Chapter 1
Introduction to Mechanistic Data Science

Abstract Mechanistic data science provides a framework to combine data science tools with underlying scientific tools for addressing a wide range of problems. Data science and machine learning are pushing the limits of the hardware and computer algorithms more and more. Additionally, mathematical science and engineering is constantly taxed with challenging the current status quo. In this perspective, it is paramount to explore the possibility to perform the same task with mechanistic data science. One promising approach to attain this ambitious goal is to make effective use of dimensional reduction through mechanistic data science where the knowledge of past scientific discovery interacts with the streams of data in comporting harmony. A successful implementation of the balanced interaction between the data and existing scientific knowledge will be beneficial to the advancement of science such as accelerating the scientific discovery. As an introductory chapter, the evolution of science, technology, engineering, and mathematics throughout history is explored, from Aristotle and the ancient Greeks to Galileo and Isaac Newton, to the present-day data science revolution. The everyday applications of data science or machine learning are ever-present, with applications ranging from suggested movie preferences to fraud detection. However, an efficient utilization of the data science to scientific discovery with proper deployment of the past acquired scientific and mathematical knowledge is relatively less explored. This chapter presents some real-world examples to demonstrate the methodical approach to solve pragmatic scientific and engineering problems in an accelerated fashion by applying the knowledge of data science in combination with existing scientific laws.

Keywords Mechanistic data science · Fundamental scientific laws · Mathematical science · Scientific method · Falling objects · Gravity · Laws of motion · Law of inertia · Law of force balance · Law of reaction force · Edisonian approach · Empirical approach · Artificial intelligence · Neural networks · Game theory · Nash

Supplementary Information: The online version of this chapter (https://doi.org/10.1007/978-3-030-87832-0_1) contains supplementary material, which is available to authorized users.

W. K. Liu et al., *Mechanistic Data Science for STEM Education and Applications*,
https://doi.org/10.1007/978-3-030-87832-0_1

equilibrium · STEM · Fatigue · Fracture · Material design · Alloys · Gold alloys · AlphaGo · 3D printing · Regression · Multivariate linear regression · Classification · Dimension reduction · Reduced order surrogate model · System and design

Most challenging problems require a combination of data and scientific or underlying principles to find a solution. The power of mechanistic data science is that the techniques which will be shown in this book apply to problems in the physical sciences and mathematics, as well as manufacturing, medical, social science, and business. As shown in Fig. 1.1, mechanistic data science combines established equations from mathematical science with data and measurements through techniques such as neural networks to address problems which were previously intractable. Although the mechanistic data science methodology will be mostly described in terms of engineering and science examples, the same methodology is applicable to any other walks of life in which data is available and decisions are required.

This book is presented in a manner which will be applicable to high school students and teachers, college students and professors, and working professionals. Sections that are more advanced will be noted to inform the reader that additional background knowledge may be required to fully understand that particular section.

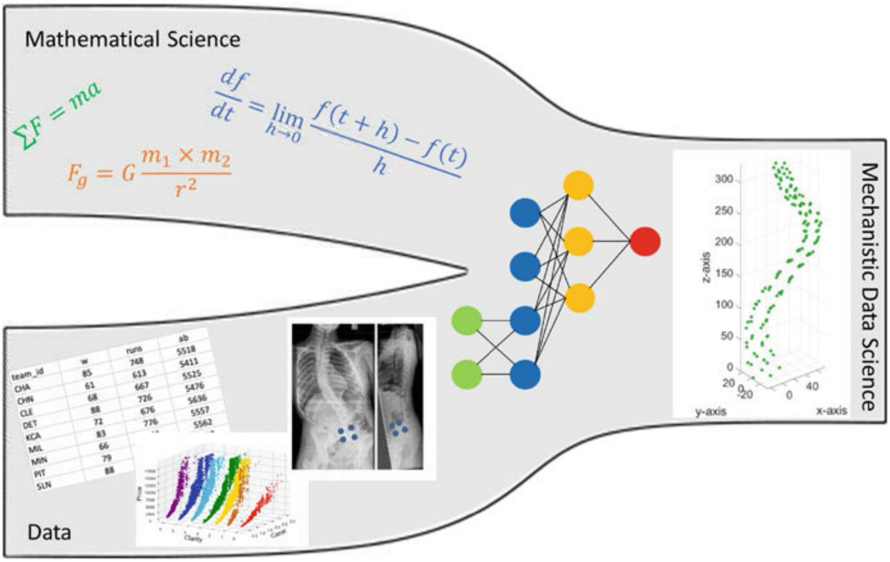


Fig. 1.1 Schematic of Mechanistic Data Science

1.1 A Brief History of Science: From Reason to Empiricism to Mechanistic Principles and Data Science

The history of science has been one of observations, leading to theories, leading to new technologies. In turn, the new technologies have enabled people to make new observations, which led to new theories and new technologies. This cycle has been repeated for thousands of years, sometimes at a slow pace, and other times very rapidly.

Ancient philosophers such as Aristotle (384–322 BC) believed that scientific laws could be discerned through reason and logic. Aristotle reasoned that the natural state of an object was to be at rest and that heavier objects fell faster than lighter objects because there was more downward force on them. Aristotle believed in a geocentric universe (centering around earth) and that the heavens were made of the quintessence, which was perfect and unalterable. In other words, there could be no supernovae or comets.

These scientific ideas remained the benchmark for centuries until scientists in the Renaissance began questioning them. The astronomer Nicolaus Copernicus (1473–1543 AD) proposed a heliocentric model of the universe in which the sun was the center, not the earth. The idea that the earth was not the center of the universe was very difficult for humans to accept. However, the publication of Copernicus' seminal work *On the Revolutions of the Celestial Spheres*, published in 1543, around the time of his death, set off the Copernican Revolution in science which resulted in a major paradigm shift away from the Ptolemaic geocentric model of the universe.

The work of Copernicus was later supported by the observations of Tycho Brahe (1546–1601 AD). Tycho was a talented astronomer who recorded many accurate measurements of the solar system which became the foundation for future astronomical theories. Tycho was followed by Johannes Kepler (1571–1630), who used the data collected by Tycho to formulate scientific laws of planetary motion which could predict the past or future position of the planets. Of particular note, he determined that the planets moved in an elliptical orbit around the sun, not a circular orbit, in contrast to the ancient Greeks, who thought the universe was geocentric and planetary motion was circular.

Galileo Galilei (1564–1642) is known as the father of the scientific method because of his systematic combination of experimental data and mathematics. He was a contemporary of Kepler who was the first scientist to use a telescope to observe celestial bodies and championed the heliocentric model of the universe. In 1638, he published *Discourses and Mathematical Demonstrations Relating to Two New Sciences* (better known by the abbreviated name *The Two New Sciences*) where he laid out the fundamentals for strength of materials and motion.

1.2 Galileo's Study of Falling Objects

One of Galileo's major contributions was his study of motion and his ability to discern primary forces such as gravity from secondary forces such as friction and wind resistance. A notable example was Galileo's study of falling objects. Most people have seen a dense, weighty object like a baseball fall faster than a lightweight object like a feather or a piece of paper. As discussed previously, ancient philosophers such as Aristotle had postulated that heavier objects fall faster than lighter objects in proportion to their mass. This remained the generally accepted theory of gravity until Galileo began studying falling objects in the late 1500s. Around 1590, when Galileo was a professor of mathematics at the University of Pisa in Italy, he reportedly conducted experiments (according to his student Vincenzo Viviani) by dropping objects of different masses from the Leaning Tower of Pisa to demonstrate that they would fall at the same speed [1].

Many years later in 1971, astronaut David Scott performed a similar experiment on the moon, in which he dropped a feather and a hammer at the same time. Because the moon has almost no atmosphere (and thus no air resistance to slow the feather), the feather and the hammer hit the ground at the same time [2].

1.3 Newton's Laws of Motion

Isaac Newton (1642–1726) was a scientist and mathematician best known for his laws of motion and the invention of calculus. Newton's laws of motion are a classic example of a law that was developed through the scientific method. He synthesized many years, decades, and centuries of observations, experimental data, and theories by scientists and mathematicians such as Galileo, Kepler, and Copernicus, into a new understanding of motion. In 1687, Newton published his work *Philosophiae Naturalis Principia Mathematica* (better known by its abbreviated name *Principia*), which has become one of the most classic scientific texts in history. In this book, Newton laid out three fundamental laws of motion:

1. Law of inertia—an object in motion tends to stay in motion and an object at rest tends to stay at rest unless some force is applied to it.
2. Law of force balance—changing the motion requires a force to be applied, which leads to the classic equation: Force = mass × acceleration.
3. Law of reaction forces—for every action/force, there is an equal and opposite reaction.

These three seemingly simple laws of motion account for and describe nearly all the motion seen and experienced in the world around us even to this day. In fact, it was not significantly modified until the early twentieth century when Albert Einstein's theory of relativity was needed to describe the motion of objects traveling at a significant fraction of the speed of light.

Joseph Fourier (1768–1830) was a scientist and mathematician born 42 years after the death of Isaac Newton who made fundamental contributions in the areas of heat conduction and dynamics. Fourier recognized that the any function, including the equation for periodic dynamic motion, can be approximated as a set of sine and cosine functions, called a *Fourier series*. The coefficients for each of the sine and cosine terms in the series determine their relative contributions/amplitudes. As such, the frequencies and amplitudes for a dynamic signal describe the signal and allow for other useful analyses such as filtering and frequency extraction.

Since the time of Newton, tremendous technological strides have been made by coupling fundamentals laws of science with creativity to meet the needs of people. One emblematic example is Thomas Edison (1847–1931) and the light bulb. One of the most basic human needs is to see in the dark, something that was mostly accomplished by fire until the invention of the electric light bulb. Although Edison did not invent the first light bulb, he took it from a crude concept to a mainstream technology. Early light bulbs would only last around 14 h, but through Edison's innovations, the working life improved to 1200 h. This was accomplished through a long and arduous process of trial and error. When asked about the 1000 failed attempts at inventing the lightbulb, Edison famously replied that he "didn't fail 1,000 times. The light bulb was an invention with 1,000 steps".

This Edisonian *brute force* method is a purely empirical approach for applying science for technological development. This method involves pursuing and achieving a goal by building a design and testing it, and making small modifications based on the results of the previous tests. These steps are repeated until the inventor is satisfied with the design. During this process, the inventor is learning what works and what does not work. This information can be used for calibration of parameters in conjunction with the applicable scientific laws. Another offshoot of the Edisonian style brute force method is that lots of data for the various trials is collected which can be very informative for future work. *The combination of calibrated mechanistic principles and data collected can be used to accelerate future development*.

One application of using collected data is artificial intelligence (AI) and neural networks in which data is used to guide decision making. An early success story for AI involved the chess matches between the IBM Deep Blue computer and chess champion Garry Kasparov. In 1985, Carnegie Mellon University began a project to "teach" a computer to play competitive chess. Over the next decade, the computer algorithm was trained using data from 4000 different positions and 700,000 chess games by chess grandmasters. In 1996, Deep Blue actually won a single chess game out of a six-game match against chess champion Garry Kasparov. In a 1997 rematch, Deep Blue won the entire match against Garry Kasparov.

Implicit to the computer programming for playing games like chess is *game theory*, and no one is more synonymous with game theory than John Nash (1928–2015). Nash laid out his theory for achieving an optimal solution to non-cooperative games which came to be known as Nash equilibrium. It states that in a non-cooperative game with known strategies and rational players, the game achieves a state of equilibrium if no player can improve their position by unilaterally changing their strategy. Nash equilibrium is often illustrated by the prisoner's dilemma in which two prisoners apprehended together are interrogated in separate rooms. Both prisoners are given three choices: (1) freedom if they confess before the other prisoner but their partner receives extensive jail time, (2) minimal jail time if neither confesses, or (3) extensive jail time if they do not confess but the other prisoner does confess. It can be easily seen that each prisoner will achieve their own best outcome by confessing first rather than trying to cooperate with each other by not confessing. Since its discovery, Nash equilibrium has become one of the most important concepts in the game theory approach to artificial intelligence and decision making for neural networks.

The dawn of the new millennium also brought about the information age in which information and data are collected and categorized as never before in history. No more going to the library to look up information for a research paper—just Google it. The information age could also be called the data age because of the large amounts of data being collected. The challenge is to turn that data into information and use it synergistically with the already known and established understanding of our world. In this book, this synergy is called *Mechanistic Data Science*.

Human progress has been greatly accelerated by our ability to understand and control the world around us. A large part of this is because of science and engineering. Since the times of Galileo and Newton, the fundamental principles of materials and motion have been further studied and formalized. Special technical fields of study have developed that feed and interact with each another in a symbiotic manner.

1.4 Science, Technology, Engineering and Mathematics (STEM)

Science provides a set of fundamental laws that describe nature and natural phenomena. From physics to chemistry to biology, once the natural phenomena are scientifically understood and described, humans are able to predict a particular outcome without testing for each possible outcome. Science is heavily reliant on mathematics for the "language" in which its laws are written [3] and is also dependent on engineering and technology for applying scientific findings and developing new tools to enable future discoveries.

Mathematics is the unifying language of the physical sciences, and as such, the development of mathematics is integral to scientific progress. The understanding and capability of data scientists in mathematical topics such as algebra, geometry, trigonometry, matrix algebra and calculus foster the progress of science and engineering.

Engineering is the application of scientific principles for design and problem solving. In other words, people can make things. Engineers use data collected regarding the needs and wants of society and then use the principles of science, invoke human creativity, and apply manufacturing craftsmanship to design, develop and produce products that address these societal needs and wants.

Technology is the implementation of products and capabilities developed through science and engineering. This generally takes the form of actual products on the market and in use today. Technology draws heavily on scientific discovery and engineering development to be able to address challenges, grow the economy, and improve efficiency.

The **scientific method** provides an organized methodology for studying nature and developing new scientific theories. A basic form of the scientific method is:

- Observe: the subject of interest is studied and characterized from multiple standpoints. This first stage involves data collection in order to move to the next step.
- Hypothesize: based on the observations and the early data collected, a hypothesis (proposed explanation) is developed.
- Test: experiments are conducted to evaluate and challenge the hypothesis. In this stage, extensive amounts of data are collected and analyzed.
- Theory or law: if a hypothesis is not proven to be false during the testing challenges then it is established as a theory. Theories that are considered to be fundamental and widely accepted are often described as laws. Note that often times a theory or law is established with limitations for when it is valid (e.g., Newton's laws of motion are valid for speeds much less than the speed of light, but Einstein's theory of relativity is required for objects moving at speeds close to the speed of light.)

There are often many special cases for a scientific law. In the above-mentioned example of a falling object, it is necessary to account for air and wind resistance when comparing a falling hammer versus a feather. They will fall at the same rate in a vacuum or on the surface of the moon, but in normal atmospheric conditions, the effect of air resistance makes a noticeable difference in the rate of falling. In this case, the law of gravity still applies, but it must be coupled with data on wind resistance, object density, and aerodynamics in order to properly describe the phenomena.

Once a scientific law has been established, engineers can use this information to make calculations in the continuing quest to design and build new products. For example, understanding gravity and falling objects allows engineers to design and build objects that can fly, whether they are a backyard water bottle rocket as shown in the Fig. 1.2 below, or a high-tech reusable rocket like the SpaceX rocket that can return components to Earth and land upright on a barge in the ocean.

1.5 Data Science Revolution

Recent years have seen a revolution in data science as large amounts of data have been collected on a vast array of topics. For instance, the ubiquitous smart phone is constantly collecting and transmitting information about grocery store purchases,

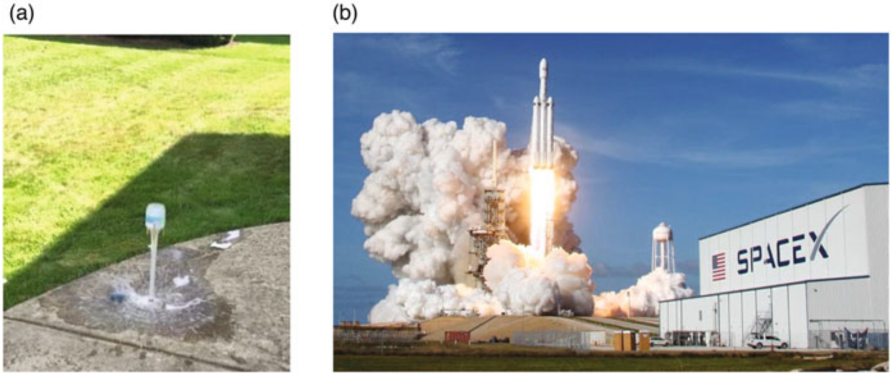


Fig. 1.2 (**a**) Backyard water bottle and baking soda rocket (a video is available in the E-book, Supplementary Video 1.1). (**b**) SpaceX Falcon Heavy rocket launch (Reuters/Thom Baur)

travel routes, and search history. Companies like Facebook, Google, and Instagram collect and utilize data posted on their site for various marketing purposes such as targeted advertisements and purchase recommendations. These sites match demographic information such as age, gender, and race with internet browsing history, purchases made, and photos and comments posted to predict future behavior, such as whether you are likely to buy a car or make some other significant purchase. These predictions can be used to target advertisements at the right audience.

Data science has been heavily used for product development, from the concept stage to engineering and manufacturing to the customer. Data collected at each stage can be used to improve future products or identify the source of problems that arise. For instance, manufacturers regularly collect customer data to understand the "voice of the customer", which is used to plan future product models. Furthermore, as a product is being manufactured, data is being collected at every step of the manufacturing process for process control. The increased use of sensors has given rise to the internet of things (IOT) in which data is automatically collected and transmitted over the internet for analysis without requiring explicit human interaction.

1.6 Data Science for Fatigue Fracture Analysis

Disasters have often been a driving force for exploration of new areas or to develop a much deeper understanding of existing areas. The data, techniques, and methods generated from these explorations in turn becomes a boon for engineers and product designers when designing new products or working to improve existing designs. One such area of scientific and engineering exploration is the study of fractures and failures due to fatigue. Fatigue is the initiation and slow propagation of small cracks into larger cracks under repeated cyclic loading. The formed cracks will continue to

Fig. 1.3 Aloha Airlines flight 243 fatigue fracture example (**a**) scheduled and actual flight paths of flight 243 (**b**) Boeing 737 airplane after fuselage separation (**c**) schematic illustration of fatigue crack growth between rivet holes (https://fearoflanding.com/accidents/accident-reports/aloha-air-243-becomes-relevant-thirty-years-later/)

get longer and longer while the product is being used under normal operating loads, until they become so large that the structure fails catastrophically.

Consequences of Fatigue On April 28, 1988, Aloha Airlines flight 243 took off from Hilo, HI on a routine flight to Honolulu, HI. The Boeing 737 airplane had just reached cruising altitude when a large section of the fuselage separated from the plane (see Fig. 1.3). The pilot was able to successfully land the plane on the island of Maui, although one flight attendant was lost in the incident. Post-incident inspection of the airplane showed that small cracks had initiated from the rivet holes that were used to join the separate pieces of the airplane fuselage. The cracks had propagated slowly from hole to hole until the resulting crack was sufficiently large that the structure could no longer support the service loads in the forward section of the plane. A commercial airplane is pressurized for every flight, which stresses the fuselage, in addition to takeoff and landing loads, and vibration loads. The subject airplane had accumulated 89,680 flight cycles and 35,496 flight hours prior to the incident [4].

It should be noted that this incident resulted in the formation of the Center for Quality Engineering and Failure Prevention, led by Prof. Jan Achenbach at Northwestern University, and with which two of the authors collaborated extensively in the past. Prof. Achenbach received both the National Medal of Technology and the National Medal of Science, partly for his important scientific work on the non-destructive detection of fatigue cracks.

Fatigue Design Methodology One of the most common methods for fatigue design and analysis is a mechanistic data driven methodology known as the stress life method. In this methodology, the material of the design has been tested at many different stress levels to determine how many load cycles to failure. That stress amplitude is then plotted on a graph vs. the number of cycles to failure for that material. This plot is commonly referred to as an S-N curve. When used for design analysis, the cyclic stress amplitude at a location of interest is either measured by laboratory testing, field testing or computed using finite element analysis or hand calculations. Using the computed or measured stress amplitude and the appropriate S-N curve, the fatigue life can be estimated.

There are many factors involved in the initiation and propagation of fatigue cracks, such as material strength, microscopic impurities and voids, and surface roughness. For each material of interest, many controlled laboratory tests are conducted on standardized specimens at a range of stress levels to measure how many cycles the material can endure before fracturing. This data-driven approach has been necessary because of the relatively large number of randomly occurring factors that are involved.

Fatigue cracks generally initiate at or near the surface of the material. It has been found that parts with rougher surfaces will have shorter fatigue lives, with the effect being more pronounced at higher fatigue lives. For many materials, a large amount of testing has been performed to characterize the effect of the surface roughness due to the manufacturing process used to make the part (eg. polished surface, machined surface, or as-cast surface finish).

1.7 Data Science for Materials Design: "What's in the Cake Mix"

The macrostructure (or physical structure you can see and hold in your hand) is composed of trillions of atoms of different elements which are mixed and organized in a certain way at the small-scale sub-structure, and this organization depends on the particular material being used. This small-scale sub-structure is referred to as the microstructure if a powerful microscope is required to see it and is called the mesoscale if you need to use a magnifying glass or just look very closely.

The overall macrostructural performance of a bulk part is controlled in large part by the microstructure of the material used to make the part. For example, if a pastry chef is baking a cake, the flavor, texture, and crumble of the cake are controlled by the ingredients, the mixing, and the baking time and temperature. One can think of the microstructure as "what's in the cake mix" (Fig. 1.4).

Engineered components are often evaluated for strength, stiffness, and fracture resistance (as opposed to taste and texture for a cake). A simple example is to consider an ice cube. If one were to make an ice cube from only water, the ice cube would likely shatter into many pieces when dropped on a rigid surface from a sufficient height. However, if other ingredients were added to the water when the ice was made, the resulting ice cube would likely be more resistant to shattering, depending on what was added. For example, if strips of newspaper were added to the ice, the resulting reinforced ice cube would not shatter when dropped from the same height as the unreinforced ice cube. This is because the mesostructure of the newspaper in the ice increases the toughness of the ice cube and resists cracks from propagating through the cube as it impacts the rigid surface (Fig. 1.5).

If one realizes that we have entered the digital age and newspaper is no longer readily available, then new composite material needs to be developed to replace newspaper filler. Given below is a sampling of alternative reinforcement materials that could be used to make a composite cube structure. To evaluate the impact fracture resistance of the various cubes, several composite ice cubes were made

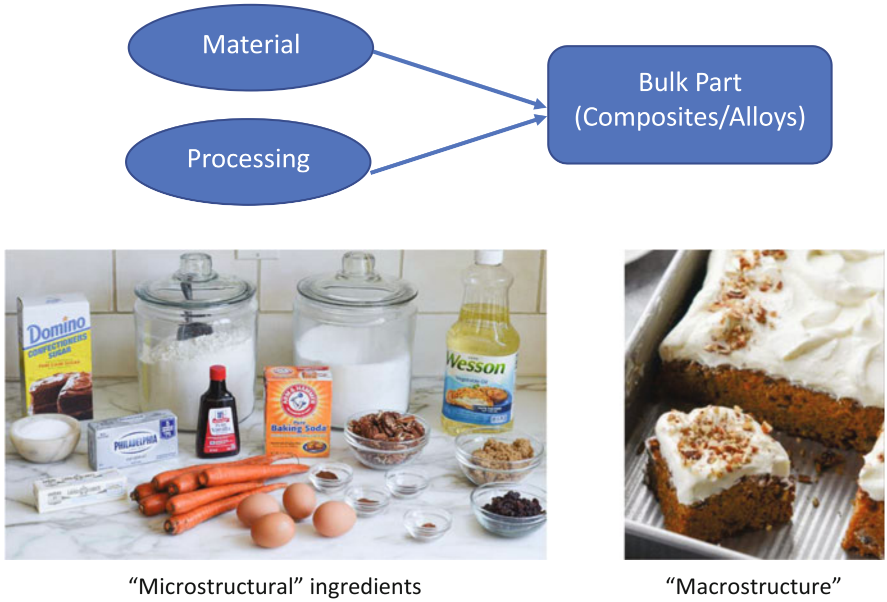


Fig. 1.4 “Cake Mix” material microstructure example (“Classic Carrot Cake with Cream Cheese Frosting.” Once upon a chef with Jenn Segal. https://www.onceuponachef.com/recipes/carrot-cake.html)

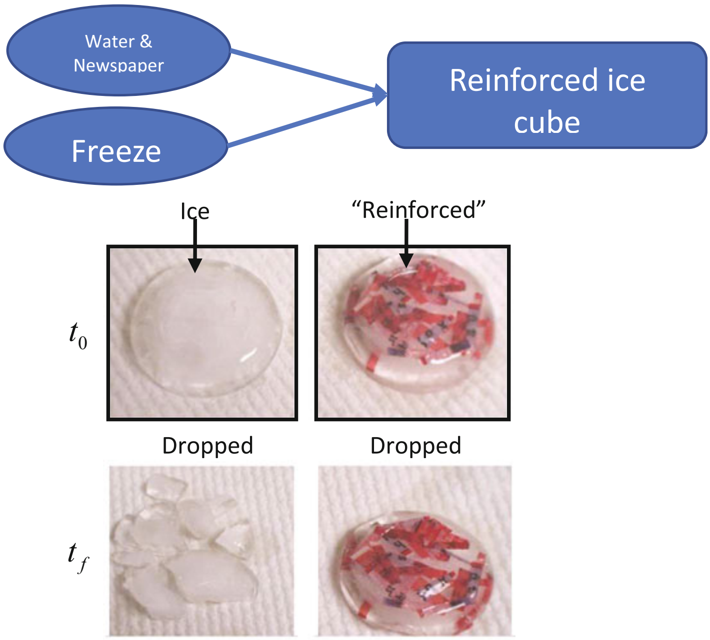


Fig. 1.5 Ice with and without reinforcement dropped to show effects of reinforcement on fracture resistance. Experiment by Northwestern University Prof. Yip Wah Chung, 2003

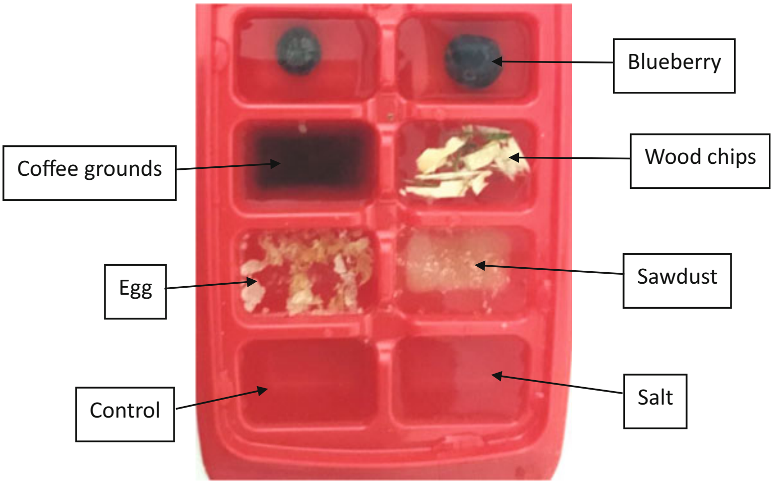


Fig. 1.6 Composite ice cube experiment by Northwestern University Prof. Mark Fleming and Carmen Fleming, 2020

Table 1.1 Drop test results for ice cubes formed with various reinforcement materials

Filler material	Result
Control	Fractured, big chunks
Salt water	Small chunks
Egg	Fractured, big chunks
Sawdust	Fractured, small chunks
Wood chips	Very little fracturing
Coffee grounds	Fractured, big chunks
Blueberry	Fractured, big chunks

(Fig. 1.6) and dropped 14 ft. (impact speed of 30 ft./s) onto a concrete surface. Results showed that the addition of wood chips was most effective in preventing fracture of the cubes on impact. This resistance to fracture is called the fracture toughness and is important when designing objects which need to have good impact resistance (Table 1.1).

1.8 From Everyday Applications to Materials Design

The reinforced ice cube can be considered as a metaphor for an engineered composite material in which the mesoscale structure enhances the overall structural performance. Consider the material substitutions shown in Fig. 1.7. A composite material in its most basic form consists of a matrix material with some reinforcing material embedded inside. If the matrix ice material is replaced with an epoxy and

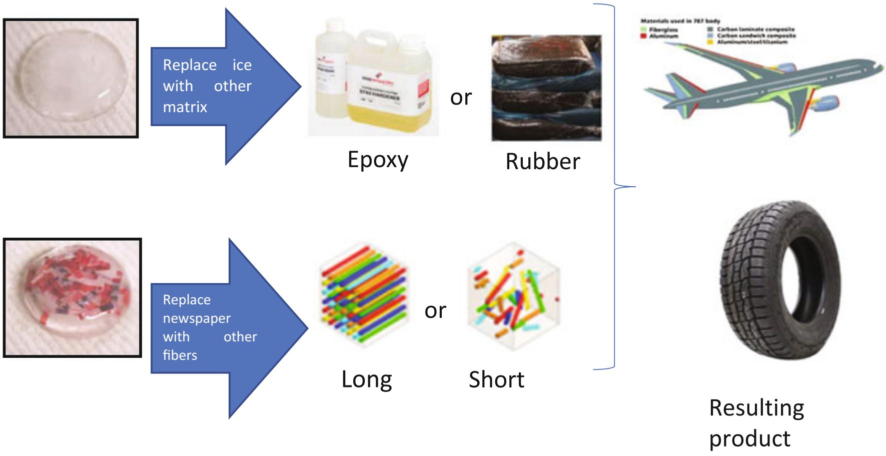


Fig. 1.7 Ice cube to engineered materials analogy for materials design

the newspaper is replaced with carbon fibers, the result is a carbon fiber reinforced composite material. On the other hand, if the matrix ice material is replaced with rubber and the newspaper is replaced with steel or polymer cords, a tire can be made.

1.8.1 Example: Tire Tread Material Design Using the MDS Framework

Tire durability is one of the fundamental questions facing the tire industry. The unpredictable weather conditions and road surface conditions that each tire faces every day have significant impact on its durability. One of the key materials property metrics that can be related with the tire material performance is called tan(δ). For tire materials, a high tan(δ) is desirable for low temperatures (better ice and wet grip) and a low tan(δ) is desirable for high temperatures (better rolling friction). It is noteworthy to mention that approximately 5–15% of the fuel consumed by a typical car is used to overcome the rolling friction of the tire on the road. Therefore, controlling the rolling friction of tires is a feasible way to save **energy** (by reducing fuel consumption) and **reduce the environmental impact** (by reducing carbon emission). Additionally, it ensures the **safe operation** of tires in providing sufficient ice or wet grip.

The key performance metric tan(δ) is a function of the matrix materials, microstructure, and the operating conditions such as temperature and frequency. It is well known that adding fillers improves the tire materials performance. But what fillers and their distribution to achieve optimized properties and performance is still an important research question. The design space combine different rubber matrices and fillers, and the microstructure and operating conditions can be enormous, making some experimental or simulation techniques not feasible. The mechanistic data science approach can provide an effective solution to explore the design space by leveraging the data science tools to reveal the mechanism and construction of accurate and efficient reduced order surrogate models. This approach will enable the industrial practitioner to perform rapid design iteration and expedite the decision-making process (Fig. 1.7).

1.8.2 Gold and Gold Alloys for Wedding Cakes and Wedding Rings

Pure gold is so ductile that it can be rolled into sheets that are so thin that it can be used to decorate cakes and subsequently eaten. While this expensive application of pure gold is an interesting possibility for a wedding cake, gold is more often used for jewelry such as a wedding ring. As such, it needs to hold a specific shape. Materials like gold are strongly influenced by the microstructure of the material. Pure 24 K gold is extremely ductile and malleable, meaning that it can be reshaped by rolling and pounding without cracking. Gold jewelry is generally made from 18 K or 14 K gold, which is strengthened by alloying (mixing with other elements) it with other metals—see periodic table of elements in Fig. 1.8. As an example, 18 K gold contains a mix of 75% pure gold, 10% copper, 8% nickel, 4.5% zinc and 2.5% silver (Fig. 1.9).

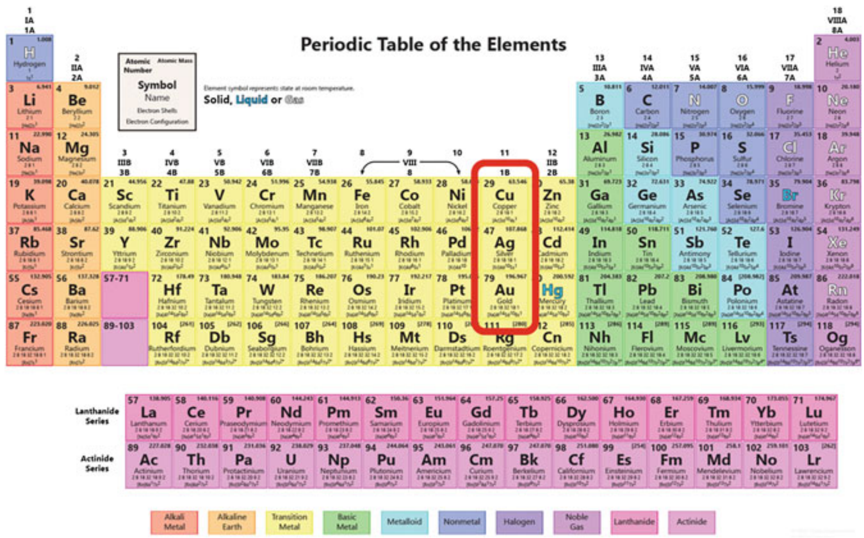


Fig. 1.8 Periodic table of elements (Source: Sciencenotes.org). The red box signifies some of the elements used for making gold alloys

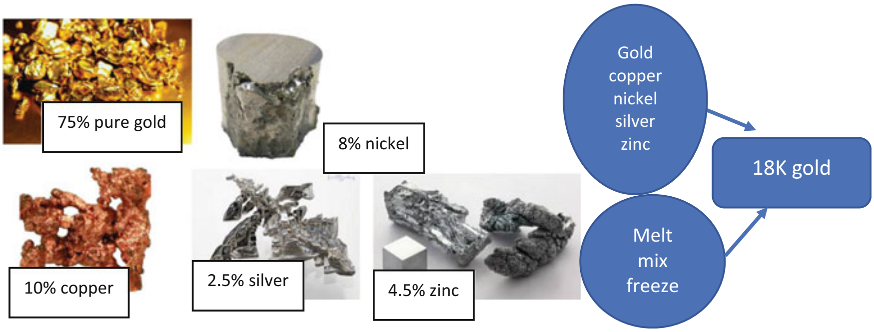


Fig. 1.9 Elements making up 18 K gold, which is a common gold alloy for jewelry

1.9 Twenty-First Century Data Science

1.9.1 AlphaGo

One very interesting data science accomplishment is the use of deep learning neural networks for the complex game of Go. Go is a strategy game created 3000 years ago in China. The game is played with black and white stones that are placed on a grid-filled board in an attempt to surround an opponents stones and to strategically occupy space. The possibilities of the game are astronomical, with 10^{170} possible configurations, which makes it much more complex than chess. Until recently, the best Go computer programs could only play at a relatively novice level.

In 2014, a company called Deep Mind began working on a project that led to a program called AlphaGo and trained it to play using deep learning neural networks. It was initially trained with many amateur games and against itself, all the time improving its ability.

In 2015, AlphaGo played a match against reigning European champion, Mr. Fan Hui, and won a game. Then in 2016, AlphaGo beat 18-time world champion Mr. Lee Sedol. Since that time, additional versions have been released, including AlphaZero, which is capable of learning other games such as chess and shogi.

1.9.2 3D Printing: From Gold Jewelry to Customized Implants

Recently, a new method of manufacturing structural parts called additive manufacturing (AM) or 3D printing has become popular. A report from the National Academy of Engineering identified 3D printing as a revolutionary new manufacturing technology capable of making complex shapes, and possibly 1 day printing new body parts [5]. Additive manufacturing generally works by building a part through

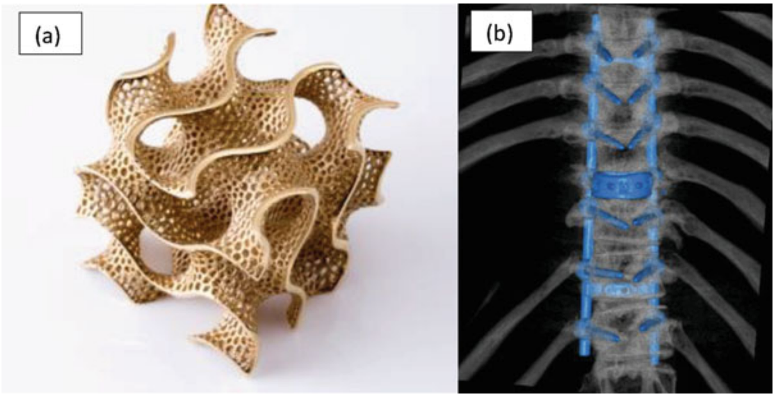


Fig. 1.10 (**a**) A example of 3D printed gold jewelry [6]. (**b**) 3D custom implant to reconstruct a vertebra destroyed by a spinal tumor [7]

depositing thin layers of material one after the other until a complete part is formed. This process is able to form very complicated shapes, and has been widely used in many industrial applications from precious metals such as gold to customized 3D spinal implants (see Fig. 1.10). This type of technology is making a difference for the environment, health, culture and more. This has included disaster relief, affordable housing, more efficient transportation (less pollution), better and more affordable healthcare, 3D bioprinted organs, cultural and archeological preservation, accessible medical and lab devices, and STEM education.

A great deal of data are generated during 3D printing process and part qualification, including process parameters, physical fields (varying in both space and time), material microstructure, and mechanical properties and performance. Data science is extremely useful for 3D printed part qualification and optimization, stablized printing processes and improved material properties.

1.10 Outline of Mechanistic Data Science Methodology

As shown in Fig. 1.11, data science for solving engineering problems can be broken down into six modules, ranging from acquiring and gathering data to processing the data to performing analysis.

Mechanistic data science is the structured use of data combined with the core understanding of physical phenomena to analyze and solve problems, with the end goal of decision-making. The problems to be solved range from purely data-driven problems to problems involving a mixture of data and scientific knowledge.

Type 1 *purely data-driven*: problems with abundant data but undeveloped or unavailable fundamental principles. This type of problem can be illustrated through using data for the features of diamonds to determine the price. There is not an explicit

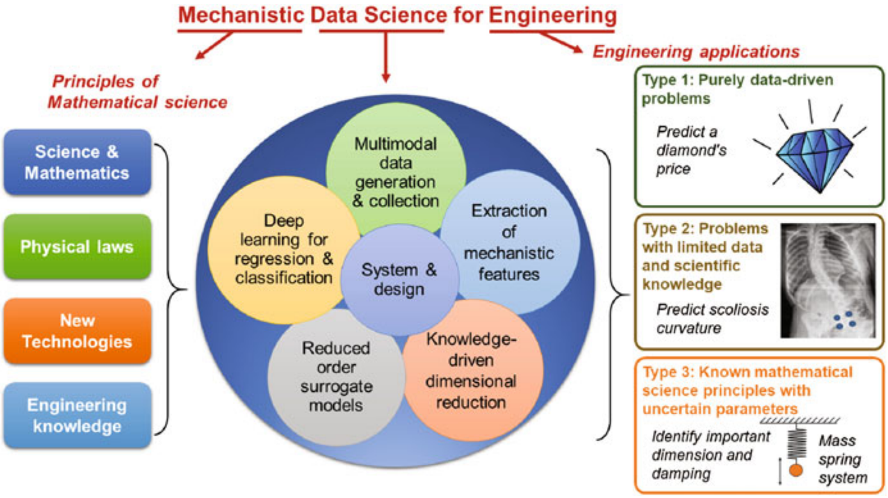


Fig. 1.11 Schematic of Mechanistic Data Science for Engineering. Scientific knowledge is combined with data science to effect engineering design with the goal of obtaining knowledge for improved decision making

"theory" associated with diamond prices. Instead, the price is determined by the complex interplay of many features such as size and sparkle.

Type 2 *limited data and scientific knowledge*: problems in which neither the data nor the scientific principles provide a complete solution. This type of problem can be illustrated by the analysis of scoliosis patients. X-rays provides some data for the progression of spine growth but combining that data with finite element surrogate models in a neural network provides a good estimate of scoliosis progression.

Type 3 *known mathematical science principles with uncertain parameters*: problems which can be computationally burdensome to solve. This type of problem can be illustrated through a spring-mass example. Physics models of spring-mass systems typically assume a point mass, a massless spring, and no damping. Data collected on an actual spring-mass system can illustrate how to use data science to identify key physical factors such as damping coefficient directly from high-dimensional noisy data collected from experimental observations.

In this book, mechanistic data science will be broken into seven chapters:

Chapter 1: Introduction of Mechanistic Data Science

Illustrative problems will be used to demonstrate the power of MDS: Determining price of a diamond based on features (pure data science—Type 1), predicting playoff contention for baseball teams (pure science—Type 1), predicting patient-specific scoliosis curvature (mixed data science and surrogate—Type 2), and identifying important dimension and damping in a mass-spring system (Type 3 problem). The details of the solutions will be provided in the following chapters. (This chapter is for both general readers and advanced readers).

Chapter 2: Multimodal Data Generation and Collection

Large quantities of data are collected related to the topic under study. Multimodal data is data from various types of sources, such as different type of measuring instruments and techniques, models, and experimental setup. Multimodal data generation and collection will be described in Chap. 2. Data and how data evolves to empiricism and to mechanism will be introduced using the story of Kepler's laws and Newton's laws or motion in the seventeenth century. Modern deep learning datasets and use those datasets to solve an engineering problem will be described using examples such as material quantification using macro/micro-indentation to measure the material hardness.

Chapter 3: Least Square Optimization

This chapter includes prerequisites for the book. We will introduce the concept of least square optimization and use a Baseball example to demonstrate linear regression. Nonlinear regression methods will be used to analyze interesting phenomena such as stock market performance and bacteria growth. A few more advanced methods such as moving least-squares and reproducing kernel will be introduced for advanced readers.

Chapter 4: Extraction of Mechanistic Features

Mechanistic features are the key pieces of data that will be used for further data science analysis. They often have to be computed from the raw data collected. It should be noted that data scientists generally describe the feature extraction process as the step where they spend a majority of their time. The concept of the traditional Fourier transformation will be shown and related to an important aspect in modern data science called convolution. Extraction of meaningful features will be shown for real life and engineering problems, such as human speech analysis and additive manufacturing.

Chapter 5: Knowledge-Driven Dimensional Reduction and Reduced Order Surrogate Models

Dimension reduction is the process of reducing the number of random variables under consideration by obtaining a set of principal variables, generally based on the mechanistic features extracted. Two types of dimensional reduction will be introduced. The first type reduces the number of data points based on clustering methods such as k-means clustering and self-organizing map (SOM). We will apply those clustering methods to real life jogging, diamond price, and additive manufacturing design. Another type of dimension reduction is to reduce the number of features by eliminating redundancy between them. Reduced order surrogate models can be built by using those dimension reduction methods. Singular value decomposition (SVD), principal component analysis (PCA), and proper generalized decomposition (PGD) will be described. Identification of intrinsic properties from a spring-mass system will be shown.

Chapter 6: Deep Learning for Regression and Classification

*Regression analysis is a set of statistical processes for estimating the relationships between a dependent variable (often called the 'outcome variable') and one or more independent variables (often called 'predicto*rs', 'covariates', or 'features'). In machine learning and statistics, classification is the problem of identifying to which of a set of categories (sub-populations) a new observation belongs, on the basis of a training set of data containing observations (or instances) whose category membership is known. Deep learning method such as feedforward neural networks (FNN) will be described and used to capture nonlinear relations in examples such as diamond price prediction. The learning process will be analyzed by Fourier transform to demonstrate some features of neural networks learning. Combining convolution layers and feedforward neural networks, convolutional neural networks (CNN) will be introduced. The power of CNN will be illustrative by examples of human behavior detection and damage identification of rolling bearings.

Chapter 7: System and Design

System and design tie the other modules together. The data science is coupled with the mechanistic principles in order to complete an analysis and make decisions. Based on the author's years of experience, state-of-the-art applications ranging from daily life (like baseball) to engineering will demonstrate the concepts of system and design:

1. Piano example with spring mass system (Type 3 general)
2. Feature-based diamond pricing (Type 1 general)
3. Additive manufacturing (Type 1 advanced)
4. Spine growth prediction (Type 2 advanced)
5. Composite design (Type 3 advanced)
6. Indentation analysis for materials property prediction (Type 2 advanced)
7. Early warning of rainfall induced landslides (Type 3 advanced)

1.11 Examples Describing the Three Types of MDS Problems

Examples are provided in this section to illustrate the application of mechanistic data science for the three types of problems described in Sect. 1.10. Examples ranging from strictly data-centric problems to medical diagnosis to physics-based problems are provided to illustrate the broad applications of MDS.

1.11.1 Determining Price of a Diamond Based on Features (Pure Data Science: Type 1)

Jim and Maddie are two young people who want to get married. Jim wants to give Maddie a traditional diamond ring that really sparkles. However, they recently graduated from college and have student loans to pay off, as well as many other expenses. Jim realizes that he needs to study what makes diamonds sparkle in order to get the best diamond ring that he can afford. Jim took a mechanistic data science course in college and decided to use that analytical capability when studying the features and prices of diamonds.

A cursory study of diamonds shows that they have some very impressive properties and can be used for industrial applications as well as jewelry. Diamonds are the hardest substance on earth, which make them popular for surface coatings in which wear resistance is important, such as cutting and drilling tools.

Diamonds for jewelry are known for their sparkle and impressiveness, which are features that are not easily quantified. However, they are functions of other features that can be quantified. The classic, best-known, quantifiable features of diamonds are the 4-C's (Cut, Clarity, Color, Carat), but other features such as depth and dimension are also reported. A combination of all these features is used when determining the price of a diamond [8].

Multimodal Data Collection Jim found a large repository of data on diamond features and prices can be found at www.kaggle.com, which is a subsidiary of Google. For his analysis, a database of 53,940 diamonds is catalogued with information for ten different features along with the price for each diamond. A sample of some of the 3 pieces of the larger dataset is shown in the table below:

Carat	Cut	Color	Clarity	Depth	Table	Price	x	y	z
0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
0.21	Premium	E	SI1	59.8	61.0	326	3.89	3.84	2.31
0.23	Good	E	VS1	56.9	65.0	327	4.05	4.07	2.31

Extraction of Mechanistic Features The raw diamond data available on Kaggle is a good first start, but Jim needed to do some feature engineering, or initial data processing, to get it into an appropriate form for mathematical analysis. This includes removing missing values and converting alphabetic and alphanumeric scores to numerical scores (e.g., the cut of a diamond is rated as Premium, Very Good, Good, Fair, and Poor. These ratings are converted to 1, 2, 3, 4, and 5.). In addition, it is often useful to normalize the data.

Dimension Reduction Jim found that for the analysis he was performing, certain features were often more useful than others. Since he was interested in how a diamond sparkles, the clarity and color were more interesting than the geometric features such as size distribution.

For his analysis, Jim created a correlation matrix to help separate the relevant from the irrelevant features. From this, Jim selected four features for further analysis: carat, clarity, color, y.

Regression and Classification Regression is the process of developing a mathematical relationship between variables in a dataset. Jim first performed a basic form called **linear regression**, which tries to determine where to draw a line through the

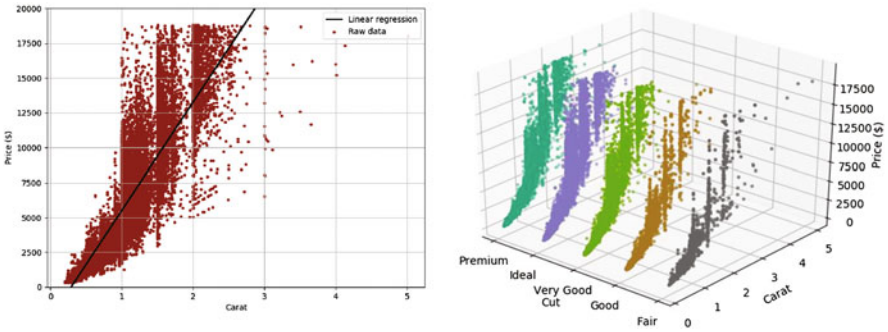


Fig. 1.12 Diamond price vs. carat for all diamonds combined (left) and separated by cut (right)

middle of the data. In Fig. 1.12a, the red dots represent the raw data for carat versus price and the black line represents a linear regression between the price and the carat of the diamonds. The **correlation** of the data describes how close the actual data is to the regression line. It can be seen from looking at the graph that when all the diamond features are lumped together, there is a lot of scatter of the red data points around the black line, resulting in a low correlation. To achieve a better fit, Jim needed to consider additional features. Figure 1.12b shows the price vs. carat data subdivided by clarity, ranging from low clarity I1 diamonds with inclusions to high clarity IF diamonds that are inclusion free. From this graph, it can be seen that the price per carat increases more quickly for higher clarity diamonds. However, there is still increased scatter at higher prices for all the clarity levels.

Multivariate Linear Regression Jim then decided to perform multivariate linear regression in order to consider the effects of multiple features simultaneously. He found that the correlation of the regression improved as more features are considered. In the graphs in Fig. 1.13 below, the prediction using a linear regression is plotted versus the actual data. It can be seen that as the number of features considered increases the amount of scatter decreases, with the least scatter noted when all 9 features are used for the multivariate linear regression.

System and Design Once the multivariate linear regression is complete, Jim was able to estimate the price of a diamond using multiple features, and the nonlinear nature of diamond pricing is seen. For instance, using the regression data, the following diamond prices are found for different sized diamonds:

1 carat diamond = \$4600
2 carat diamond = \$17,500

In short, Jim found that when it comes to diamonds $1 + 1 \neq 2$.
Jim chose . . .?

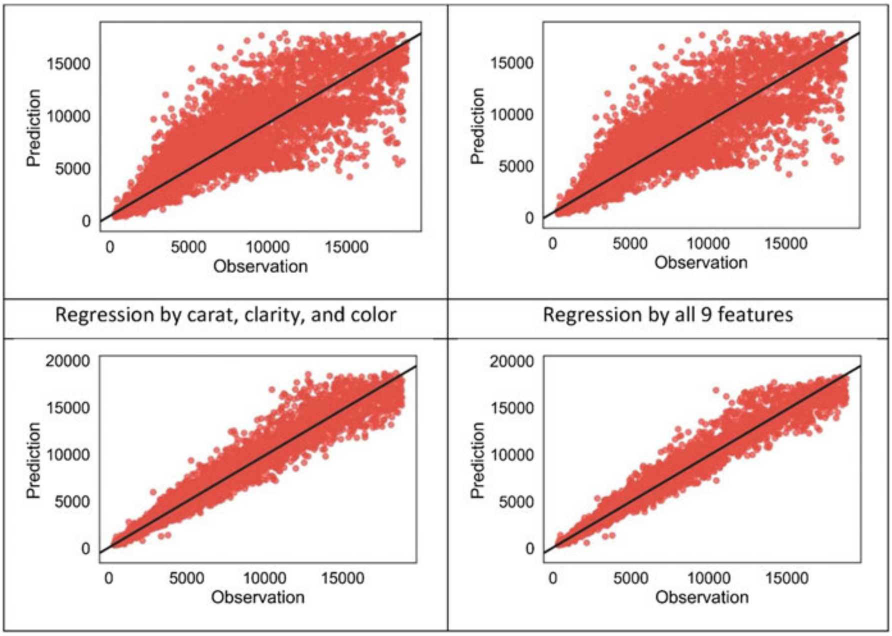


Fig. 1.13 Diamond price prediction vs. observation using regression based on various numbers of features

1.11.2 Sports Analytics

Data science and analytics are very important in sports. Whether it is deciding which star player to choose in a professional sports draft, evaluating fantasy football statistics, or evaluating shot selection in a basketball game, the insight provided from data science has been very profound in sports.

1.11.2.1 Example: "Moneyball": Data Science for Optimizing a Baseball Team Roster

Baseball is a game in which tradition is strong and data and statistics carry great weight. This allows baseball fans to compare the careers of Ty Cobb in 1911 to Pete Rose in 1968 (or anyone else for that matter). Historically, the worth of a player was largely dictated by their batting average (how many hits compared to how many time batting) and runs batted in (how many runners already on base were able to score when the batter hit the ball). However, through the use of data science, a new trend emerged.

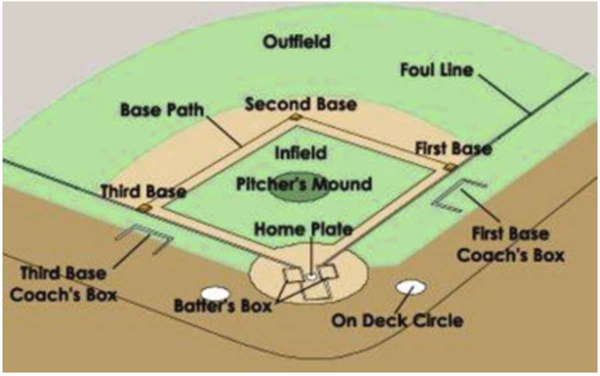
Team	League	Year	RS	RA	W	OBP	SLG	BA	Playoffs	RankSeaso	RankPlayo	G	OOBP	OSLG
ARI	NL	2012	734	688	81	0.328	0.418	0.259	0			162	0.317	0.415
ATL	NL	2012	700	600	94	0.32	0.389	0.247	1	4	5	162	0.306	0.378
BAL	AL	2012	712	705	93	0.311	0.417	0.247	1	5	4	162	0.315	0.403
BOS	AL	2012	734	806	69	0.315	0.415	0.26	0			162	0.331	0.428
CHC	NL	2012	613	759	61	0.302	0.378	0.24	0			162	0.335	0.424
CHW	AL	2012	748	676	85	0.318	0.422	0.255	0			162	0.319	0.405

The game of baseball is played with 9 players from one team in the field playing defense (see figure to right). A pitcher throws a baseball toward home plate where a batter standing next to home plate tries to hit the ball out into the field and then run to first base. If the batter hits the ball and makes it to first base before the ball is caught or picked up and thrown to first base, then the batter is awarded a hit and allowed to stay on the base, becoming a base runner. If the ball is caught in the air, picked up and thrown to first base before the batter arrives, or the batter is tagged when running to first base then the batter is out. The runner can advance to second, third, and home bases as other batters get hits. When the runner reaches home base, the team is awarded a run. The batting team continues to bat until they make three outs, at which point they go out to the field and the team in the field goes to bat.

As mentioned previously, batting average (BA) and runs batted in (RBI) have traditionally been a very important statistic for baseball teams to evaluate the worth of a player. Players with high BA's and high RBI's were paid very large salaries by the richest teams (usually large market teams) and the small market teams had trouble competing.

In 2002, Billy Beane, the general manager of the Oakland Athletics, utilized data science to field a competitive team. Although Major League Baseball (MLB) generates around $10 billion in annual revenue, the smaller market MLB teams have a much lower budget with which to recruit and sign players. In 2002, Oakland A's general manager, Billy Beane, found himself in a tough situation because of this. The Oakland A's were a small market team without a large budget for player salaries. Beane, and his data science capable assistant, Paul DePodesta, analyzed baseball data from previous seasons and determined that they needed to win 95 games to make the playoffs. To achieve this goal, they estimated they needed to score 133 more runs than their opponents. The question they had to answer was "what data should they focus on".

To field a competitive team, Beane and DePodesta looked at a combination of a player's on-base percentage (OBP), which is the percentage a batter reaches base, and the slugging percentage (SLG), which is a measure of how many bases a batter is able to reach for a hit. In formulaic terms: $SLG = (1B + 2B * 2 + 3B * 3 + HR * 4)/AB$, where $1B$, $2B$, *and* $3B$ are first, second, and third base, respectively, HR is a "home run", and AB is an "at bat". Through these two measures, it is possible to assess how often a player is getting on base in any possible way (and thus in a position to score) and how far they go each time they hit the ball.



It is possible to show through linear regression that SLG and OBP provide a good correlation with runs scored (RS). Using a moneyball baseball dataset available from Kaggle (https://www.kaggle.com/wduckett/moneyball-mlb-stats-19622012/data), a regression analysis was performed to compare the number or runs scores as a function of the batting average, and then as a function of the on base percentage and slugging percentage. A sampling of the moneyball data used for the analysis is shown below.

A linear regression analysis was first performed on the RS vs. BA. The results, which are plotted in the figure below, showed that the correlation between the RS and BA was only 0.69. BA was deemed a marginally useful statistic because it does not account for players hitting singles versus home runs and does not account for players getting on base by walks or being hit by a pitch.

By contrast, a linear regression between the RS and OBP shows a correlation of $r^2 = 0.82$. OBP accounts for all the ways a player can get on base, and as such, provides a more meaningful measure of the number or runs scored than does the batting average.

Finally, a multivariate linear regression was performed with the RS vs. the OBP and the SLG. The results of this linear regression showed a correlation of $r^2 = 0.93$, meaning that OBP combined with SLG provided a better indicator or run scoring performance than BA or the OBP by itself. It should be noted that the linear combination of OBP and SLG is called On-base Plus Slugging (OPS), and is a commonly used baseball statistic in the game today ($OPS = OBP + SLG$). With this measure of OPS, the amount of time a player reaches base is accounted for as well as how many bases they are able to reach when they do get on base.

Using these data science techniques, Beane and DePodesta and the Oakland A's were able to win 103 games in 2002 (including a record-setting 20-game win streak), finish in first place, and make the playoffs. Today, OPS and OBP and SLG are some of the most closely watched baseball statistics by baseball insiders and fans alike.

1.11.3 *Predicting Patient-Specific Scoliosis Curvature (Mixed Data Science and Surrogate: Type 2)*

Mechanistic data science can be used to analyze the progression of Adolescent Idiopathic Scoliosis (AIS), and someday soon provide a way to virtually assess the effectiveness of patient-specific treatments before starting the actual treatment. AIS is a condition in which the adolescent spine curves in an unnatural manner. Recently, mechanistic data science has been used to study the progression of this condition.

Multimodal Data Generation and Collection The analysis and diagnosis of AIS begins with medical imaging of the spine. Two types of images are used for the analysis—X-rays and magnetic resonance imaging (MRI). X-rays of a patient are taken from the front or anteroposterior (AP) view and the side or lateral (LAT) view to capture the position of the vertebrae that make up the spine (see Fig. 1.14). The X-rays are repeated to document the progression of the scoliosis condition over time. An outline of each vertebra can be extracted from the 2D X-ray data. The 2D data points are projected to 3D data.

In addition to X-rays, MRI's taken of a few patients provide a detailed 3D image of the entire spine. The MRI data from one patient's spinal vertebrae can then be used as reference or surrogate images for other patients. The surrogate model of the vertebrae can be adjusted to be patient-specific by combining it with data from the 2D X-rays.

The 2D X-ray data points for each vertebra that have been located in 3D space are overlaid on the surrogate model of the vertebra. For this analysis, the surrogate model of the vertebrae is taken from the MRI data of the spine, but there are other

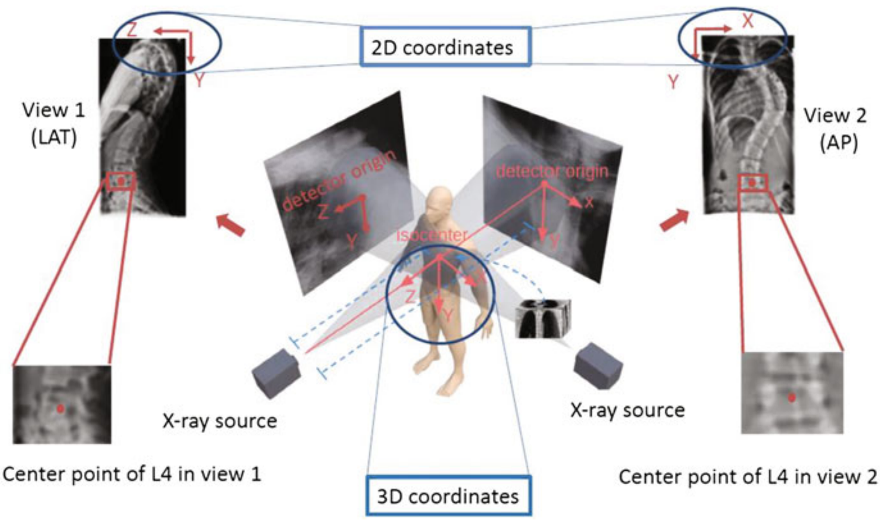


Fig. 1.14 X-ray projections for collecting data on scoliosis progression

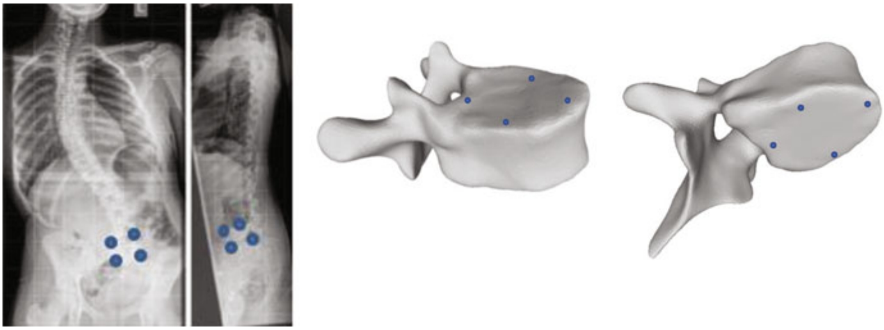


Fig. 1.15 Surrogate geometry of a vertebra. The vertebra on the left is before being adjusted by the collected data. The vertebra on the right has been adjusted through the collected data

sources of vertebrae data that can also be used. Once the 3D X-ray data has been overlaid on the surrogate model of the vertebra, the surrogate model is scaled and adjusted to yield a patient specific model of the vertebra (see Fig. 1.15). This process is repeated for each vertebra until the entire spine has been mapped for a particular patient by combining 2D X-ray images with generic models from a surrogate.

Extraction of Mechanistic Features Key reference points, or landmarks, for each vertebra are also extracted from the 2D X-rays using image processing. The intersection of the two 2D projection is then used to locate the data points as 3D data points (see Fig. 1.15).

Knowledge-Driven Dimension Reduction Using the 3D simplified model of the spine derived from the X-ray images, a detailed model of the spine was created using the atlas model vertebra. The generic surrogate model of each vertebra is updated based on the actual vertebra size and shape as shown in Fig. 1.15. These more detailed vertebrae are assembled to form a patient-specific spine model.

Reduced Order Surrogate Models The patient specific geometry of the spine can be used to generate a finite element model to compute the pressure distribution on each vertebra due to scoliosis. According to the Hueter-Volkmann (HV) principle, areas of a vertebra with higher stress grow more and areas with lower stress grow less. The pressure distribution computed using finite element analysis and the geometries of the vertebrae are updated based on the HV principle. The gravity load and the material properties of the spine material are also updated over time to reflect the changes in a specific patient due to aging. The computed stress results can be used as the input to a neural network (Fig. 1.16), along with other factors such as the landmarks, the global angles, and the patient age, to predict how the spine would move over time. At this time, not enough information is known about modeling the materials of the growing spine and it is not possible to measure the pressure distribution of the vertebrae pressing on each other. However, the results below show that through a combination of finite element computer simulation to compute the pressure distribution on the vertebrae and the data from the patient X-rays, it is possible to accurately predict the progression of scoliosis (Figs. 1.16 and 1.17).

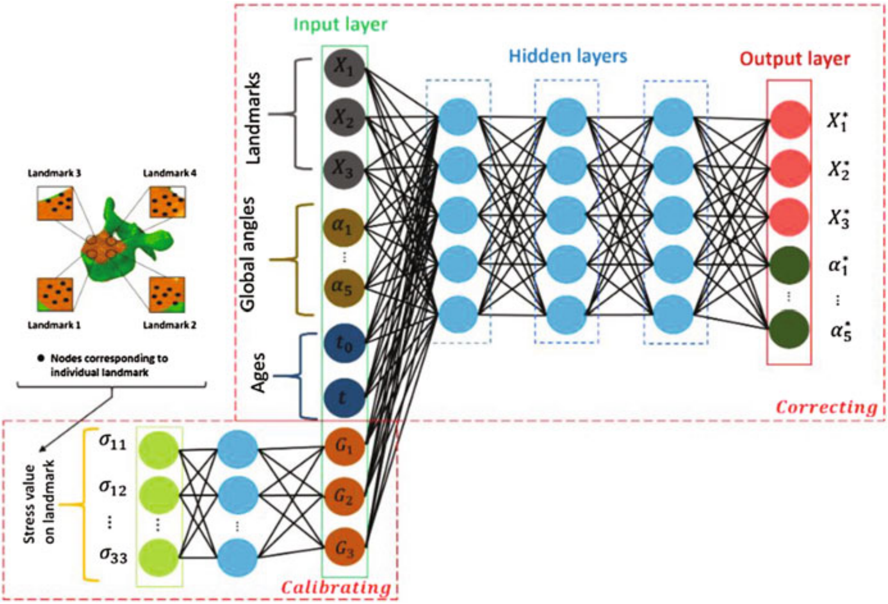


Fig. 1.16 Neural network combining mechanistic models with X-ray data to predict spine growth in scoliosis patients

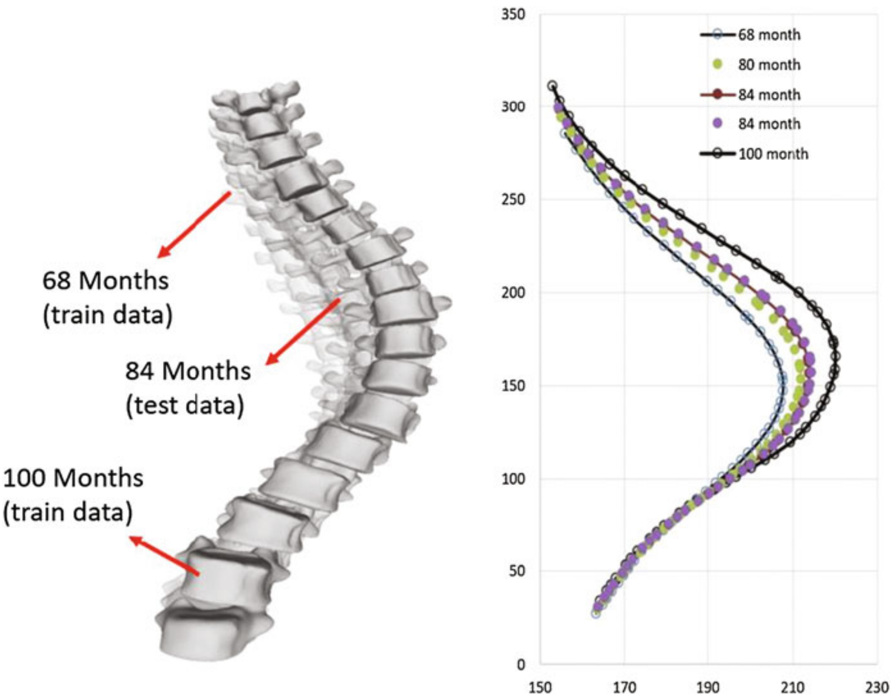


Fig. 1.17 Predicted progression of spine growth in scoliosis patients

1.11.4 *Identifying Important Dimensions and Damping in a Mass-Spring System (Type 3 Problem)*

A young engineer is given the job of running experiments with physical components and then analyzing the data. Unfortunately, the test data does not always clearly match with the theory learned in school. This is not a trivial problem, but rather a fundamental challenge in empirical science. Examples abound from complex systems such as neuroscience, web indexing, meteorology, and oceanography - the number of variables to measure can be unwieldy and, at times, even deceptive, because the underlying relationships can often be quite simple.

One such example is a spring-mass system shown in Fig. 1.18. The engineer learned in school to model a system like this with an ideal massless spring, assuming that all the mass at a point and no mass or damping in the spring. For an ideal system, when the weight is released a small distance away from equilibrium (i.e. the spring is stretched), the ball will bounce up and down along the length of the spring indefinitely. The frequency of the motion will be constant based on the stiffness of the spring and the mass of the attached weight. An actual spring-mass system does not perfectly match the ideal specifications since the attached mass is not a point mass, and the spring has some mass, and there will be some damping due to friction.

Multimodal Data Generation and Collection The engineer needs to make measurements of an actual spring-mass system to determine the frequency of the motion and the amount of damping in the spring. To make these measurements, he decided to record the position of the weight from three different angles and orientations using video cameras, which would record the position of the weight in three dimensions

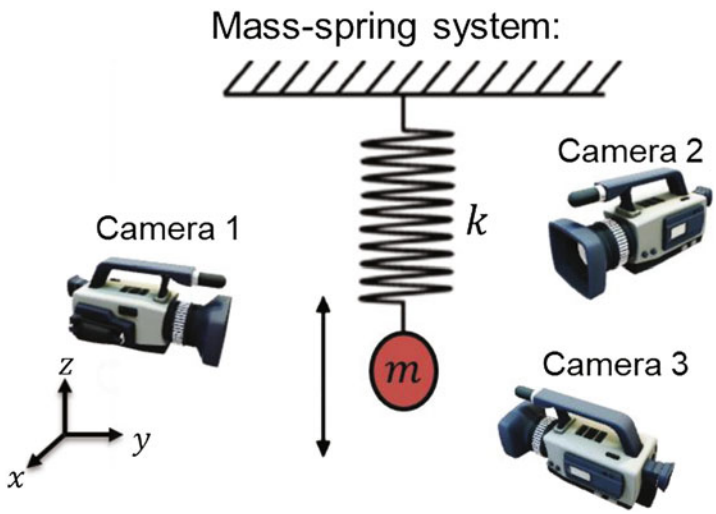


Fig. 1.18 A spring-mass motion example. The position of a ball attached to a spring is recorded using three cameras 1, 2 and 3. The position of the ball tracked by each camera is depicted in each panel. (a video is available in the E-book, Supplementary Video 1.2)

(since the world is inherently a three-dimensional world). He placed three video cameras around the spring-mass system and recorded the motion at 120 frames per second, which provided three distinct projections of the two-dimensional position of the ball. Unfortunately, the engineer placed the video cameras as three arbitrary locations, which meant that the angles between the measurements were not necessarily at right angles! After recording the motion, the engineer was faced with the big question: how to get one-dimensional motion data from the two-dimensional projection data collected from three different angles?

Extraction of Mechanistic Features The engineer first needed to obtain digital data from the videos that had been recorded. To do this, the motion of the weight is extracted from the videos using the computer vision motion capture capabilities of Physlet Tracker [9]. The motion data from each camera is shown in the lefthand plots in Fig. 1.19. To aid in further analysis, the data from each camera is centered on the mean.

Knowledge-Driven Dimension Reduction The useful data for analyzing the spring mass system is the motion of the weight along the orientation of the spring. To extract the 1D data along the longitudinal direction of the spring, the dimension reduction techniques of singular value decomposition (SVD) and principal component analysis (PCA) can be used (these techniques will be described in detail in Chap. 5). These techniques evaluate which motion directions are key and compute the data associated with those directions. A dimension reduction of the data is achieved by only maintaining a few of these key motion components.

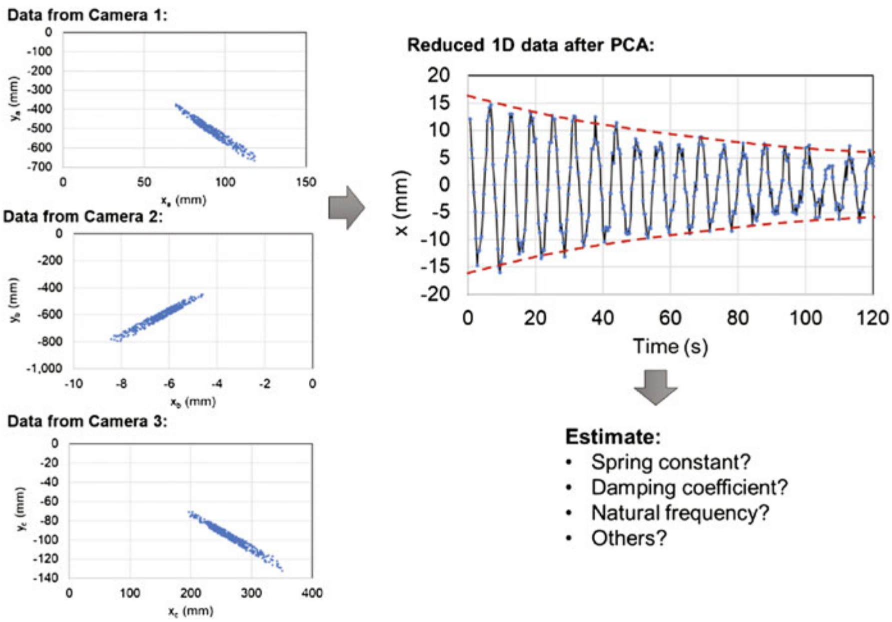


Fig. 1.19 An example reducing high-dimensional data to 1D data and then estimate important parameters in the system

Reduced Order Surrogate Models After the dimension reduction, the engineer created a reduced order model with a smaller number of features can be developed to represent all of the original data. For this spring mass system, one dimensional data for the motion along the longitudinal direction of the spring is desired to represent the original set of data from the three cameras. This is possible because the displacements from different local coordinate systems are highly correlated since they recorded the same spring-mass.

Regression and Classification The engineer then used the data from the spring-mass experiments to determine the natural frequency and damping coefficients by considering the mechanistic principles of a spring-mass system. A spring-mass system will oscillate in a pattern that matches a sine wave. If the system has some damping, the sine wave can be multiplied by an exponential function

$$z(t) = A * \sin(2\pi f t) * \exp(-bt)$$

where A is the starting amplitude, f is the natural frequency in Hz, and b is the damping coefficient. The natural frequency is the inverse of the peak-to-peak distance of the sine wave. The damping coefficient is computed by considering the rate of the exponential decay of the sine wave.

Systems and Design The engineer was able to use the reduced order model of the spring, along with the regression data based on mechanistic principles, to determine the critical properties of the spring-mass system. The data show that the natural frequency is $f = 0.158$ Hz, and the damping coefficient is $\gamma = 0.008$. These data are overlaid on the reduced order model data in Fig. 1.20.

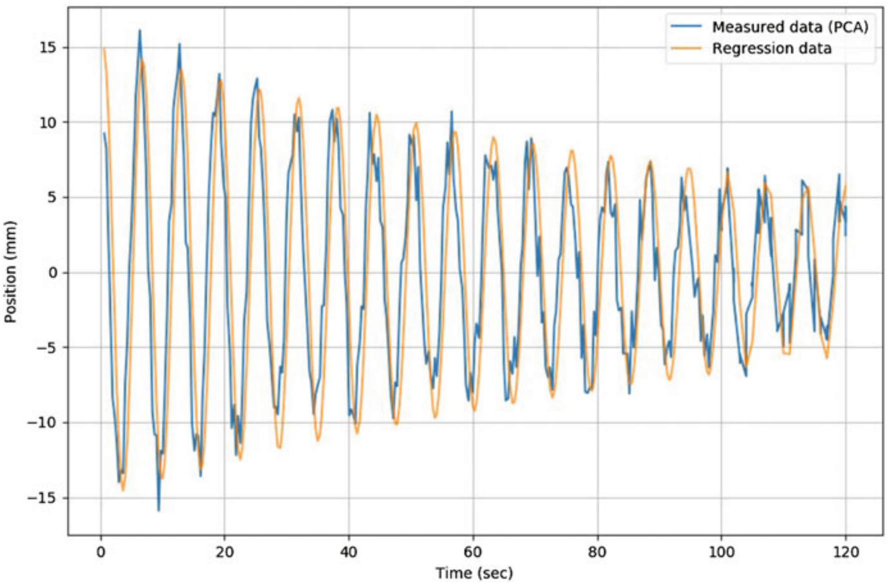


Fig. 1.20 Spring-mass motion data with damped sinusoidal plot overlaid

References

1. Galileo's Leaning Tower of Pisa experiment. Wikipedia. 2020 [Online]. https://en.wikipedia.org/w/index.php?title=Galileo%27s_Leaning_Tower_of_Pisa_experiment&action=history. Accessed 8 Sep 2020.
2. https://nssdc.gsfc.nasa.gov/planetary/lunar/apollo_15_feather_drop.html
3. Galileo, Opere Il Saggiatore.
4. Aloha Airlines Flight 243. Wikipedia. 2020. https://en.wikipedia.org/wiki/Aloha_Airlines_Flight_243. Accessed 8 Sep 2020.
5. NAE Report on Making a World of Difference-Engineering Ideas Into Reality.
6. https://3dprintingindustry.com/news/eos-cooksongold-put-bling-3d-printing-precious-metal-printer-33041/
7. https://www.spineuniverse.com/resource-center/spinal-cancer/3d-spinal-implants-glimpse-future
8. https://www.diamonds.pro/education/gia-certification/
9. https://physlets.org/tracker/