event simulation. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 4, pages 6–pp. IEEE, 2002.
(Back reference page: 1)

[13] Sanjiv Bavishi and Edwin KP Chong. Automated fault diagnosis using a discrete event systems framework. In *Proceedings of 1994 9th IEEE International Symposium on Intelligent Control*, pages 213–218. IEEE, 1994.
(Back reference page: 1)

[14] Maxwell Toothman, Birgit Braun, Scott J Bury, Michael Dessauer, Kaytlin Henderson, Ray Wright, Dawn M Tilbury, James Moyne, and Kira Barton. Trend-based repair quality assessment for industrial rotating equipment. In *2021 American Control Conference (ACC)*, pages 502–507. IEEE, 2021.
(Back reference page: 3)

[15] Jens Windau and Laurent Itti. Inertial machine monitoring system for automated failure detection. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 93–98. IEEE, 2018.
(Back reference pages: 4, 13)

[16] Yuehua Liu, Wenjin Yu, Tharam S Dillon, Wenny Rahayu, and Ming Li. Empowering iot predictive maintenance solutions with ai: A distributed system for manufacturing plant-wide monitoring. *IEEE Transactions on Industrial Informatics*, 2021.
(Back reference pages: 4, 9, and 13)

[17] Hao Wang, Yassine Qamsane, James Moyne, and Kira Barton. Merging subject matter expertise and deep convolutional neural network for state-based online machine-part interaction classification. In *International Manufacturing Science and Engineering Conference*, volume 85079, page V002T09A006. American Society of Mechanical Engineers, 2021.
(Back reference pages: 4, 13)

[18] Wenjin Yu, Yuehua Liu, Tharam Dillon, Wenny Rahayu, and Fahed Mostafa. An integrated framework for health state monitoring in a smart factory employing iot and big data techniques. *IEEE Internet of Things Journal*, 2021.
(Back reference pages: 4, 9, and 14)

[19] Yassine Qamsane, James Phillips, Clare Savaglio, David Warner, Scott James, and Kira Barton. Open process automation- and digital twin-based performance monitoring of a process manufacturing system. *IEEE Access*, Submitted, under revision.
(Back reference pages: 4, 14)

[20] Mikel Canizo, Angel Conde, Santiago Charramendieta, Raul Minon, Raul G Cid-Fuentes, and Enrique Onieva. Implementation of a large-scale platform for cyber-physical system real-time monitoring. *IEEE Access*, 7:52455–52466, 2019.
(Back reference pages: 4, 9, and 14)

[21] Jiafu Wan, Shenglong Tang, Di Li, Shiyong Wang, Chengliang Liu, Haider Abbas, and Athanasios V Vasilakos. A manufacturing big data solution for active preventive maintenance. *IEEE Transactions on Industrial Informatics*, 13(4):2039–2047, 2017.
(Back reference pages: 4, 10, and 14)

[22] James Moyne, Yassine Qamsane, Efe C Balta, Ilya Kovalenko, John Faris, Kira Barton, and Dawn M Tilbury. A requirements driven digital twin framework: Specification and opportunities. *IEEE Access*, 8:107781–107801, 2020.
(Back reference pages: 4, 14)

[23] Yassine Qamsane, James Moyne, Maxwell Toothman, Ilya Kovalenko, Efe C Balta, John Faris, Dawn M Tilbury, and Kira Barton. A methodology to develop and implement digital twin solutions for manufacturing systems. *IEEE Access*, 9:44247–44265, 2021.
(Back reference pages: 4, 14)

[24] Yassine Qamsane, Chien-Ying Chen, Efe C Balta, Bin-Chou Kao, Sibin Mohan, James Moyne, Dawn Tilbury, and Kira Barton. A unified digital twin framework for real-time monitoring and evaluation of smart manufacturing systems. In *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, pages 1394–1401. IEEE, 2019.
(Back reference pages: 4, 14)

[25] Kung-Jeng Wang, Ying-Hao Lee, and Septianda Angelica. Digital twin design for real-time monitoring–a case study of die cutting machine. *International Journal of Production Research*, 59(21):6471–6485, 2021.
(Back reference pages: 4, 10, and 14)

[26] Emanuel F Alsina, Manuel Chica, Krzysztof Trawiński, and Alberto Regattieri. On the use of machine learning methods to predict component reliability from data-driven industrial case studies. *The International Journal of Advanced Manufacturing Technology*, 94(5):2419–2433, 2018.
(Back reference pages: 4, 10, and 14)

[27] Thyago P Carvalho, Fabrízzio AAMN Soares, Roberto Vita, Roberto da P Francisco, João P Basto, and Symone GS Alcalá. A systematic literature review of machine learning methods applied to predictive maintenance. *Computers & Industrial Engineering*, 137:106024, 2019.
(Back reference pages: 4, 10, and 14)

[28] Olga Fink, Qin Wang, Markus Svensen, Pierre Dersin, Wan-Jui Lee, and Melanie Ducoffe. Potential, challenges and future directions for deep learning in prognostics and health management applications. *Engineering Applications of Artificial Intelligence*, 92:103678, 2020.
(Back reference pages: 4, 10, and 14)

[29] Samir Khan and Takehisa Yairi. A review on the application of deep learning in system health management. *Mechanical Systems and Signal Processing*, 107:241–265, 2018.
(Back reference pages: 4, 10, 13, and 14)

[30] Yaguo Lei, Bin Yang, Xinwei Jiang, Feng Jia, Naipeng Li, and Asoke K Nandi. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138:106587, 2020.
(Back reference pages: 4, 10, and 14)

[31] Rui Zhao, Ruqiang Yan, Zhenghua Chen, Kezhi Mao, Peng Wang, and Robert X Gao. Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*, 115:213–237, 2019.
(Back reference pages: 4, 10, and 14)

[32] Guang Wang and Shen Yin. Data-driven fault diagnosis for an automobile suspension system by using a clustering based method. *Journal of the Franklin Institute*, 351(6):3231–3244, 2014.
(Back reference pages: 4, 10, and 14)

[33] Andreas Theissler, Judith Pérez-Velázquez, Marcel Kettelgerdes, and Gordon Elger. Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability engineering & system safety*, 215:107864, 2021.
(Back reference pages: 4, 10, and 14)

[34] Danielle Smith Bassett and ED Bullmore. Small-world brain networks. *The neuroscientist*, 12(6):512–523, 2006.
(Back reference page: 5)

[35] Benjamin Roche, Jean-François Guégan, and François Bousquet. Multi-agent systems in epidemiology: a first step for computational biology in the study of vector-borne disease transmission. *BMC bioinformatics*, 9(1):1–9, 2008.
(Back reference page: 5)

[36] Célian Colon, David Claessen, and Michael Ghil. Bifurcation analysis of an agent-based model for predator–prey interactions. *Ecological Modelling*, 317:93–106, 2015.
(Back reference page: 5)

[37] DP Noren, AH Johnson, D Rehder, and A Larson. Close approaches by vessels elicit surface active behaviors by southern resident killer whales. *Endangered Species Research*, 8(3):179–192, 2009.
(Back reference page: 5)

[38] Michael J Noad, Rebecca A Dunlop, DAVID Paton, and Douglas H Cato. An update of the east australian humpback whale population (e1) rate of increase. *Santiago Paper submitted to the International Whaling Commission Scientific Committee*, pages 1–13, 2008.
(Back reference pages: 5, 10, and 43)

[39] Solène Derville, Leigh G Torres, Rémi Dodémont, Véronique Perard, and Claire Garrigue. From land and sea, long-term data reveal persistent humpback whale (megaptera novaeangliae) breeding habitat in new caledonia. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 29(10):1697–1711, 2019.
(Back reference pages: 5, 10, and 43)

[40] Ana Širović, John A Hildebrand, and Sean M Wiggins. Blue and fin whale call source levels and propagation range in the southern ocean. *The Journal of the Acoustical Society of America*, 122(2):1208–1215, 2007.
(Back reference pages: 5, 10)

[41] Jennifer B Tennessen, Marla M Holt, M Bradley Hanson, Candice K Emmons, Deborah A Giles, and Jeffrey T Hogan. Kinematic signatures of prey capture from archival tags reveal sex differences in killer whale foraging activity. *Journal of Experimental Biology*, 222(3):jeb191874, 2019.
(Back reference pages: 5, 10)

[42] Alison K Stimpert, Stacy L DeRuiter, Erin A Falcone, John Joseph, Annie B Douglas, David J Moretti, Ari S Friedlaender, John Calambokidis, Glenn Gailey, Peter L Tyack, et al. Sound production and associated behavior of tagged fin whales (balaenoptera physalus) in the southern california bight. *Animal Biotelemetry*, 3(1):1–12, 2015.
(Back reference pages: 5, 11, and 43)

[43] Jeremy A Goldbogen, John Calambokidis, Robert E Shadwick, Erin M Oleson, Mark A McDonald, and John A Hildebrand. Kinematics of foraging dives and lunge-feeding in fin whales. *Journal of Experimental Biology*, 209(7):1231–1244, 2006.
(Back reference pages: 5, 11, 43, and 47)

[44] Ann N Allen, Jeremy A Goldbogen, Ari S Friedlaender, and John Calambokidis. Development of an automated method of detecting stereotyped feeding events in multisensor data from tagged rorqual whales. *Ecology and Evolution*, 6(20):7522–7535, 2016.
(Back reference pages: 5, 11, and 43)

[45] Alison K Stimpert, Marc O Lammers, Adam A Pack, and Whitlow WL Au. Variations in received levels on a sound and movement tag on a singing humpback whale: Implications for caller identification. *The Journal of the Acoustical Society of America*, 147(5):3684–3690, 2020.
(Back reference page: 5)

[46] Malene Simon, Mark Johnson, and Peter T Madsen. Keeping momentum with a mouthful of water: behavior and kinematics of humpback whale lunge feeding. *Journal of Experimental Biology*, 215(21):3786–3798, 2012.
(Back reference page: 5)

[47] M Munir Ahmad and Nasreddin Dhafr. Establishing and improving manufacturing performance measures. *Robotics and Computer-Integrated Manufacturing*, 18(3-4):171–176, 2002.
(Back reference page: 9)

[48] Alyson Fleming and Jennifer Jackson. Global review of humpback whales (megaptera novaeangliae). 2011.
(Back reference page: 10)

[49] LY Rizzo and D Schulte. A review of humpback whales' migration patterns worldwide and their consequences to gene flow. *Journal of the Marine Biological Association of the United Kingdom*, 89(5):995–1002, 2009.
(Back reference page: 10)

[50] Elena Valsecchi, Peter Hale, Peter Corkeron, and William Amos. Social structure in migrating humpback whales (megaptera novaeangliae). *Molecular Ecology*, 11(3):507–518, 2002.
(Back reference page: 10)

[51] Ari Friedlaender, Alessandro Bocconcelli, David Wiley, Danielle Cholewiak, Colin Ware, Mason Weinrich, and Michael Thompson. Underwater components of humpback whale bubble-net feeding behaviour. *Behaviour*, 148(5-6):575–602, 2011.
(Back reference pages: 10, 44)

[52] Peter Tyack and Hal Whitehead. Male competition in large groups of wintering humpback whales. *Behaviour*, 83(1-2):132–154, 1983.
(Back reference page: 10)

[53] L Bejder, S Videsen, L Hermannsen, M Simon, D Hanf, and PT Madsen. Low energy expenditure and resting behaviour of humpback whale mother-calf pairs highlights conservation importance of sheltered breeding areas. *Scientific Reports*, 9(1):1–11, 2019.
(Back reference page: 10)

[54] Amanda Johnson, Glenn Salvador, John Kenney, Jooke Robbins, Scott Kraus, Scott Landry, and Phil Clapham. Fishing gear involved in entanglements of right and humpback whales. *Marine Mammal Science*, 21(4):635–645, 2005.
(Back reference page: 10)

[55] Andrew Kusiak. Smart manufacturing must embrace big data. *Nature*, 544(7648):23–25, 2017.
(Back reference page: 12)

[56] Ilya Kovalenko, James Moyne, Mingjie Bi, Efe C Balta, Wenyuan Ma, Yassine Qamsane, Xiao Zhu, Z Morley Mao, Dawn M Tilbury, and Kira Barton. Towards an automated learning control architecture for cyber-physical manufacturing systems. *IEEE Access*, 2022.
(Back reference page: 12)

[57] Deogratias Kibira, Michael P Brundage, Shaw Feng, and KC Morris. Procedure for selecting key performance indicators for sustainable manufacturing. *Journal of Manufacturing Science and Engineering*, 140(1), 2018.
(Back reference page: 13)

[58] Xiaoning Jin, David Siegel, Brian A Weiss, Ellen Gamel, Wei Wang, Jay Lee, and Jun Ni. The present status and future growth of maintenance in US manufacturing: results from a pilot survey. *Manufacturing review*, 3, 2016.
(Back reference page: 13)

[59] Bulent Dal, Phil Tugwell, and Richard Greatbanks. Overall equipment effectiveness as a measure of operational improvement–a practical analysis. *International Journal of Operations & Production Management*, 2000.
(Back reference page: 13)

[60] Peter Muchiri and Liliane Pintelon. Performance measurement using overall equipment effectiveness (OEE): literature review and practical application discussion. *International journal of production research*, 46(13):3517–3535, 2008.
(Back reference page: 13)

[61] Yassine Qamsane, Efe C Balta, James Moyne, Dawn Tilbury, and Kira Barton. Dynamic rerouting of cyber-physical production systems in response to disruptions based on sdc framework. In *2019 American Control Conference (ACC)*, pages 3650–3657. IEEE, 2019.
(Back reference page: 15)

[62] Carlos M Jarque and Anil K Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259, 1980.
(Back reference page: 20)

[63] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. *Applied logistic regression*, volume 398. John Wiley & Sons, 2013.
(Back reference page: 24)

[64] Erotokritos Skordilis and Ramin Moghaddass. A condition monitoring approach for real-time monitoring of degrading systems using kalman filter and logistic regression. *International Journal of Production Research*, 55(19):5579–5596, 2017.
(Back reference page: 24)

[65] Rocco Langone, Alfredo Cuzzocrea, and Nikolaos Skantzos. Interpretable anomaly prediction: Predicting anomalous behavior in industry 4.0 settings via regularized logistic regression tools. *Data & Knowledge Engineering*, 130:101850, 2020.
(Back reference page: 24)

[66] Alakh Sethi. *One-Hot Encoding vs. Label Encoding Using Scikit-Learn.*, 2020. [Online] https://www.analyticsvidhya.com/blog/2020/03/one-hot-encoding-vs-label-encoding-using-scikit-learn/, Accessed: Jul. 30, 2022.
(Back reference page: 26)

[67] Bradley Efron and Robert Tibshirani. The bootstrap method for assessing statistical accuracy. *Behaviormetrika*, 12(17):1–35, 1985.
(Back reference page: 29)

[68] What species will i see on a whale watch?
(Back reference page: 43)

[69] Mark F Baumgartner, Frederick W Wenzel, Nadine SJ Lysiak, and Melissa R Patrician. North atlantic right whale foraging ecology and its role in human-caused mortality. *Marine Ecology Progress Series*, 581:165–181, 2017.
(Back reference page: 44)

[70] James HW Hain, Sara L Ellis, Robert D Kenney, Phillip J Clapham, Belinda K Gray, Mason T Weinrich, and Ivar G Babb. Apparent bottom feeding by humpback whales on stellwagen bank. *Marine Mammal Science*, 11(4):464–479, 1995.
(Back reference page: 44)

[71] Colin Ware, Ari S Friedlaender, and Douglas P Nowacek. Shallow and deep lunge feeding of humpback whales in fjords of the west antarctic peninsula. *Marine Mammal Science*, 27(3):587–605, 2011.
(Back reference page: 44)

[72] Aleria S Jensen, Gregory Keith Silber, and John Calambokidis. Large whale ship strike database. 2003.
(Back reference page: 44)

[73] David W Laist, Amy R Knowlton, James G Mead, Anne S Collet, and Michela Podesta. Collisions between ships and whales. *Marine Mammal Science*, 17(1):35–75, 2001.
(Back reference page: 44)

[74] Gregory K Silber, Angelia SM Vanderlaan, Ana Tejedor Arceredillo, Lindy Johnson, Christopher T Taggart, Moira W Brown, Shannon Bettridge, and Ricardo Sagarminaga. The role of the international maritime organization in reducing vessel threat to whales: Process, options, action and effectiveness. *Marine Policy*, 36(6):1221–1233, 2012.
(Back reference page: 44)

[75] Marine mammal stock assessment reports by species/stock.
(Back reference page: 44)

[76] Mark P Johnson and Peter L Tyack. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE journal of oceanic engineering*, 28(1):3–12, 2003.
(Back reference page: 46)

[77] Briana H Witteveen, Robert J Foy, Kate M Wynne, and Yann Tremblay. Investigation of foraging habits and prey selection by humpback whales (megaptera novaeangliae) using acoustic tags and concurrent fish surveys. *Marine Mammal Science*, 24(3):516–534, 2008.
(Back reference page: 47)

[78] Michael J Moore. How we can all stop killing whales: a proposal to avoid whale entanglement in fishing gear. *ICES Journal of Marine Science*, 76(4):781–786, 2019.
(Back reference page: 56)

[79] G Antoniak, E Xargay, K Barton, B Popa, and KA Shorter. Estimating whole-body kinematics of swimming bottlenose dolphins. In *INTEGRATIVE AND COMPARATIVE BIOLOGY*, volume 61, pages E21–E22. OXFORD UNIV PRESS INC JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA, 2021.
(Back reference page: 60)

[80] Joaquin Gabaldon, Ding Zhang, Lisa Lauderdale, Lance Miller, Matthew Johnson-Roberson, Kira Barton, and K Alex Shorter. Computer-vision object tracking for monitoring bottlenose dolphin habitat use and kinematics. *PloS one*, 17(2):e0254323, 2022.
(Back reference page: 60)

[81] Lisa K Lauderdale, K Alex Shorter, Ding Zhang, Joaquin Gabaldon, Jill D Mellen, Michael T Walsh, Douglas A Granger, and Lance J Miller. Habitat characteristics and animal management factors associated with habitat use by bottlenose dolphins in zoological environments. *PloS one*, 16(8):e0252010, 2021.
(Back reference page: 60)