

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

**APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN
DEVELOPMENT OF AN ANALGESIC
SOLID DOSAGE FORM AND PREDICTION OF ANALYSIS
RESULTS**

ESİLA DOLDUR
**A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE
DEPARTMENT OF CHEMICAL ENGINEERING**

GEBZE
2023

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

**APPLICATION OF ARTIFICIAL NEURAL
NETWORKS IN DEVELOPMENT OF AN
ANALGESIC SOLID DOSAGE FORM AND
PREDICTION OF ANALYSIS RESULTS**

ESİLA DOLDUR
**A THESIS SUBMITTED FOR THE DEGREE OF
MASTER OF SCIENCE**
DEPARTMENT OF CHEMICAL ENGINEERING

THESIS SUPERVISOR
ASST. PROF. DR. ÖZGÜN YÜCEL

GEBZE
2023

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 02/02/2023 tarih ve 2023/09 sayılı kararıyla oluşturulan jüri tarafından 21/02/2023 tarihinde tez savunma sınavı yapılan Esila DOLDUR'ın tez çalışması Kimya Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Dr. Öğr. Üyesi Özgün YÜCEL

ÜYE

: Dr. Öğr. Üyesi Hasan ŞILDIR

ÜYE

: Doç. Dr. Sedat YALÇINKAYA

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

SUMMARY

In this study, it is aimed to make the drug product development process fast, effective and reliable within the scope of R&D studies in industry of pharmaceutical, which is the most important work area in Türkiye and in the world. Studies have been implement on the estimation of the physical stability of solid dosage forms with the machine learning method, which make possible repeated stability studies to be revised more readily and quickly

In this thesis, at the first stage, the increase in hardness, which is a significant problem in stability of drug product, accordingly the increase of moisture are discussed. The drug product, which was stabilized frequently and for a short time, was exposed to temperatures with different packages. At the same time, considering that the parameter affecting the stability of the DP may the moisture value in the in-process stage, it analyzed for different granule moisture values. The data obtained affect both the determination of the chemical analysis process in R&D activities and the "time & cost" parameters, which are one of the most important parameters of the development process for the industry. At the same time, the data obtained were used in the second stage studies.

In the first stage, data were collected by performing analyzes for three parameters (moisture, dispersion and hardness). In the second stage, the estimation of these data was provided. For this purpose, regression analyzes were performed. Then, it was estimated using machine learning methods "Decision Tree", "Random Decision Forest" and "Generalized Additive Model".

Key Words: Pharmaceutical Industry, Drug Stability, Machine Learning, Decision Tree, Generalized Additive Model, Random Decision Forest.

ÖZET

Bu çalışmada ülkemiz ve dünyada en önemli sektörlerden biri olan ilaç sektöründe, Ar-Ge çalışmaları kapsamında ilaç ürünü geliştirme sürecini hızlı, etkin ve güvenilir kılmak adına; tekrarlanan stabilite çalışmalarının daha kolay ve hızlı bir şekilde revize edilmesini sağlayacak olan makine öğrenmesi yöntemi ile katı dozaj formlarının fiziksel stabilitesinin tahmin edilmesi üzerine çalışmalar yapılmıştır.

Bu tezde ilk aşamada, ilaç ürününün stabilitesinde belirgin bir sorun olarak üzerine çalışılan sertlik artışı ve buna bağlı olarak nem artışı ele alınmıştır. Sık ve kısa süreli stabiliteye kaldırılan ilaç ürünü, farklı ambalajlarla, farklı sıcaklıklara maruz bırakılmıştır. Aynı zamanda ilaç ürününün stabilitesini etkileyen parametrenin, in-proses aşamasında nem değeri olabileceği düşünülerek, farklı granül nemi değerleri için de analizlenmiştir. Elde edilen veriler, hem Ar-Ge faaliyetlerinde kimyasal analiz sürecinin belirlenmesinde, hem geliştirme sürecinin sektör için en önemli parametrelerinden biri olan "zaman & maliyet" parametrelerini etkileyecektir. Aynı zamanda elde edilen veriler, ikinci aşama çalışmalarımızda kullanılmıştır.

Birinci aşamada üç parametre için (nem, dağılma ve sertlik) analizler yapılarak veri toplanmıştır. İkinci aşamada ise bu verilerin tahmini regresyon analizleri, makine öğrenmesi metotları olan "Karar Ağacı", "Rastgele Karar Ormanları" ve "Genelleştirilmiş Toplamsal Model" kullanılarak tahmin edilmiştir.

Anahtar Kelimeler: İlaç Sektörü, İlaç Stabilitesi, Makine Öğrenmesi, Karar Ağacı, Genelleştirilmiş Toplamsal Model, Rastgele Karar Ormanları.

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my supervisor, Asst. Prof. Özgün YÜCEL, who not only shared his profound scientific knowledge with me but also taught me great lessons of life. His support, suggestions and encouragement gave me the drive and will to complete this work.

I wish to express my sincere thanks to my manager, Ms. Seda ÇALIŞKAN, for their collaboration for this research.

I would also like to thank my family for their love and faith in me.

Finally, I want thank my precious husband Batuhan DOLDUR, who showed me the beauty of love.

TABLE OF CONTENTS

	<u>Page</u>
SUMMARY	iv
ÖZET	v
ACKNOWLEDGEMENTS	vi
TABLE OF CONTENTS	vii
LIST of ABBREVIATIONS AND ACRONYMS	ix
LIST OF FIGURES	x
LIST OF TABLES	xi
1. INTRODUCTION	1
1.1. Hypothesis	4
1.2. Objective	5
2. HISTORICAL BACKGROUND	6
2.1. Drug Product Approvals	8
3. METHODOLOGY OF DRUG PRODUCTS	10
3.1. Analgesic Drug Products	11
3.2. Components of Drug Products	11
3.3. Excipient Selection	14
3.4. Stability of Drug Products	15
3.5. Machine Learning Method For Drug Stability	17
3.6. Decision Tree	19
3.7. General Additive Method	20
3.8. Random Forest	20
4. RESULTS AND DISCUSSION	21
4.1. Instruments	21
4.1.1. Moisture Content Analysis	21
4.1.2. Disintegration Analysis	23
4.1.3. Hardness Analysis	24
4.2. Collection Of Data	26
4.2.1. First Trial Batch Stability Data	27
4.2.2. Second Trial Batch Stability Data	29

4.2.3. Third Trial Batch Stability Data	30
4.3. Formulas	31
4.4. Results	31
4.4.1. Hardness Results	31
4.4.2. Moisture Content Results	37
4.4.3. Disintegration Results	43
5. CONCLUSION	49
REFERENCES	52
BIOGRAPHY	57
APPENDICES	58



LIST OF ABBREVIATIONS AND ACRONYMS

<u>Abbreviations</u>	<u>Explanations</u>
<u>and Acronyms</u>	
API	: Active Pharmaceutical Ingredients
CAGR	: Compound Annual Growth Rate
DP	: Drug Product
DT	: Decision Tree
EMA	: European Medicines Agency
FDA	: Food & Drug Administration
GAM	: Generalized Additive Model
ICH	: International Council of Harmonization
ML	: Machine Learning
QSAR	: Quantitative Structure Activity Relationship
RF	: Random Forest
WHO	: World Health Organization

LIST OF FIGURES

<u>Figure No:</u>	<u>Page</u>
2.1: Timeline indicating the years meanwhile many of the pharmacopoeias were performed.	8
3.1: Excipient classification for selection.	14
3.2: ICH stability climate conditions and zones in world.	16
3.3: A representative binary decision tree: simple univariate.	19
4.1: Device used for moisture content analysis.	21
4.2: Moisture content analysis.	22
4.3: Device used for disintegration analysis.	23
4.4: Disintegration analysis.	24
4.5: Hardness analysis.	25
4.6: Device used for hardness analysis.	26
4.7: Hardness analysis results obtained using the random forest model.	36
4.8: Hardness analysis results obtained using the decision tree model.	36
4.9: Hardness analysis results obtained using the general additive model.	36
4.10: Moisture content (KF) analysis results obtained using the random forest model.	42
4.11: Moisture content (KF) analysis results obtained using the decision tree model.	43
4.12: Moisture content (KF) analysis results obtained using the general additive model.	48
4.13: Disintegration analysis results obtained using the random forest model.	48
4.14: Disintegration analysis results obtained using the decision tree model	48
4.15: Disintegration analysis results obtained using the general additive model.	49

LIST OF TABLES

<u>Table No:</u>	<u>Page</u>
3.1: Excipient classification for selection.	14
3.2: Proposed conditions for long-term stability test by stability zones.	16
4.1: Process parameters for three trial batches.	27
4.2: In process control analysis.	27
4.3: First trial batch stability analysis results.	28
4.4: Second trial batch stability analysis results.	29
4.5: Third trial batch stability analysis result.	30
4.6: Formulas that used for study.	31
4.7: Data collection for hardness results.	32
4.8: Regression analysis results.	35
4.9: ANOVA analysis.	35
4.10: Analysis moisture content results.	35
4.11: Data collection for moisture content result.	38
4.12: Regression analysis results.	41
4.13: ANOVA results.	41
4.14: Parameters of moisture content (KF) analysis results.	41
4.15: Data collection for disintegration analysis.	44
4.16: Regression analysis result.	47
4.17: ANOVA results.	47
4.18: Parameters of disintegration analysis results.	47
5.1: Summary of regression and machine learning analysis results.	50

1. INTRODUCTION

The pharmaceutical industry is one of the industries where research is most intense. The pharmaceutical industry is a sector that is constantly trying to produce new products with the mission of saving lives and improving the quality of life. Basic scientific knowledge is very important in new drug discovery. In this context, connections were established with places where basic research was carried out, such as national laboratories and universities, in the final product development processes.

In the simplest terms; Pharmaceutical industry is a unit that discovers or develops pharmaceutical products applied to errors and to cure patients or relieve some symptoms of discomfort. At the same time, the pharmaceutical industry markets the products it develops as it produces [1].

According to the "Global Pharmaceutical Market Report" research prepared in 2020, global pharmaceutical market, produced end-products worth 1,228 billion dollars and showed a CAGR of 1.8% [2]. Pharmaceutical technology development units and R&D studies of the pharmaceutical industry continue in Türkiye and in the world. Pharmaceutical development is of great importance in these studies. Development work is extensive and strategically organized.

The guidelines ensure only the common necessity for, the rating of both new drug substances and new drug product (ICH, International Conference on Harmonization). WHO (World Health Organization) guide-lines to strived both active pharmaceutical ingredients (APIs) and finished pharmaceutical products (FPPs). While others, such as the FDA (Food and Drug Administration), assist in the preparation of more specific analytical procedures, the mission of the European Medicines Agency (EMA) is to promote scientific excellence in the evaluation and inspection of drugs for the benefit of public and animal health. in the European Union (EU) [3].

It is very important that the active substance be practicable in the pharmaceutical development phase in Türkiye and in the world. For approval of an active substance, and that active substance is new and valid focuses on guidelines. However, there are several older generation compounds still in use in contemporary medical practice [4].

Therefore, Some new developments in analytics should cover the entirety, not just new products under development. These analytical developments should cover stability issues for both new and legacy compounds. Stability studies have a great place in analytical development. One of the most basic requirements for pharmaceutical development and commercialization is the stability of the drug formulation.

For decades, manufacture of pharmaceutical products has been controlled by a regulative construction that protects quality of final products by testing raw materials, in-process materials, final product properties, and drug stability. Drug stability, which is one of the critical-importance quality and performance characteristics, needs to be evaluated during pharmaceutical development. It also needs to be controlled during clinical research and marketing. Stability studies are a parameter that can change the quality of a drug substance (active pharmaceutical ingredient, also called API) or a drug product (DP) depending on the controlled (ICH stability zones determined according to the climate assessments of the countries.) temperature and relative humidity provides evidence of how it changes over time. This helps determine the re-test day (for API) or shelf life and storage situation.[5].

Determination of instability from the pharmaceutical quality side, it involves performing a large number of chemical as well as physical tests that are specifically disclosed in Pharmacopeias and regulations by other industrial boards. Since the determination of the physical stability affect the decision of the chemical stability follow-up, the physical stability follow-up is important in terms of cost, time and planned work. Stability is also described as a state, state, where the pharmaceutical active ingredient, as well as the drug product, does not deteriorate as a result of exposure to certain factors or remains intact throughout its shelf life. If , raw material does not deteriorate, the follow-up of both physical and chemical tests are compatible. Follow-up of physical tests is important at this stage. Physical stability monitoring is less costly and it would be more logical to specify chemical stability monitoring to evaluate its results.[6].

Physical stability, thrugout with chemical stability or stability of microbiological, has long been of industrialist as well regulatiative concern. Appropriate stability of physical is required as dosage form to achieve its intended performance throughout along its shelf life. Physical changes, with the sense import organs such as color change, can impress the perceived efficacy and consumer user

reliance of a drug product, even without changing drug strength. In the researches, the main physical properties of common pharmaceutical dosage forms are summarized as follows: appearance and odor are paramount in stability for all dosage forms. Dispersion, moisture and hardness parameters are important for solid dosage forms. For dosage forms such as suspension, emulsion, inhalation, nasal aerosol/spray, liposome and other nano-carriers, pH, viscosity, and particle size distribution are essential.

Now, physical stability of drug substances requires at least three - six months with trial-and-error experiments that are time-consuming and incalculable. It fails, the long loop is retested. Moreover, the mechanism of physical stability of solid dosage forms still not fully understood. In recent years, various theories regarding the stability of drug substances such as solubility parameters and Tg prediction model have been disputed. However, these theoretical models ask for a large amount of physicochemical information notice and plenty professional knowledge for each component. Also, the predictive ability of these models was reasonably limited by unchecked error due to the mathematical hypothesis [7].

Machine Learning algorithms are capable of "learning" and predicting complex systems from large amounts of data. In recent years, ML has been successfully applied to pharmaceutical research such as quantitative structure activity relationship (QSAR), drug-drug interaction, drug discovery and pharmacogenomics.

In the domain of formulation, successfully advanced high-correctness models to soothsay the disintegration time of orally disintegrating solid tablets and orally disintegrating film tablets with Machine Learning method, disintegration curves of oral sustained-release tablets, and the binding free energy of the complex active pharmaceutical ingredients such as anticancer drug-carrying cells by ML techniques. In addition, transfer learning and multitasking approximations of learning have been developed to forecast humanity pharmacokinetic parameters [8].

1.1. Hypothesis

It is important to improve economic conditions and save time in the pharmaceutical industry. Accurate, reliable and rapid development of a drug product is essential. For this reason, the possibilities offered by today's technology open some

new doors. One of them is the development of a new predictive model for the physical stability of solid dosage formulations together with machine learning techniques.

In this way, it is advantageous in terms of both time and cost. By predicting the physical stability of the drug with machine learning method, it makes predictions for product development before the long-term stability process, and accurate, fast and reliable formulation development times are shortened in the pharmaceutical industry.

In this study, the aim is to make the drug development process fast, effective and reliable; The aim is to predict the solid dosage forms physical stability with the machine learning method, which allow repetitive stability studies to be revised more easily and quickly.

In this study, the stability study of this analgesic solid dosage form product, which has been used continuously for many years in Türkiye and in the world, was discussed. The final drug product is a centrally acting pain reliever used to treat muscle spasm and the resulting pain or discomfort.

1.2. Objective

The aim of this study is to predict the physical stability of an immediate-release analgesic drug product in solid dosage form by machine learning method. Every area of life is constantly open to change. One of the main purposes of humans is to control these changes for our own benefit. This is especially important in the field of medicine and pharmacy.

The physical stability parameters considered are hardness, dispersion and moisture (LOD). In physical stability, the estimation of these parameters for solid dosage forms provides advantages in terms of time and cost, as it provides the estimation of both chemical stability parameters and stability processes in the pharmaceutical industry.

As it is known, the presence of moisture affects the stability of tablet products. The amount of moisture affects the stability of the tablets. Too much moisture grounds powder particles to clump together and form a weak, crumbly tablet; too little moisture grounds the tablet to disintegrate.

Tablet hardness testing is important in that: As the compression force increases, the hardness of the tablet increases and the porosity decreases, which can result in reduced water ingress and slower disintegration rate. Therefore, in stability, the

hardness of a tablet can be a risk factor. In this study, hardness analysis was carried out for short periods to see whether can be observe a hardness change in the stability of product in solid dosage form. The product has been subjected to stability condition at different packaging and temperatures. By evaluating the results in this way, it was tried to reach the source of any problem.

Disintegration is a solid dosage form indicates that the tablet or capsule must first be broken into smaller pieces. The disintegration test accurately measures the ability of a sample to break up into smaller particles under standard conditions. This test is typically performed on tablets, capsules, and enteric-coated tablets. The disintegration result of the final product is very important in in-process controls and stability controls. A final product with an out-of-limit dispersion parameter not exhibit correct disintegration in the mouth, stomach and intestinal environment at the analytical stage.

2. HISTORICAL BACKGROUND

In the early 1800s, pharmaceutical industry activities began with the applied and research discoveries of herbal medicines such as morphine and quinine. One of the first pharmaceutical products, Morphine, an analgesic and sleep-warm spy, was explored by the German pharmacist Sertürner as well entitled afterwards Morpheus, the Greek god of imaginations.

In the 1880's, the purification of organic compounds reached a good level, and the science of pharmacology emerged with the development of synthetic chemicals.

In the 1890's, adrenal extracts were developed and the blood pressure-raising effects of these products were investigated. The effect of this product in the treatment of shock has attracted the attention of surgeons. Epinephrine was discovered, which saved the lives of many people, by identifying the active substance as epinephrine, by John Abel of J. Hopkins University.

In the 1900's, with the discovery of Amphetamine; A drug has been discovered that is indicated for narcolepsy, post-encephalitic parkinsonism, and collapse and other psychiatric signs. It admitted approval from the American Medical Association as a New and Unofficial Solution for use in 1937. Amphetamine was successfully sold as the first antidepressant in the late 1930's and 1940's, in common with a private understanding of depression as anhedonia [9].

Such a development of the pharmaceutical industry ignored possible risks. The "Thalidomide" disaster that occurred in the 1960s caused many legal arrangements. Thousands of newborns in Europe and elsewhere have suffered from the teratogenic effects of thalidomide. The thalidomide tragedy created the mandate for the FDA to establish "Good Manufacturing Practices." According to this law, all drugs produced between 1938 - 1962 should have been effective, and about 40% did not prove effective [10].

For these reasons, a series of legal regulations and pharmacopoeias have emerged on the way of developing new products or becoming a generic manufacturer. Globally, pharmacopoeia harmonization processes still continue. In this way, vital quality data such as the efficacy and safety levels of drugs and the way they reach patients are carefully examined.

Over the years, due to the fact that each country has its own pharmacopoeia, efforts to go to one type of pharmacopoeia have begun to accomplish the nuisances and perils of the wealth of three different pharmacopoeias in the UK and Ireland, London, Edinburgh and Dublin.

The British Pharmacopoeia (BP) was created in 1858. Even if earlier, in 1820's, the United States Pharmacopoeia (USP) was founded by a group of 12 physicians. Because they get to know the need for coherent standards for the pharmacopoeia. Modern pharmacopoeias become distant of from historical "recipes" that describe development methods for drugs, instead focusing on the properties of drug products and the ingredients that help ensure their identity, potency, quality. The foreword first edition of the USP addresses the need and value of harmonization to ensure consistency in pharmacopoeial standards [11].

An even broader view of today's global community is further complicated by problems such as the various pharmacopoeias around the world, and due to a wide and overwhelming lack of alignment among regulatory agencies. The publication of the USP in 1820 and the BP (abbreviated for the British Pharmacopoeia) in 1858 aimed to bring consistency and more controlled and quality-oriented work to all drug standards, which is of great benefit to pharmacy, medical practitioners and the healing of their patients. This goal is very important. It has been followed by other countries and has kept pace with modern times and even continued until modern technological times with activities such as initial development and subsequent updates to their pharmacopoeias. As mentioned earlier, there are about 40 pharmaceutical pharmacopoeias published around the world live in, each one rich in its own unique and its own history, while also sharing the common goal of providing quality standards for drugs that benefit patients and provide good returns. exists.

As can be seen in the Figure 2.1. below, timeline information of pharmacopoeias that regulate drug development is given. In today's drug development, mostly American Pharmacopoeia (USP) and European Pharmacopoeia (EP) are taken into consideration.

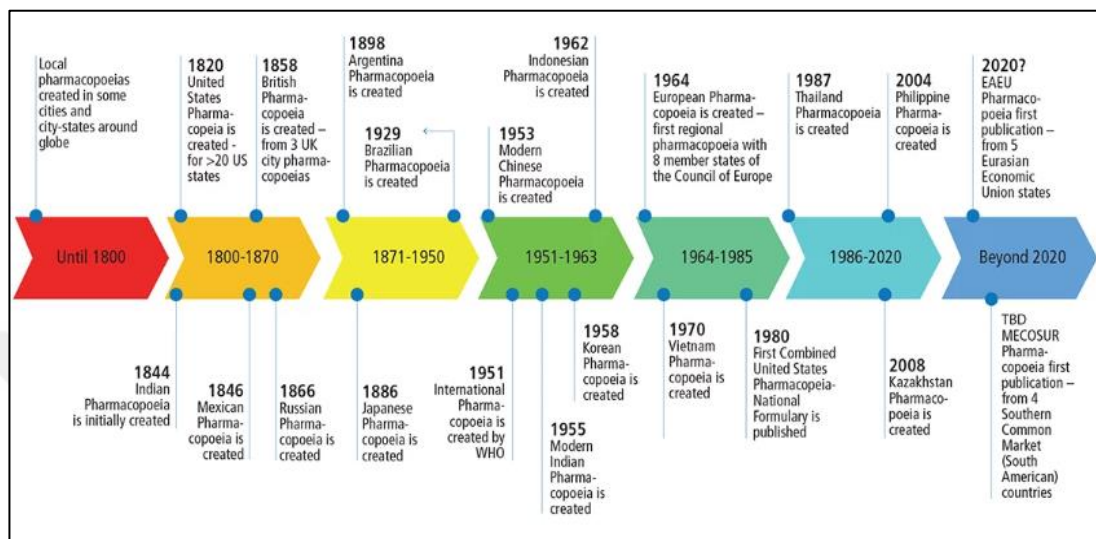


Figure 2.1: Timeline Indicating the years meanwhile many of the pharmacopoeias were performed.

2.1. Drug Product Approvals

It applies the knowledge obtained as a result of research to develop products and obtain market approval, and therefore invests in clinical studies.

The agency known as the Food and Drug Administration (called the FDA for short) is one of the most important organizations to self-approve a targeted drug product that is both safe and effective. In short, FDA requires a research dossier containing sufficient pre-clinical data and periodic stability analyzes of the product for the approval process stages. In clinical studies; Phase I studies toxicity, usually using healthy volunteers.

Phase II is important. Phase II patients may include pharmacokinetic mechanism of action and dosing, and Phase III is a very large efficacy study in the intended patient population. The FDA reviews the data and if the product comes back positive with benefit-risk data, it is approved for marketing in the USA.

Before approving a product, the drug product must be both safe and effective. It should determine whether the drug product is for use in the specified conditions, according to the product label and the instructions it contains [13].

These are to consistently obtain a drug product of acceptable and stable quality so that the applicant for the drug product is both safe and effective for use. It contains product quality requirements that demonstrate that a drug product and active substance, manufacturing process, and control strategy have been developed.

Secondly, European Medicines Agency, EMA for short, is an innovative and high-tech organization valid in the EU. Responsible for scientific evaluation of drugs. This organization cooperates with experts from scientific committees and working groups. Some of the guidelines offered by EMA reflect the latest and most current thinking about developments in the biomedical and pharmaceutical sciences.

It guides the development programs of all drug developers and R&D who wish to apply for marketing and regulatory clearance. It is always available to ensure stable (physically, chemically and microbiologically stable) and high quality development for all conditions. EMA, has developed a gathered list of existente guidelines and opportunities for interplay of during the development of a medicinal product.

The EMA provides drug developers with several opportunities for early dialogue and consultation before submitting a marketing authorization application. This aims to provide regulatory and scientific support to facilitate the preparation of applications and ensure a smooth verification and evaluation procedure [3].

In order for the product to be approved by the authorities and to initiate phase studies and clinical trials, the product must be effective, high quality and safe. It is determined by a series of analyzes of this product, and certain analysis results are expected to be appropriate within certain specification ranges of the product. In this context, the product must be stable throughout its promised shelf life.

This organization is defined as the International Council for Harmonization of Technical Requirements for Medicines for Human Use, and ICH for short, which is the organization that brings together pharmaceutical industry activity groups and some important regulatory authorities. It is an initiative that brings together regulatory authorities and the pharmaceutical industry to discuss the scientific and technical aspects of pharmaceutical product development and registration.

ICH's mission is to improve public health through greater compliance by developing technical Guidelines and requirements for pharmaceutical product licensing [14].

The mission of ICH is to provide recommendations for greater compliance in the interpretation and fine implementation of some technical guidelines and some requirements for the substantive and rigorous registration of pharmaceutical products and the very good care of pharmaceutical products. In addition, the task of this organization is to monitor and update harmonized technical requirements and Update new developments, leading to greater mutual acceptance of research and development data. or replace existing apps. or to facilitate the adoption of advanced technical research and development approaches.

It addresses stability testing and results of new of substances of drug and products with the stability guide, the analytical validation guide and the factual experimental data necessary with statistical interpretation to validate analytical procedures, pharmaceutical development and harmonization, and good manufacturing practices [15].

3. METHODOLOGY OF DRUG PRODUCTS

A drug product provides the performance intended by its developer in an effective and stable manner and creates a quality product and production process for this; important for pharmaceutical development. Knowledge from manufacturing studies and trials, pharmaceutical development studies, specifications, and generation of manufacturing information provides scientific understanding.

Looking at all pharmaceutical product/product development; Changes in production processes, ie change of production site (including changes in factory relocation, device changes, etc.) should be considered as opportunities to gain additional knowledge and further support the creation of the required design space.

In fact, it can be very important to incorporate knowledge from trial studies that yielded unexpected problems into the development phase. Aspects of drug substances, excipients, container closure systems and manufacturing processes critical to product quality, safety, efficacy and stability should be identified and control strategies justified [16].

The drug quality and critical quality characteristics of a pharmaceutical product are, at a minimum, defined by the excipients, the packaging, the container-lid systems, and the aspects of the manufacturing processes.

Pharmaceutical product control strategies must be justified, which is a critical element. Critical formulation characteristics and process parameters are often defined by evaluating the extent to which their variations can have an impact on the quality of the drug product [17].

3.1. Analgesic Drug Products

The pain, feel in our body is one of the most important symptoms of diseases and some injuries. It can be felt in many organs or skin. Pain can be examined in two stages: chronic pain and acute pain. In order to completely eliminate the felt pain, it is necessary to eliminate the factors that cause the pain.

For example, the best way to relieve a sore throat is to eliminate the bacteria that cause the pain with the use of antibiotics. In this case, the pain is also eliminated.

Analgesics are drugs used to tolerate pain. Simply called analgesic, pain reliever. An analgesic drug is any member of the group of drugs used to relieve pain. Local anesthetics interrupt pain transmission, general anesthetics usually relieve pain with inactivity and central nervous system depressants, such as central muscle relaxants, tranquilizers, and antihistamines, also relieve pain by an as yet unknown mechanism. Even placebo treatments have been shown to relieve pain. This explains the psychological effect [18].

3.2. Components of Drug Products

Although not required, the pharmaceutical product is called a drug product. Containing active drug component and excipients together with inactive ingredients; It is a drug product in tablet, capsule or solution, etc. forms.

The physicochemical and biological properties (eg, solid state properties) of the active substance that may affect the performance and manufacturability of the drug product or that are designed specifically for the drug substance should be identified and discussed [19].

Examples of physicochemical and biological properties that may need to be examined include solubility, water content, particle size, crystalline properties, biological activity, and permeability. These features may be interrelated and need to be considered together [20].

The efficacy and safety of a drug product is determined by its stability results. Although the most important compound affecting the stability of the drug product is the raw material, a very important part is the selection of the excipient.

Excipients are inert pharmaceutical ingredients used in product formulations. Excipients can play a variety of functional roles in the pharmaceutical product. Each excipient serves a specific purpose (eg, binder, disintegrant, or pH adjustment) for proper performance of the dosage form. The properties (ie stability) of the final dosage form depend largely on the excipients chosen, their concentrations, and their interactions both with the active compound and with each other.

Excipients selected, their concentrations, and properties that may affect drug product performance (eg, stability, bioavailability) or manufacturability should be discussed with respect to the respective function of each excipient. This should include all substances used in the manufacture of the drug product (eg, processing aids), whether or not they appear in the finished product [21].

The present invention also relates to methods for providing effective pain relief, including an analgesic effect, by administering the paracetamol formulation.

Paracetamol, also known as N-acetyl-p-aminophenol, acetaminophen, and APAP, is an analgesic and antipyretic agent and is widely used in prescription and over-the-counter drugs. Paracetamol was first marketed in the 1950s and is now a widely used agent.

Improving the rate and extent of absorption of oral formulations of compounds has been the subject of substantial research. In general, once a solid ingestion composition reaches the stomach, it undergoes degradation and/or disintegration and passes into the small intestine, where the active ingredient is absorbed into the circulatory system via the portal vein and the liver, before reaching its site of action.

Paracetamol oral formulation and excipient selection are mentioned in the patent numbered US8461210B2 of the United States, which is an exemplary formulation patent [22].

Paracetamol is a widely available analgesic drug product that has been used in most of the countries for more than 40 years. It has been produced by R&D and formulators as a result of many long and rich trials.

If a patient has mild to moderate pain and fever; It can be used as a standard antipyretic and analgesic. It is produced, sold and marketed in many countries over the counter and liquid as capsules, tablets and coated dosage forms. Following ingestion of paracetamol in solid form such as a tablet or capsule, the rate of drug absorption and onset of pharmacological activity may vary from patient to patient.

Dosage forms who are many formulated today are complex systems containing the active pharmaceutical ingredient (API) along with many other ingredients; these compounds are often added together with the active pharmaceutical ingredients to maintain, support or increase the stability of the formulation. It is often observed that the pharmaceutical active substance cannot maintain its stability in its pure form for a long time, as a result of which it denaturates or becomes unsuitable by adhering to the container wall, so excipients are added to stabilize the API. This helps to maintain the stability of the product and allows the API to maintain its stability for a significant period of time, thus extending the shelf life of the dosage formulation.

Pharmaceutical excipients help to increase the bioavailability of the active drug: Excipients are usually In most cases, an active ingredient (like aspirin) is not readily absorbed by the human body, in such cases the active ingredient is dissolved or mixed with an excipient that can act as a solvent or aid in absorption of the drug in the human body [23].

These ingredients are commonly referred to as excipients, and according to the International Pharmaceutical Excipient Council, Excipient is defined as "Any substance other than the active drug or prodrug included in the manufacturing process or contained in finished pharmaceutical dosage forms". US Pharmacopoeia-National formularies classify excipients according to the functions they perform in formulations; Connectors, Shredders, etc.

Excipients can be classified in the Table below according to their origin, their use in dosage form and the functions they perform. Figure 3.1. below shows some excipients. They are classified according to their origin, dosage form, usage areas and functional properties that they can perform.

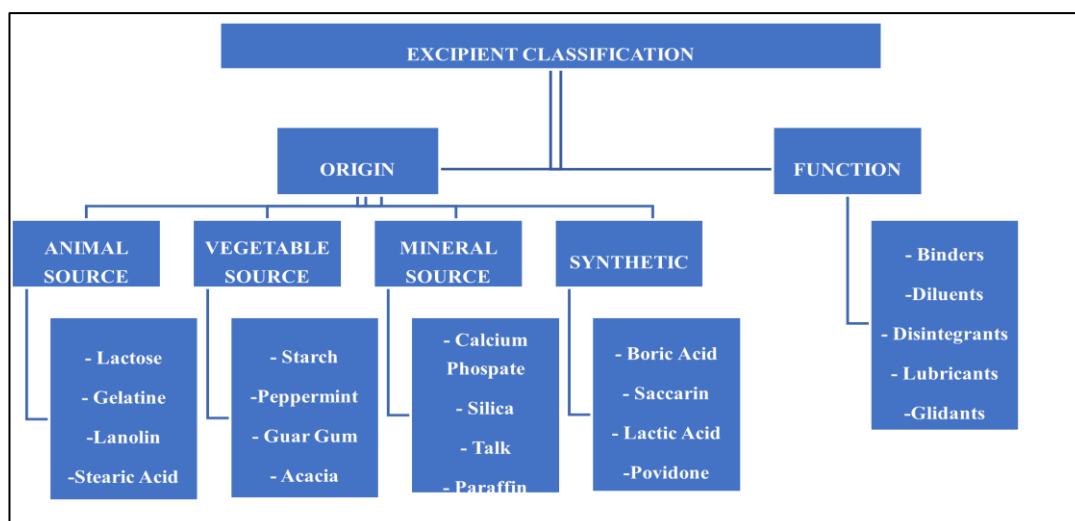


Figure 3.1: Excipient classification for selection.

3.3. Excipient Selection

Excipients are joined to a formulation to reach the desired filler gravity of a dosage form, evolve workability, or affect drug release behavior in the body.

They are materials that are considered indispensable components of medical drug products. They are present in more than the active ingredient in most pharmaceutical drug products. It is important that they are chosen correctly, as they always make up a large part of the formulation. Selection of auxiliary material; It is important in terms of legal acceptances, functionality, stability and effectiveness of the drug product throughout its shelf life. An irritating ideal defined as an excipient should have the following properties. An excipient should be: stable, behave chemically stable, not be reactive. On the contrary, it should be inert, have low equipment and process precision, durable, inert to the human body, non-toxic so that the employees and patients using it and the active substance are not affected by excipian, economical having efficiency in regards with the intended use [25].

Excipients are considered inert in each formulation. However, it should be known that drug components tend to interact with auxiliary materials and packaging materials. For example, if an excipient has the ability to absorb moisture and the active ingredient of the drug is affected by moisture; This situation may create some negativities throughout the stability of the drug product.

In the simplest case, an active ingredient affected by moisture may exhibit degradation as a result of incorrect packaging or excipient selection. Physically, it may deviate from the physical stability parameters outside the quality specifications determined during the stability process.

In short, excipients may contain various impurities that can cause the active pharmaceutical ingredients in the formulation to decompose, thereby altering the shelf life and stability of the formulation. In another way, by affecting the physical stability parameters; chemical stability parameters may not even need to be examined [26].

3.4. Stability of Drug Products

Stability assessment is a critical factor in pharmaceutical development to ensure the identity, potency, purity and quality of the product. Equally important, these studies need to be done in a way that is acceptable to regulatory bodies in a global setting. Product quality is an essential requirement for an effective and safe product and requires stability assessment from the point of compound selection during the preclinical phase.

Stability studies should be performed on the active substance packaged in a container closure system that is the same or simulates the packaging recommended for storage and distribution [5].

Stability studies should include testing for properties of the active substance that are subject to change during storage and are likely to affect quality, safety and/or efficacy. The test should cover physical, chemical, biological and microbiological properties, as appropriate. It is important that the active substance/auxiliary materials absorb moisture during this process, comply with the hardness limits, and that the dispersion/disintegration values of the drug product do not go beyond the limit. Analytical procedures demonstrating verified stability should be applied [27].

Acceptance criteria are numerical limits, ranges and other criteria for the specific tests described and should include individual and total upper limits for impurities and degradation products. As you can see in the Figure 3.2., the average kinetic temperature anywhere in the world can be derived from climate data, and the world can be divided into four climate zones I-IV in terms of long-term stability conditions along with national/regional requirements.

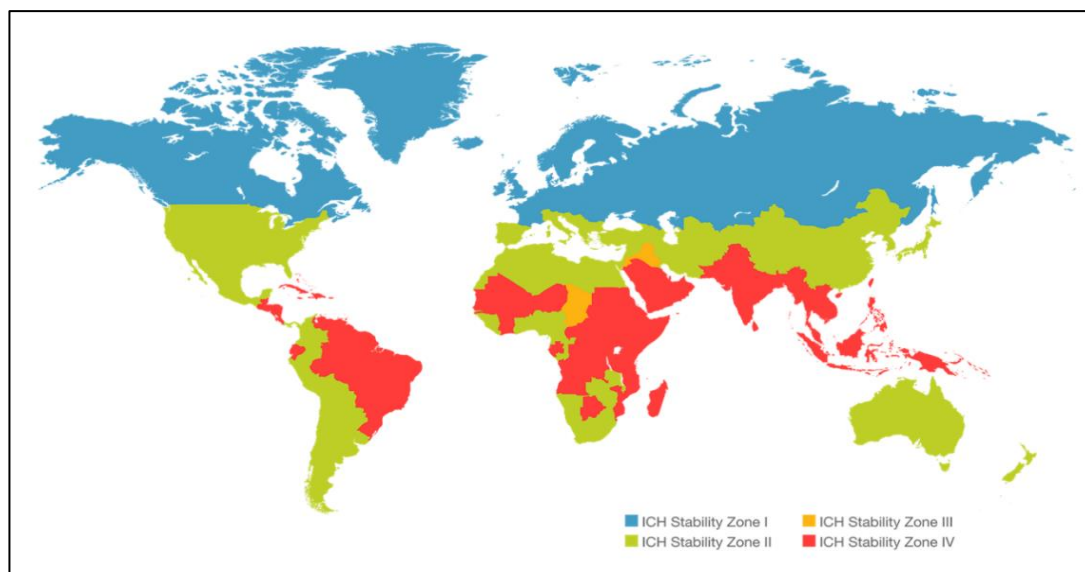


Figure 3.2: ICH stability climate conditions and zones in world.

ICH classifies the temperature and humidity conditions determined according to the regions it divides into four climate zones, I-IV in terms of long-term stability conditions, as given in the Table 3.1. below [29].

Table 3.1: Proposed conditions for long-term stability test by stability zones.

Climatic Zone	Type of Climate	Recommended Conditions
1 st Zone	Temperate	21 °C / 45 % RH
2 nd Zone	Mediterranean	25 °C / 60 % RH
3 rd Zone	Hot & dry	30 °C / 35 % RH
4 th Zone	Higher Humidity	30 °C / 65 % RH

Stability studies designed by the formulation developers according to this guide can be quite lengthy. An idea can be obtained by doing short-term stability studies beforehand. Or, when an out-of-limit situation is observed in drug product analyzes during this study, stability studies can be performed at short intervals.

3.5. Machine Learning Method for Drug Stability

In the last decade, With the rapid development of technology and especially in the last 10 years, learning with data has become very important in the pharmaceutical industry, as in every field of science.

Estimation and estimation problems often arise in learning. For these problems, regression theory based mainly on the idea of least squares or maximum likelihood estimation is used, but classification theory is also applied.

It can be said that such studies were first started twenty years ago. The first studies started with the application of artificial neural networks in the stability of pharmaceutical products. Since then, modeling methods have been used for issues such as drug release and the production of solid dosage forms. It is preferred in other new techniques, especially in artificial neural networks. To explain with an example, traditional stability methods start with a minimum of 3 months and the current stability process is determined by the shelf life. However, with a machine learning method that can predict beforehand; labor can be earned. This is because the traditional approach to drug formulation development relies on iterative trial and error, which requires a large number of resource-intensive and time-consuming in vitro and in vivo experiments.

Machine learning is a branch of artificial intelligence. It is a scientific discipline concerned with the design and development of algorithms that allow computers to develop behaviors based on experimental data [30].

Machines "learn" to automatically recognize complex patterns, distinguish samples based on their different patterns, and make intelligent decisions. Decision tree learning, association rule learning, neural networks, genetic programming, support vector machines, clustering, Bayesian networks etc. There are several machine learning disciplines such as Among these artificial neural networks are the most frequently applied in pharmaceutical development.

The most effective way to identify and control design spaces is to complement the parts to be identified and analyzed with some of the more advanced machine learning techniques such as ANN. Design of Experiments (DoE) leads to easy adjustment of level and bounds of experimental data attributes and process parameters that affect critical quality attributes (CQAs). However, the connection between input(s) and output(s) often requires advanced nonlinear modeling techniques compared to Design of Experiments approaches. DoE models are polynomial, whereas machine learning offers a wide variety of classification and modeling techniques [31].

The simplest ML model is perhaps linear regression, which assumes that predictive properties, such as API resolution, are linearly dependent on input

properties. Given its simplicity, this model is easy to interpret, but it also has serious limitations in modeling capabilities. Nonlinear models generally provide better performance.

ML models differ in their assumptions about the complexity of relationships between data, the process and features that compose it, and goals. In general, building an ML model requires a balance between predictive capacity, interpretability, and robustness. Some models can reproduce more complex relationships but tend to be more difficult to interpret, while others are easier to interpret but may be too simple to resolve all the variations in the underlying physical relationship [32].

The earliest applications of ML in drug product development date back to the 1990s, when ANNs were used to predict the properties of immediate-release (IR) oral tablets. These studies included the preparation and evaluation of a range of tablet formulations, and the data obtained was then used to supplement ANNs and/or decision trees to predict various outcomes (e.g. disintegration time, disintegration rate and friability).

Some of the first studies, and even the most popular, to show how ML can be used to predict drug formulation performance include a set of in-process control tablet specifications (i.e., tensile strength, brittleness, capping, disintegration time, and disintegration at various time points) for the drug direct printing process) is the use of ANNs to predict. The relative weight percentages of the excipients in the formulation are used as the input properties [33].

3.6. Decision Tree

Artificial neural networks (ANNs) have received increased interest in recent years due to their ability to represent relationships that are difficult to model and nonlinear. In addition, it has advantages such as ease of application, robustness and practicality of fast calculation.

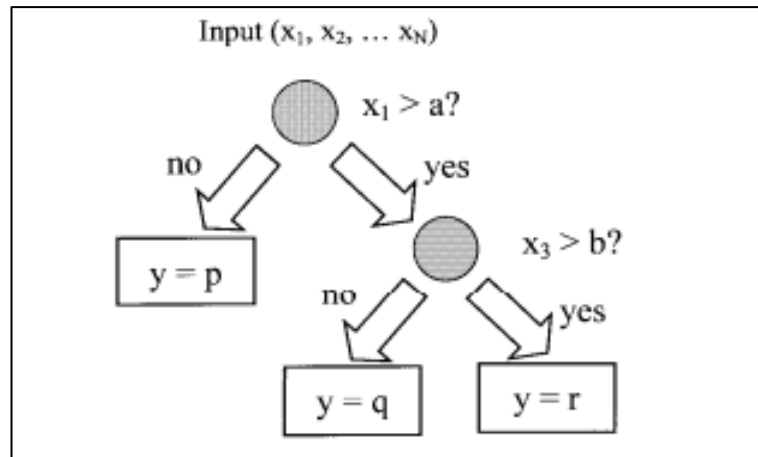


Figure 3.3: A representative binary decision tree: simple univariate.

The Decision Tree algorithm can be summarized as follows. In simplest terms, as you can see at Figure 3.3. above, a univariate (representative as in the figure) decision tree is produced with a trained neural network. Let us consider a particular decision node in the decision tree. The only variable specified in this node is "true" if it exceeds a trained threshold, otherwise "false" if it does not. For each node that can be added, the decision tree algorithm takes the dataset and decides how to classify its variable. It then sets a threshold for this variable [34].

Such decision trees can be used in the pharmaceutical industry to enable information gathering early in the drug development process. As a strategic tool, it allows the development of optimal analytical controls. It is the responsibility of the industry to determine the relevant pharmaceutical property of the drug product, to choose the appropriate method, to store it in the appropriate package, to select the appropriate excipient and to present it to the relevant international committees. With this approach, it is effective on a wide range and may be superior to performing long tests [35].

3.7. General Additive Model

The General Additive Model (GAM) is an algorithm system with regular scatter plots proposed by Hastie and Tibshirani (1987). The General Additive Model (GAM) is part of statistical learning and is a modern technique that functions in a wide sequence of fields such as finance, mathematics, biology, medicine, chemistry, and environmental engineering. In these models, response expectation is dependent on

predictors through a link function. GAMs are confident methods in order to low count time series like repeat admissions.

Within the framework of regression modeling techniques, the univariate General Additive Model is used quite frequently. Both models are modeled as a function of the explanatory variables of the mean μ of y . Given by $V(y) = \phi \quad v(\mu)$, it assumes that the variance of y depends on a constant and that there is an exponential function distribution for the response variable y . (distribution parameter ϕ and mean μ , variance function $v(\mu)$)

Thus, in GAM models, variance, skewness, and kurtosis are not explicitly modeled in terms of explanatory variables, but indirectly through their dependencies on μ [36].

3.8. Random Forest

A Random Forest is an unpruned collection of classification or regression trees created using bootstrap samples of training data and random feature selection in tree induction. Prediction is made by summing (major vote or averaging) the guesses of the community.

A new classification and regression tool, Random Forest, is introduced and investigated for predicting a compound's quantitative or categorical biological activity based on a quantitative description of the compound's molecular structure. Random Forest is an ensemble of unpruned classification or regression trees created by using bootstrap samples of the training data and random feature selection in tree induction. Prediction is made by aggregating (majority vote or averaging) the predictions of the ensemble. Predictive models were built for six cheminformatics data sets [37].

4. RESULTS AND DISCUSSION

4.1. Instruments

In this part of the study, moisture, disintegration and hardness analyzers that created the data set are given. These devices are devices that are calibrated every day within the body of X pharmaceutical company and keep up with the latest technology.

4.1.1. Moisture Content Analysis

The Sartorius device shown in Figure 4.1. was used as the device for moisture analysis. As seen in Figure 4.2., some powder is loaded into the device and alternative drying mode is selected. A Moisture Analyzer builds moisture analysis using the 'loss on drying' (LOD) way method. This is the most joint method for moisture analysis.



Figure 4.1: Device used for moisture content analysis.

After drying, which is one of the critical steps of the process, the sample that take at a certain weight is placed in the chamber or oven of the device, and it is weighed again by the device, followed by a pre-set program determined completely depending on the powder content. When the weight is fixed unchanged, the difference between the initial weight and the final weight refers to the evaporated moisture. This value is also used to calculate the moisture percentage.

The sample is weighed before and after dehumidification using the balance unit of the Moisture Analyzer as you can see at the Figure 4.2. below. This is done by continuously recording the weight as the sample is heated and dried using a halogen lamp. When the sample no longer loses weight, the difference between the two weights is measured, which is used to calculate the moisture content. This is also known as the thermo-gravimetric principle.



Figure 4.2: Moisture content analysis.

Here, the meaning of moisture determination does not only refer to the evaporable water in its content. The loss on drying method is also important in the determination of some essential oils and solvents such as alcohol used in the process.

4.1.2. Disintegration Analysis

The pharmaceutical product is solid dosage form and printed with different excipients at the desired fill weight. When this complex system comes into contact with physiological fluid, they are exposed to different mechanisms. The performance of my product is primarily affected by the disintegration, swelling and dissolution behavior of the powder compact it contains. The disintegration process is critical for immediate release dosage forms like the product. This is how estimate the release kinetics of the product.

The Disintegration Tester is a tool used to detect the dispersion of solid dosage products. In the analysis, the Distek device shown in Figure 4.3. was used as the device.



Figure 4.3: Device used for disintegration analysis.

As seen in Figure 4.4., when the water in the device reaches 37 °C, 6 tablets are loaded into the device reservoir and dispersion is observed in a disc manner. As soon as the product is completely dispersed, the stopwatch is checked. The dispersion time is determined. It is an easy-to-use and flexible distribution determination tester with reliable temperature control. It can also be used in industry to check the consistency and suitability of a batch of tablets or capsules. This also helps the production team adhere to established standards such as the USP requirement. Overall, the Disintegration Tester plays an important role in quality control and R&D.



Figure 4.4: Disintegration analysis.

The disintegration parameter refers to the disintegration (disintegration) of the compressed tablet into small granules after swallowing. After the tablets are compressed; It is characterized by the disintegration of the interparticle bonds formed.

4.1.3. Hardness Analysis

It is a laboratory technique used in the pharmaceutical industry to determine the breaking point and structural integrity of a tablet product and to learn how it change under "pre-use storage, transportation, packaging and handling conditions".

The device shown in Figure 4.5. was used as the device in the analysis. The operation of the majority of tablet hardness testing equipment is a basic action of a jaw or bar that presses the tablet between two plates. This forces the tablet to break under pressure thereby creating the necessary force.



Figure 4.5: Hardness analysis.

During the hardness test of the tablet, the tablet is placed between the two jaws. The motorized jaw orients the tablet forward, squeezing it against the stationary jaw until it cracks. After breaking, the motorized plate retracts and the device registers the force required to break the tablet.



Figure 4.6: Device used for hardness analysis.

Tablets should not be too hard or too soft. An excessively hard tablet may indicate the potential for excessive binding between the active ingredients and excipients, preventing proper dissolution of the tablet required for an accurate dosage.

4.2. Collection of Data

The data that collected for the study and the results of the analysis are given below. However, first of all, the values that see in the Table include some parameters that affect the analysis process and the stability of the drug product. During the process, kept the air flow and temperature parameters constant. In order to monitor the stability of the powder dried in the fluidized bed dryer at 3 different moisture levels, the drying times are variable in 3 trial productions, as given in the Table 4.1. below.

Table 4.1: Process parameters for three trial batches.

	Drying Time (sec)	Air Flow (m3/h)	Temperature (°C)
1st Trial Batch	79	10	60
2nd Trial Batch	58	10	60
3rd Trial Batch	41	10	60

Given in Table 4.2. are the following in-process control parameters. As can be seen, the moisture value decreases as the drying time increases. The data in Table 4.2. (in process control data) is the initial analysis.

Table 4.2: In process control analysis.

	Moisture (%)	Disintegration (sec)	Hardness (N)
1st Trial	0.689	118	68
2nd Trial	1.200	80	70
3rd Trial	2.002	58	63

4.2.1. First Trial Batch Stability Data

The first trial production, as explained and as you can see in Table 4.3. at below, is the results of short-term and frequent stability analyzes of the tablets obtained by pressing the dried powder mixture for 79 minutes with the appropriate parameters, packaging with two different packaging materials (PVC and PVC-PVDC)

Table 4.3: First trial batch stability analysis results.

1st Trial Batch Stability							
Granule Moisture		0.60 %					
Packaging		PVC - Al Foil			PVC/PVDC		
RESULTS							
25°C	INITIAL	0.689	118	68	0.689	118	68
	DAY 1	1.247	115	68	1.096	118	68
	DAY 2	1.306	115	65	1.100	115	68
	DAY 3	1.242	115	68	1.102	115	68
	DAY 7	1.205	115	66	0.977	113	67
	DAY 15	1.296	110	65	1.199	115	66
	1 MONTH	1.350	114	65	1.200	114	68
	2 MONTHS	1.470	114	66	1.206	116	68
	3 MONTHS	1.500	115	68	1.300	118	72
	4 MONTHS	1.502	116	68	1.292	118	71
60°C	INITIAL	1.247	118	68	1.200	118	68
	DAY 1	1.006	117	84	0.990	107	92
	DAY 2	0.928	119	94	1.000	110	95
	DAY 3	0.761	123	92	0.896	117	96
	DAY 7	0.643	123	94	0.707	115	100
	DAY 15	0.587	123	92	0.812	113	101
	1 MONTH	0.590	125	92	0.900	115	100
	2 MONTHS	0.500	125	93	0.881	110	101
	3 MONTHS	0.496	137	99	0.912	115	102
	4 MONTHS	0.502	143	95	1.000	100	106

4.2.2. Second Trial Batch Stability Data

The second trial production, as explained and as you can see in Table 4.4. at below, is the results of short-term and frequent stability analyzes of the tablets obtained by pressing the dried powder mixture for 58 minutes with the appropriate parameters, packaging with two different packaging materials.

Table 4.4: Second trial batch stability analysis results.

2nd Trial Batch Stability							
Moisture		1.20%					
Packaging		PVC – Al Foil			PVC/PVDC – Al Foil		
RESULTS							
		M	D	H	M	D	H
25 °C	INITIAL	1.200	80	70	1.200	80	70
	DAY 1	1.659	85	67	1.541	83	70
	DAY 2	1.572	90	72	1.556	93	72
	DAY 3	1.560	91	77	1.559	95	70
	DAY 7	1.574	92	72	1.519	99	75
	DAY 15	1.526	100	80	1.556	106	78
	1 MONTH	1.602	100	83	1.556	107	78
	2 MONTHS	1.605	100	83	1.600	107	79
	3 MONTHS	1.600	100	83	1.600	106	79
	4 MONTHS	1.600	102	84	1.600	100	81
60°C	INITIAL	1.650	80	70	1.444	80	70
	DAY 1	1.277	85	95	1.425	90	101
	DAY 2	0.914	85	99	1.380	87	100
	DAY 3	0.786	82	101	1.412	98	106
	DAY 7	0.707	85	100	0.870	85	111
	DAY 15	0.710	86	102	0.974	92	109
	1 MONTH	0.710	90	102	1.000	100	108
	2 MONTHS	0.705	90	103	1.200	100	111
	3 MONTHS	0.705	90	103	1.115	102	118
	4 MONTHS	0.802	105	117	1.122	103	115

4.2.3. Third Trial Batch Stability Data

The third trial production, as explained and as you can see in Table 4.5. at below, is the results of short-term and frequent stability analyzes of the tablets obtained by pressing the dried powder mixture for 41 minutes with the appropriate parameters, packaging with two different packaging materials.

Table 4.5: Third trial batch stability analysis result.

3rd Trial Batch Stability							
Granule Moisture		2.00%					
Packaging		PVC - Al Foil			PVC/PVDC - Al Foil		
RESULTS		<u>RESULTS</u>					
		M	D	H	M	D	H
25°C	INITIAL	2.002	58	63	2.002	58	63
	DAY 1	2.168	60	87	2.123	54	97
	DAY 2	2.085	75	101	2.200	72	100
	DAY 3	1.827	76	107	1.965	80	99
	DAY 7	1.923	80	109	1.977	81	110
	DAY 15	1.960	80	110	1.906	80	116
	1 MONTH	2.000	83	110	1.942	82	119
	2 MONTHS	2.040	83	111	1.962	86	120
	3 MONTHS	2.300	99	110	1.962	86	120
	4 MONTHS	2.306	102	112	1.988	90	123
60°C	INITIAL	2.002	58	63	2.002	58	63
	DAY 1	1.736	65	135	2.100	90	138
	DAY 2	1.119	69	135	2.080	95	135
	DAY 3	0.882	71	140	2.500	96	134
	DAY 7	0.813	82	141	2.489	99	146
	DAY 15	0.787	86	138	2.161	95	151
	1 MONTH	0.779	90	139	2.000	92	145
	2 MONTHS	0.780	92	140	2.459	110	142
	3 MONTHS	0.780	167	145	2.666	157	145
	4 MONTHS	0.824	174	152	2.812	165	157

4.3. Formulas

The analysis results data collected in this study and the data taught to the systems were calculated and interpreted with the help of the formulas in the table below. Some formulas that used in the study are given in Table 4.6 below.

Table 4.6: Formulas that used for study.

Name	Abbreviation	Formula ^(a)
Coefficient of Determination	R^2	$R^2 = 1 - \left[\frac{\sum_1^n (y_i - y_i^{pred})^2}{\sum_1^n (y_i - \bar{y})^2} \right]$
Adjusted Coefficient of Determination	Adj_R^2	$Adj.R^2 = 1 - \frac{[(1 - R^2) * (m - 1)]}{m - n - 1}$
Root Mean Squared Error	RMSE	$RMSE = \sqrt{\frac{\sum_1^n (y_i - y_i^{pred})^2}{n}}$
Normalized Root Mean Squared Error	NRMSE	$NRMSE = \sqrt{\frac{\sum_1^n (\frac{y_i - y_i^{pred}}{y_i})^2}{n}}$
Average Absolute Error	AAE	$AAE = \frac{1}{n} \sum_{i=1}^n X_i - X $

4.4. Results

4.4.1. Hardness Results

As a result of the hardness studies, coded the packaging materials in 2 different ways, as in the Table 4.7. below. (0: PVC packaging and 1: PVC/PVDC packaging) Then, the recorded stability data for the 1st, 3rd, 5th, 7th, 15th, 30th, 60th, 90th and 120th days.

Table 4.7: Data collection for hardness results.

Packaging	Moisture	Temperature	Initial	Day	Hardness
0.000	0.600	25.000	68	1	68
0.000	0.600	25.000	68	3	65
0.000	0.600	25.000	68	5	68
0.000	0.600	25.000	68	7	66
0.000	0.600	25.000	68	15	65
0.000	0.600	25.000	68	30	65
0.000	0.600	25.000	68	60	66
0.000	0.600	25.000	68	90	68
0.000	0.600	25.000	68	120	68
1.000	0.600	25.000	68	1	68
1.000	0.600	25.000	68	3	68
1.000	0.600	25.000	68	5	68
1.000	0.600	25.000	68	7	67
1.000	0.600	25.000	68	15	66
1.000	0.600	25.000	68	30	68
1.000	0.600	25.000	68	60	68
1.000	0.600	25.000	68	90	72
1.000	0.600	25.000	68	120	71
0.000	1.200	25.000	70	1	67
0.000	1.200	25.000	70	3	72
0.000	1.200	25.000	70	5	77
0.000	1.200	25.000	70	7	72
0.000	1.200	25.000	70	15	80
0.000	1.200	25.000	70	30	83
0.000	1.200	25.000	70	60	83
0.000	1.200	25.000	70	90	83
0.000	1.200	25.000	70	120	84
1.000	1.200	25.000	70	1	70

Table 4.7: Continued.

Packaging	Moisture	Temperature	Initial	Day	Hardness
1.000	1.200	25.000	70	3	72
1.000	1.200	25.000	70	5	70
1.000	1.200	25.000	70	7	75
1.000	1.200	25.000	70	15	78
1.000	1.200	25.000	70	30	78
1.000	1.200	25.000	70	60	79
1.000	1.200	25.000	70	90	79
1.000	1.200	25.000	70	120	81
0.000	2.000	25.000	63	1	87
0.000	2.000	25.000	63	3	101
0.000	2.000	25.000	63	5	107
0.000	2.000	25.000	63	7	109
0.000	2.000	25.000	63	15	110
0.000	2.000	25.000	63	30	110
0.000	2.000	25.000	63	60	111
0.000	2.000	25.000	63	90	110
0.000	2.000	25.000	63	120	112
1.000	2.000	25.000	63	1	97
1.000	2.000	25.000	63	3	100
1.000	2.000	25.000	63	5	99
1.000	2.000	25.000	63	7	110
1.000	2.000	25.000	63	15	116
1.000	2.000	25.000	63	30	119
1.000	2.000	25.000	63	60	120
1.000	2.000	25.000	63	90	120
1.000	2.000	25.000	63	120	123
0.000	0.600	60.000	68	1	84
0.000	0.600	60.000	68	3	94
0.000	0.600	60.000	68	5	92
0.000	0.600	60.000	68	7	94
0.000	0.600	60.000	68	15	92
0.000	0.600	60.000	68	30	92
0.000	0.600	60.000	68	60	93
0.000	0.600	60.000	68	90	99
0.000	0.600	60.000	68	120	95
1.000	0.600	60.000	68	1	92
1.000	0.600	60.000	68	3	95
1.000	0.600	60.000	68	5	96
1.000	0.600	60.000	68	7	100
1.000	0.600	60.000	68	15	101
1.000	0.600	60.000	68	30	100
1.000	0.600	60.000	68	60	101

Table 4.7: Continued.

Packaging	Moisture	Temperature	Initial	Day	Hardness
1.000	0.600	60.000	68	90	102
1.000	0.600	60.000	68	120	106
0.000	1.200	60.000	70	1	95
0.000	1.200	60.000	70	3	99
0.000	1.200	60.000	70	5	101
0.000	1.200	60.000	70	7	100
0.000	1.200	60.000	70	15	102
0.000	1.200	60.000	70	30	102
0.000	1.200	60.000	70	60	103
0.000	1.200	60.000	70	90	103
0.000	1.200	60.000	70	120	103
1.000	1.200	60.000	70	1	101
1.000	1.200	60.000	70	3	100
1.000	1.200	60.000	70	5	106
1.000	1.200	60.000	70	7	111
1.000	1.200	60.000	70	15	109
1.000	1.200	60.000	70	30	108
1.000	1.200	60.000	70	60	111
1.000	1.200	60.000	70	90	118
1.000	1.200	60.000	70	120	117
0.000	2.000	60.000	63	1	135
0.000	2.000	60.000	63	3	135
0.000	2.000	60.000	63	5	140
0.000	2.000	60.000	63	7	141
0.000	2.000	60.000	63	15	138
0.000	2.000	60.000	63	30	139
0.000	2.000	60.000	63	60	140
0.000	2.000	60.000	63	90	145
0.000	2.000	60.000	63	120	152
1.000	2.000	60.000	63	1	138
1.000	2.000	60.000	63	3	135
1.000	2.000	60.000	63	5	134
1.000	2.000	60.000	63	7	146
1.000	2.000	60.000	63	15	151
1.000	2.000	60.000	63	30	145
1.000	2.000	60.000	63	60	142
1.000	2.000	60.000	63	90	145
1.000	2.000	60.000	63	120	157

In drug stability, the parameters are highly dependent on each other. For example, as a result of the increase in humidity, the powders in the tablet adhere more tightly. Thus, the hardness value increases. Disintegration takes longer as the powders in the tablets with higher hardness values adhere to each other more tightly.

It may be easy to examine this situation in linear regression. For this reason, linear regression was examined as "preliminary information". The parameters obtained as a result of the regression analysis are given in the Table 4.8. at below.

Table 4.8: Regression analysis results.

Statistics of Regression	
R Multiple	0.98
R ²	0.97
R ² (Adjusted)	0.96
Standart Error	4.73
Observation	108

The results of the ANOVA analysis performed to analyze how the independent variables interact with each other and the effects of these interactions on the dependent variable are in Table 4.9. at below.

Table 4.9: ANOVA Analysis.

	df	SS	MS	F
Regression	5	15298.97	3059.795	69.32889
Difference	18	794.4206	44.13448	-
Total	23	16093.39	-	-

Although the accuracy of the results is good, it is not always suitable for application. In some cases, stability parameters may deviate for many unknown reasons (which should be investigated). It is essential to develop a system that allows to catch this deviation from the beginning.

As can be seen in the Table 4.10. below. three different machine learning methods were used. Decision Tree method. Random Forest method and General Additive Method were used. The results are as given below.

Table 4.10: Analysis hardness results.

Hardness					
	R^2	Adjusted R^2	RMSE	norm_RMSE	AAE
DT	0.971	0.962	4.509	0.042	0.029
GAM	0.941	0.921	6.244	0.064	0.047
RF	0.962	0.950	4.280	0.044	0.030

The results of the analyzes made with the Random Forest method are as in the Figure 4.7. and with the Decision Tree method in the Figure 4.8. and with the General Additive Method in the Figure 4.9.

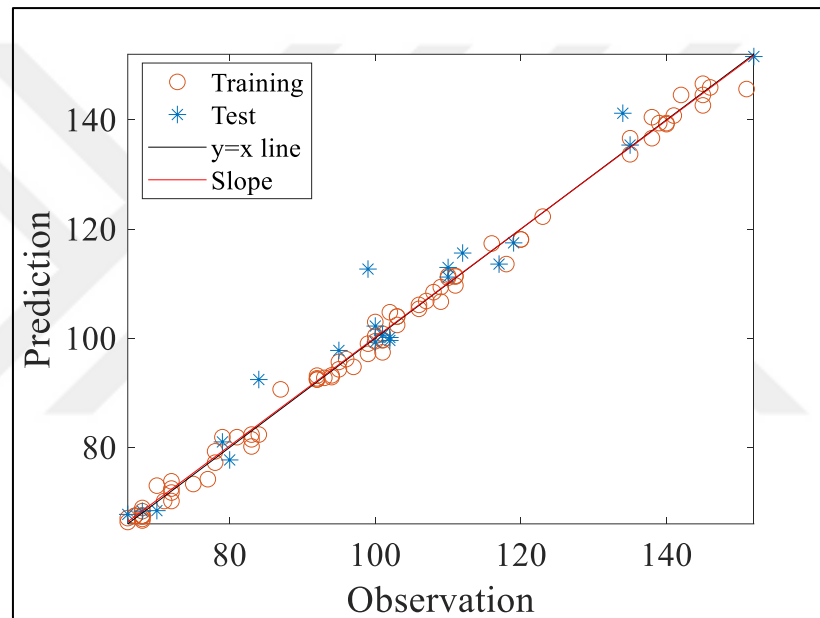


Figure 4.7: Hardness analysis results obtained using the random forest model.

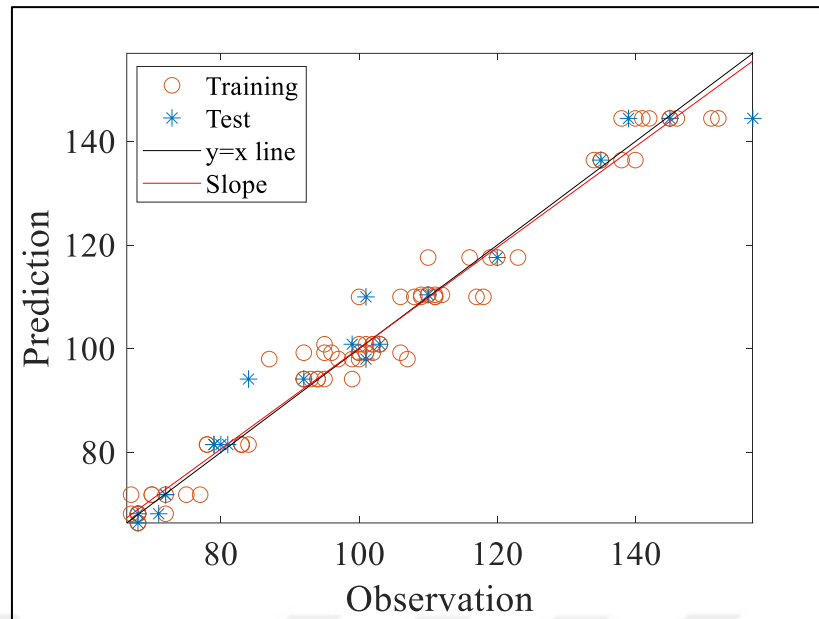


Figure 4.8: Hardness analysis results obtained using the decision tree model.

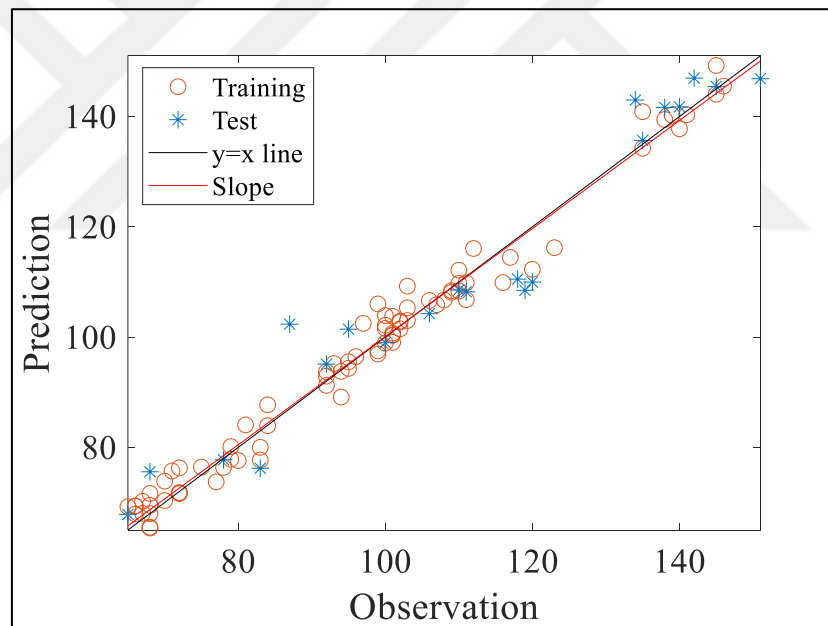


Figure 4.9: Hardness analysis results obtained using the general additive model.

4.4.2. Moisture Content Result

When the short-term stability results were examined, the following data were collected for the moisture content results.

As a result of the Moisture Content studies, coded the packaging materials in 2 different ways, as in the Table 4.11. below. (0: PVC packaging and 1: PVC/PVDC packaging) Then, the recorded stability data for the 1st, 3rd, 5th, 7th, 15th, 30th, 60th, 90th and 120th days.

Table 4.11: Data collection for moisture content result.

Packaging	Moisture	T	Initial	Day	Moisture
0.000	0.600	25.000	0.689	1	1.247
0.000	0.600	25.000	0.689	3	1.306
0.000	0.600	25.000	0.689	5	1.242
0.000	0.600	25.000	0.689	7	1.205
0.000	0.600	25.000	0.689	15	1.296
0.000	0.600	25.000	0.689	30	1.350
0.000	0.600	25.000	0.689	60	1.470
0.000	0.600	25.000	0.689	90	1.500
0.000	0.600	25.000	0.689	120	1.502
1.000	0.600	25.000	0.689	1	1.096
1.000	0.600	25.000	0.689	3	1.100
1.000	0.600	25.000	0.689	5	1.102
1.000	0.600	25.000	0.689	7	0.977
1.000	0.600	25.000	0.689	15	1.199
1.000	0.600	25.000	0.689	30	1.200
1.000	0.600	25.000	0.689	60	1.206
1.000	0.600	25.000	0.689	90	1.300
1.000	0.600	25.000	0.689	120	1.292
0.000	1.200	25.000	1.200	1	1.659
0.000	1.200	25.000	1.200	3	1.572
0.000	1.200	25.000	1.200	5	1.560
0.000	1.200	25.000	1.200	7	1.574
0.000	1.200	25.000	1.200	15	1.526
0.000	1.200	25.000	1.200	30	1.602
0.000	1.200	25.000	1.200	60	1.605
0.000	1.200	25.000	1.200	90	1.600
0.000	1.200	25.000	1.200	120	1.600
1.000	1.200	25.000	1.200	1	1.541
1.000	1.200	25.000	1.200	3	1.556

Table 4.11: Continued.

Packaging	Moisture	T	Initial	Day	Moisture
1.000	1.200	25.000	1.200	5	1.559
1.000	1.200	25.000	1.200	7	1.519
1.000	1.200	25.000	1.200	15	1.556
1.000	1.200	25.000	1.200	30	1.556
1.000	1.200	25.000	1.200	60	1.600
1.000	1.200	25.000	1.200	90	1.600
1.000	1.200	25.000	1.200	120	1.600
0.000	2.000	25.000	2.002	1	2.168
0.000	2.000	25.000	2.002	3	2.085
0.000	2.000	25.000	2.002	5	1.827
0.000	2.000	25.000	2.002	7	1.923
0.000	2.000	25.000	2.002	15	1.960
0.000	2.000	25.000	2.002	30	2.000
0.000	2.000	25.000	2.002	60	2.040
0.000	2.000	25.000	2.002	90	2.300
0.000	2.000	25.000	2.002	120	2.306
1.000	2.000	25.000	2.002	1	2.123
1.000	2.000	25.000	2.002	3	2.200
1.000	2.000	25.000	2.002	5	1.965
1.000	2.000	25.000	2.002	7	1.977
1.000	2.000	25.000	2.002	15	1.906
1.000	2.000	25.000	2.002	30	1.942
1.000	2.000	25.000	2.002	60	1.962
1.000	2.000	25.000	2.002	90	1.962
1.000	2.000	25.000	2.002	120	1.988
0.000	0.600	60.000	1.247	1	1.006
0.000	0.600	60.000	1.247	3	0.928
0.000	0.600	60.000	1.247	5	0.761
0.000	0.600	60.000	1.247	7	0.643
0.000	0.600	60.000	1.247	15	0.587
0.000	0.600	60.000	1.247	30	0.590
0.000	0.600	60.000	1.247	60	0.500
0.000	0.600	60.000	1.247	90	0.496
0.000	0.600	60.000	1.247	120	0.502
1.000	0.600	60.000	1.200	1	0.990
1.000	0.600	60.000	1.200	3	1.000
1.000	0.600	60.000	1.200	5	0.896
1.000	0.600	60.000	1.200	7	0.707
1.000	0.600	60.000	1.200	15	0.812
1.000	0.600	60.000	1.200	30	0.900
1.000	0.600	60.000	1.200	60	0.881
1.000	0.600	60.000	1.200	90	0.912

Table 4.11: Continued.

Packaging	Moisture	T	Initial	Day	Moisture
0.000	1.200	60.000	1.650	1	1.277
0.000	1.200	60.000	1.650	1	1.277
0.000	1.200	60.000	1.650	3	0.914
0.000	1.200	60.000	1.650	5	0.786
0.000	1.200	60.000	1.650	7	0.707
0.000	1.200	60.000	1.650	15	0.710
0.000	1.200	60.000	1.650	30	0.710
0.000	1.200	60.000	1.650	60	0.705
0.000	1.200	60.000	1.650	90	0.705
0.000	1.200	60.000	1.650	120	0.802
0.000	1.200	60.000	1.650	1	1.277
1.000	1.200	60.000	1.444	1	1.425
1.000	1.200	60.000	1.444	3	1.380
1.000	1.200	60.000	1.444	5	1.412
1.000	1.200	60.000	1.444	7	0.870
1.000	1.200	60.000	1.444	15	0.974
1.000	1.200	60.000	1.444	30	1.000
1.000	1.200	60.000	1.444	60	1.200
1.000	1.200	60.000	1.444	90	1.115
1.000	1.200	60.000	1.444	120	1.122
0.000	2.000	60.000	2.002	1	1.736
0.000	2.000	60.000	2.002	3	1.119
0.000	2.000	60.000	2.002	5	0.882
0.000	2.000	60.000	2.002	7	0.813
0.000	2.000	60.000	2.002	15	0.787
0.000	2.000	60.000	2.002	30	0.779
0.000	2.000	60.000	2.002	60	0.780
0.000	2.000	60.000	2.002	90	0.780
0.000	2.000	60.000	2.002	120	0.824
1.000	2.000	60.000	2.002	1	2.100
1.000	2.000	60.000	2.002	3	2.080
1.000	2.000	60.000	2.002	5	2.500
1.000	2.000	60.000	2.002	7	2.489
1.000	2.000	60.000	2.002	15	2.161
1.000	2.000	60.000	2.002	30	2.200
1.000	2.000	60.000	2.002	60	2.459
1.000	2.000	60.000	2.002	90	2.666
1.000	2.000	60.000	2.002	120	2.812

In drug stability, the parameters are highly dependent on each other. For example, as a result of the increase in humidity, the powders in the tablet adhere more tightly. Thus, the hardness value increases. Disintegration takes longer as the powders in the tablets with higher hardness values adhere to each other more tightly.

It may be easy to examine this situation in linear regression. For this reason, linear regression was examined as "preliminary information".

The parameters obtained as a result of the regression analysis are given in the Table 4.12. at below.

Table 4.12: Regression analysis results.

Statistics of Regression	
R Multiple	0.81
R ²	0.65
R ² (Adjusted)	0.55
Standart Error	0.39
Observation	24

The results of the ANOVA analysis performed to analyze how the independent variables interact with each other and the effects of these interactions on the dependent variable are in Table 4.13. at below.

Table 4.13: ANOVA results.

	df	SS	MS	F
Regression	5	5.10296	1.02059	6.7122
Difference	18	2.73691	0.15205	-
Total	23	7.83987	-	-

Although the accuracy of the results is good, it is not always suitable for application. In some cases, stability parameters may deviate for many unknown reasons (which should be investigated). As can be concluded from the table, the accuracy value is low. It is understood that the deviation in moisture content cannot be predicted very well by linear regression. It is essential to develop a system that allows to catch this deviation from the beginning.

As can be seen in the Table 4.14. below. three different machine learning methods were used. Decision Tree method. Random Forest method and General Additive Method were used. The results are as given below.

Table 4.14: Parameters of moisture content (kf) analysis results.

Moisture Content (KF)					
	R^2	Adjusted R^2	RMSE	norm_RMSE	AAE
DT	0.894	0.858	0.178	0.141	0.127
GAM	0.867	0.823	0.197	0.146	0.118
RF	0.944	0.926	0.146	0.077	0.064

The results of the analyzes made with the Random Forest method are as in the Figure 4.10.. and with the Decision Tree method in the Figure 4.11. and with the General Additive Method in the Figure 4.12.

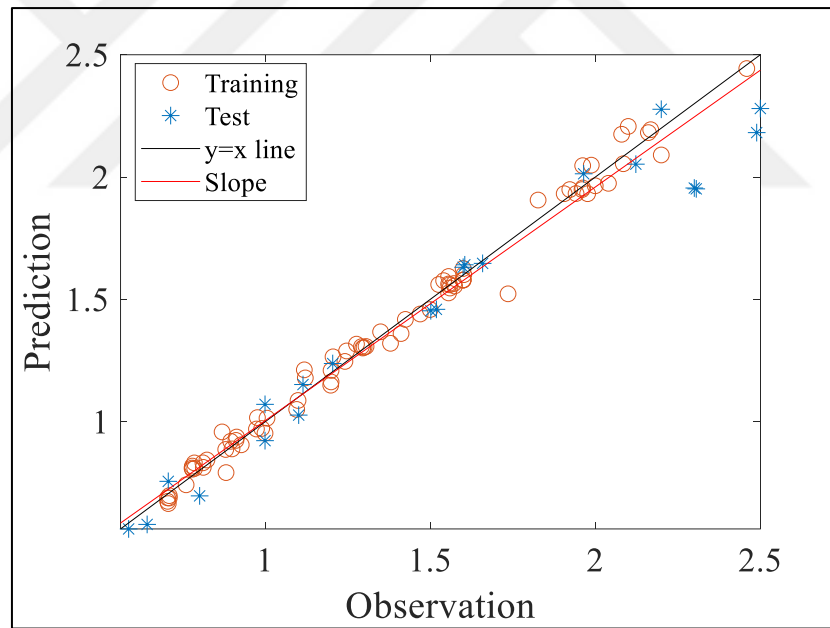


Figure 4.10: Moisture content (KF) analysis results obtained using the random forest model.

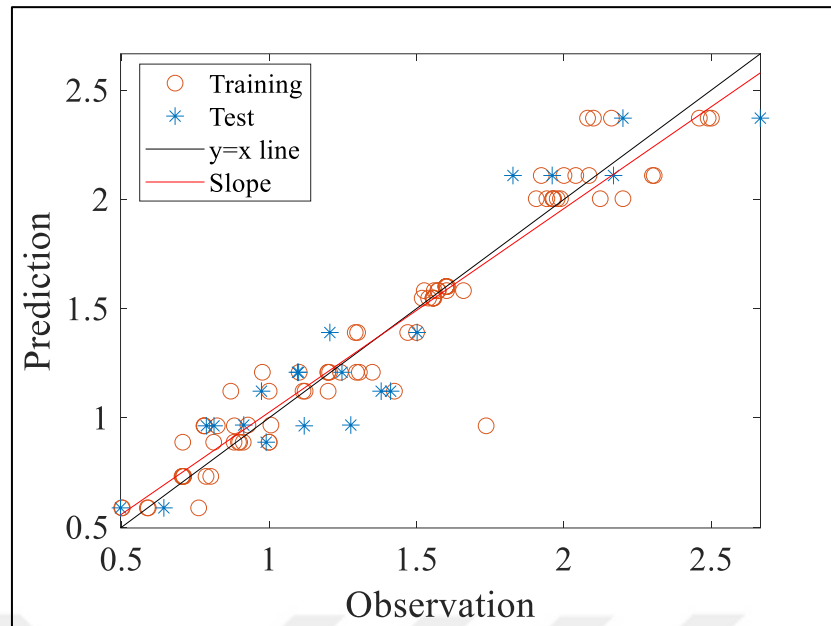


Figure 4.11: Moisture content (KF) analysis results obtained using the decision tree model.

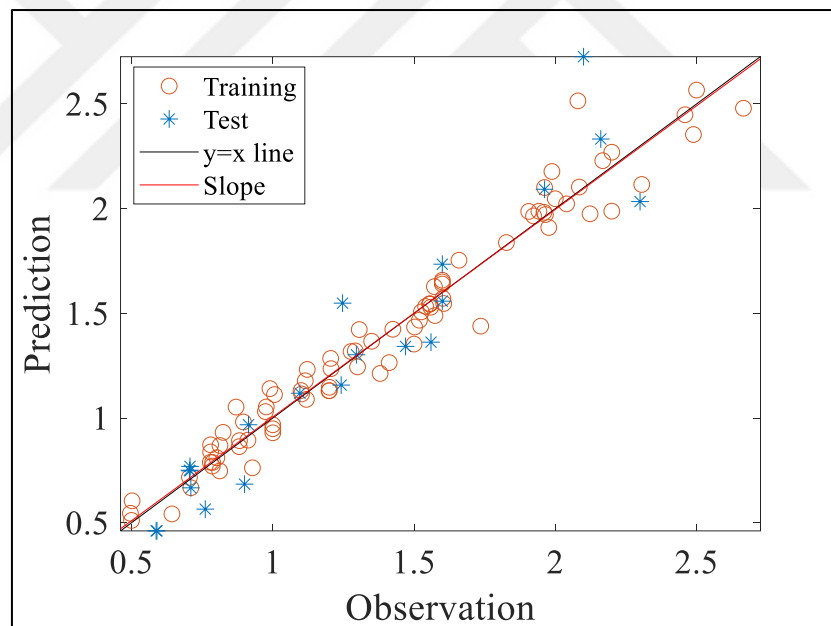


Figure 4.12: Moisture content (KF) analysis results obtained using the general additive model.

4.4.3. Disintegration Results

As a result of the disintegration studies. coded the packaging materials in 2 different ways. as in the Table 4.15. below. (0: PVC packaging and 1: PVC/PVDC packaging) Then, the recorded stability data for the 1st. 3rd. 5th. 7th. 15th. 30th. 60th. 90th and 120th days.

Table 4.15: Data collection for disintegration analysis.

Packaging	Moisture	T	Initial	Day	Disintegration
0.000	0.600	25.000	118	1	115
0.000	0.600	25.000	118	3	115
0.000	0.600	25.000	118	5	115
0.000	0.600	25.000	118	7	115
0.000	0.600	25.000	118	15	110
0.000	0.600	25.000	118	30	114
0.000	0.600	25.000	118	60	114
0.000	0.600	25.000	118	90	115
0.000	0.600	25.000	118	120	116
1.000	0.600	25.000	118	1	118
1.000	0.600	25.000	118	3	115
1.000	0.600	25.000	118	5	115
1.000	0.600	25.000	118	7	113
1.000	0.600	25.000	118	15	115
1.000	0.600	25.000	118	30	114
1.000	0.600	25.000	118	60	116
1.000	0.600	25.000	118	90	118
1.000	0.600	25.000	118	120	118
0.000	1.200	25.000	80	1	85
0.000	1.200	25.000	80	3	90
0.000	1.200	25.000	80	5	91
0.000	1.200	25.000	80	7	92
0.000	1.200	25.000	80	15	100
0.000	1.200	25.000	80	30	100
0.000	1.200	25.000	80	60	100
0.000	1.200	25.000	80	90	100
0.000	1.200	25.000	80	120	102
1.000	1.200	25.000	80	1	83
1.000	1.200	25.000	80	3	93
1.000	1.200	25.000	80	5	95
1.000	1.200	25.000	80	7	99

Table 4.15: Continued.

Packaging	Moisture	T	Initial	Day	Disintegration
1.000	1.200	25.000	80	15	106
1.000	1.200	25.000	80	30	107
1.000	1.200	25.000	80	60	107
1.000	1.200	25.000	80	90	106
1.000	1.200	25.000	80	120	100
0.000	2.000	25.000	58	1	60
0.000	2.000	25.000	58	3	75
0.000	2.000	25.000	58	5	76
0.000	2.000	25.000	58	7	80
0.000	2.000	25.000	58	15	80
0.000	2.000	25.000	58	30	83
0.000	2.000	25.000	58	60	83
0.000	2.000	25.000	58	90	99
0.000	2.000	25.000	58	120	102
1.000	2.000	25.000	58	1	54
1.000	2.000	25.000	58	3	72
1.000	2.000	25.000	58	5	80
1.000	2.000	25.000	58	7	81
1.000	2.000	25.000	58	15	80
1.000	2.000	25.000	58	30	82
1.000	2.000	25.000	58	60	86
1.000	2.000	25.000	58	90	86
1.000	2.000	25.000	58	120	90
0.000	0.600	60.000	118	1	117
0.000	0.600	60.000	118	3	119
0.000	0.600	60.000	118	5	123
0.000	0.600	60.000	118	7	123
0.000	0.600	60.000	118	15	123
0.000	0.600	60.000	118	30	125
0.000	0.600	60.000	118	60	125
0.000	0.600	60.000	118	90	137
0.000	0.600	60.000	118	120	143
1.000	0.600	60.000	118	1	107
1.000	0.600	60.000	118	3	110
1.000	0.600	60.000	118	5	117
1.000	0.600	60.000	118	7	115
1.000	0.600	60.000	118	15	113
1.000	0.600	60.000	118	30	115
1.000	0.600	60.000	118	60	110
1.000	0.600	60.000	118	90	115

Table 4.15: Continued.

Packaging	Moisture	T	Initial	Day	Disintegration
1.000	0.600	60.000	118	120	100
0.000	1.200	60.000	80	1	85
0.000	1.200	60.000	80	3	85
0.000	1.200	60.000	80	5	82
0.000	1.200	60.000	80	7	85
0.000	1.200	60.000	80	15	86
0.000	1.200	60.000	80	30	90
0.000	1.200	60.000	80	60	90
0.000	1.200	60.000	80	90	90
0.000	1.200	60.000	80	120	105
1.000	1.200	60.000	80	1	90
1.000	1.200	60.000	80	3	87
1.000	1.200	60.000	80	5	98
1.000	1.200	60.000	80	7	85
1.000	1.200	60.000	80	15	92
1.000	1.200	60.000	80	30	100
1.000	1.200	60.000	80	60	100
1.000	1.200	60.000	80	90	102
1.000	1.200	60.000	80	120	103
0.000	2.000	60.000	58	1	65
0.000	2.000	60.000	58	3	69
0.000	2.000	60.000	58	5	71
0.000	2.000	60.000	58	7	82
0.000	2.000	60.000	58	15	86
0.000	2.000	60.000	58	30	90
0.000	2.000	60.000	58	60	92
0.000	2.000	60.000	58	90	167
0.000	2.000	60.000	58	120	174
1.000	2.000	60.000	58	1	90
1.000	2.000	60.000	58	3	95
1.000	2.000	60.000	58	5	96
1.000	2.000	60.000	58	7	99
1.000	2.000	60.000	58	15	95
1.000	2.000	60.000	58	30	92
1.000	2.000	60.000	58	60	110
1.000	2.000	60.000	58	90	157
1.000	2.000	60.000	58	120	165

In drug stability, the parameters are highly dependent on each other. For example, as a result of the increase in humidity, the powders in the tablet adhere more tightly. Thus, the hardness value increases. Disintegration takes longer as the powders in the tablets with higher hardness values adhere to each other more tightly.

It may be easy to examine this situation in linear regression. For this reason, linear regression was examined as "preliminary information". The parameters obtained as a result of the regression analysis are given in the Table 4.16. at below.

Table 4.16: Regression analysis result.

Statistics of Regression	
R Multiple	0.77
R ²	0.60
R ² (Adjusted)	0.49
Standart Error	15.79
Observation	24

The results of the ANOVA analysis performed to analyze how the independent variables interact with each other and the effects of these interactions on the dependent variable are in Table 4.17. at below.

Table 4.17: ANOVA results

	df	SS	MS	F
Regression	5	6663.403	1332.681	5.348287
Difference	18	4485.222	249.179	-
Total	23	11148.63	-	-

Although the accuracy of the results is good, it is not always suitable for application. In some cases, stability parameters may deviate for many unknown reasons (which should be investigated). As can be concluded from the table, the accuracy value is low. It is understood that the deviation in moisture content cannot be predicted very well by linear regression.

As can be seen in the Table 4.18. below. three different machine learning methods were used. Decision Tree method. Random Forest method and General Additive Method were used. The results are as given below.

Table 4.18: Parameters of disintegration analysis results.

Disintegration					
	R^2	Adjusted R^2	RMSE	norm_RMSE	AAE
DT	0.239	-0.014	18.153	0.216	0.139
GAM	0.735	0.647	8.500	0.102	0.077
RF	0.806	0.741	9.376	0.081	0.060

The results of the analyzes made with the Random Forest method are as in the Figure 4.13.. and with the Decision Tree method in the Figure 4.14. and with the General Additive Method in the Figure 4.15.

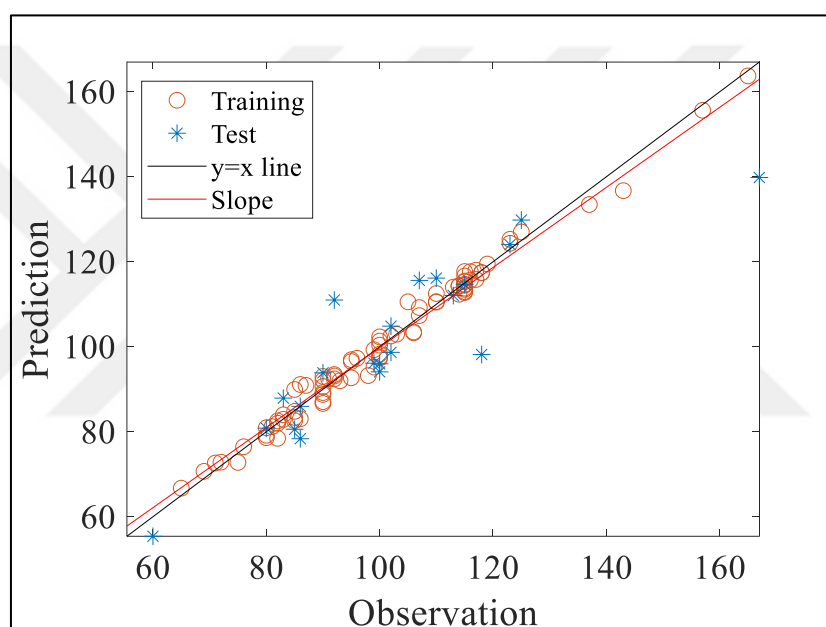


Figure 4.13: Disintegration analysis results obtained using the random forest model.

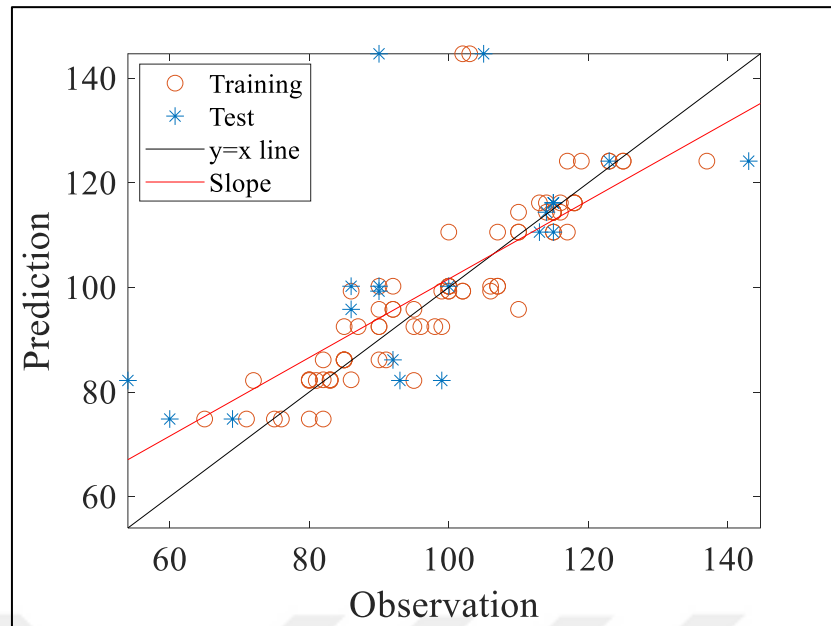


Figure 4.14: Disintegration analysis results obtained using the decision tree model.

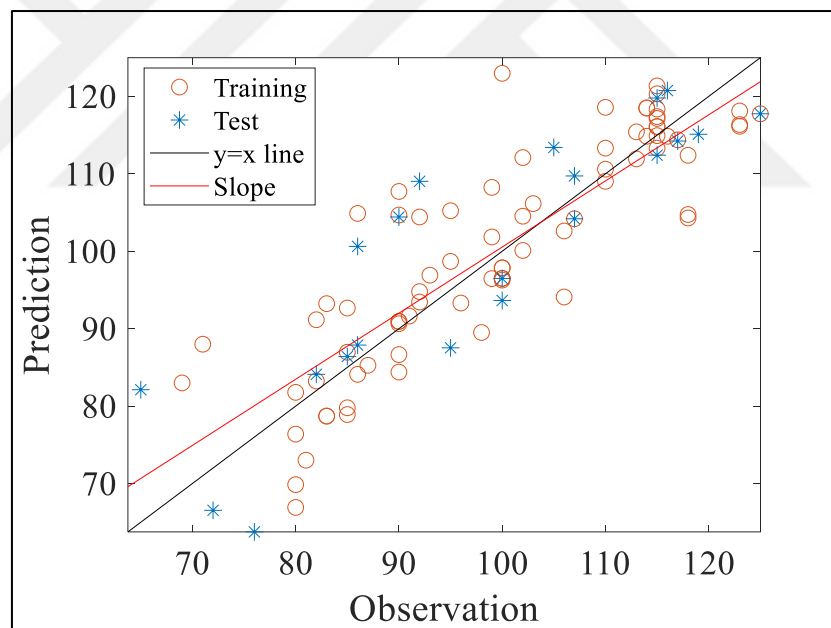


Figure 4.15: Disintegration analysis results obtained using the general additive model.

5. CONCLUSION

The most important sector in the world and in Türkiye is the Pharmaceutical Sector in terms of human health, reliable and continuous service. Scientific and theoretical knowledge is very important in the pharmaceutical industry. However, applied studies are also very important in developing new or generic products. While drug discovery and development is carried out in the pharmaceutical industry and R&D, first preliminary research and then development studies are carried out.

Some rules have been established due to the discovery of pharmaceutical products and some unfortunate and unfortunate accidents in the historical past. These rules have been transformed into laws or guidelines over time. Some harmonization organizations, ICH (International Council of Harmonization), are intended to achieve global harmonization.

It is important that an active drug substance is applicable in Türkiye and in the world. This is based on some guidelines and principles. Where the active ingredient is applicable, the finished drug product obtained and formulated must also be physically, chemically and microbiologically stable throughout its shelf life. Drug stability, which is one of the critical-importance quality and performance characteristics.

Drug stability starts with physical initial analyzes and stability analysis in the first stage. If the physical stability is appropriate and valid, it is decided to perform chemical stability analyzes in the next step. In the final stage, microbiological analysis is done. Therefore, physical stability, moisture and disintegration and hardness analysis values evaluated in operation are very important.

This is the case in every sector, however, since the pharmaceutical sector is a sector that serves humanity with respect, it is very important to obtain the most reliable and highest quality results in the least amount of time.

In order to serve this purpose or to open a way; A method has been chosen that can predict the results of the analysis during drug stability. Data sets are collected from the analysis results obtained. From this point of view, it is seen that the more data, the more accurate results obtained.

During this study, Decision Tree, Random Forest and General Additive Model methods from Machine Learning methods were used. Deviation plots and regression model fit parameters were examined. The analysis results of these models are summarized in the Table 5.1. below.

Table 5.1: Summary of regression and machine learning analysis results.

	Method	R²	RMSE	AAE
Hardness	Decision Tree	0.971	4.509	0.029
	General Additive Model	0.941	6.244	0.047
	Random Forest	0.962	4.280	0.030
Moisture Content	Decision Tree	0.894	0.178	0.127
	General Additive Model	0.867	0.197	0.118
	Random Forest	0.944	0.146	0.064
Disintegration	Decision Tree	0.239	18.153	0.139
	General Additive Model	0.735	8.500	0.077
	Random Forest	0.806	9.376	0.060

It is aimed to find the distance between the actual values and the estimated values and to choose the most accurate and reliable method. The more datasets, the more reliable the results.

When the Decision tree, General Additive Method and Random Forest methods are examined, it is seen that the model that gives the best similarity result for hardness analysis is Random Forest method. In addition, the error value calculated in the decision tree method is also low. (AAE) It is an important step in deciding whether to continue long-term stability. When Figure 4.6, Figure 4.7 and Figure 4.8 are examined, it is seen that the Random Forest method is appropriate when the deviations between the trained and tested parameters in the graph are controlled.

For moisture content analysis, the best similarity (R squared) between the data taught to the systems and the parameters estimated by the system was seen in the Random Forest model. When Figure 4.10, Figure 4.11 and Figure 4.12 are examined, it is seen that the Random Forest method is appropriate when the deviations between the trained and tested parameters in the graph are controlled.

For the disintegration analysis, the best similarity (R squared) between the data taught to the systems and the parameters estimated by the system was seen in the Random Forest model.

In addition, the calculated error value in this method is low. For the disintegration analysis before the long-term stability process, the way to save time and obtain effective safe analysis results may be to use the random forest method, one of the machine learning methods. When Figure 4.13, Figure 4.14 and Figure 4.15 are examined, it is seen that the Random Forest method is appropriate when the deviations between the trained and tested parameters in the graph are controlled.

Machine learning methods have been used in the pharmaceutical industry for more than 10 years. The frequent use of these methods in processes such as analysis, discovery and estimation of pharmaceutical products provide the industry with time, cost and labor savings. Since the estimations of these systems are made from the terms R squared and Absolute error and the values calculated between the parameters estimated by the system and trained the system, every data that obtain in the first stage allows us to grow the data set and get the most accurate result.

In the pharmaceutical industry; Increasing the frequency of use of these new and rational technological developments, which positively affect the most important terms such as time, labor and cost, and reducing the repetitions of analysis help patients to provide medicine, which is one of the most basic needs, in the most effective and rapid way.

REFERENCES

- [1] Sherer F. M., (2000), “The Pharmaceutical Industry”, In: A., Julyer, J., Newhouse, Editors, “Hanbook of Health Economics”, Elsevier
- [2] Web 1, (2022), [globenewswire.com](https://www.globenewswire.com), (Date of Access: 24/04/2022).
- [3] Web 2, (2022), <https://www.ema.europa.eu/en/partners>, (Date of Access: 24/04/2022).
- [4] Anonymous., (2007), “Quality assurance of pharmaceuticals”, “Good manufacturing practices and inspection”, WHO.
- [5] Marzena J., Piekowsk K., (2019), “Recent breakthroughs in the stability testing of pharmaceutical compounds”, TrAC Trends in Analytical Chemistry, 111, 118-127.
- [6] Davit B. M., Kanfer I., Tsang Y. C., (2016), “BCS biowaivers: Similarities and differences among EMA, FDA and WHO requirements”, The AAPS Journal, 18, 612-618.
- [7] Rhodes T. C., Dawoodbhai S., (2008), “The effect of moisture on powder flow and on compaction and physical stability of tablets”, Drug Development and Industrial Pharmacy, 15 (10), 1577-1600.
- [8] Banigan M., Aldegi P., Bao F., Häse A., Aspuru-Guzik C. A., (2021), “Machine learning directed drug formulation development”, Advanced Drug Delivery Reviews, 175, 113-836.
- [9] Web 3, (2022), <https://cen.acs.org/>, (Date of Access: 20/04/2022).
- [10] Kim H., J., Scialli R., A., (2011), “ Thalidomide: The tragedy of birth defects and the effective treatment of disease”, Toxicological Sciences, 122 (1), 1-6.
- [11] Web 4, (2022), www.who.int/medicines/publications, (Date of Access: 20/04/2022)

- [12] Wiggins J., M., Albanese A., J., (2019), "A Brief History of Pharmacopoeias: Global Perspective", In: S., Schniepp, Editor, "Regulatory Sourcebook", PharmTech.
- [13] Web 5, <https://www.fda.gov/drugs/guidance>, (Date of Access: 01/07/2022).
- [14] Web 6, <https://www.ema.europa.eu/en>, (Date of Access: 01/05/2023).
- [15] Valerie J., (2005), "Clinical drug trials - Studying the safety and efficacy of new pharmaceuticals", Master thesis, Geneva Law School.
- [16] Woolf C. J., Max M. B., (2001), "Mechanism-based Pain Diagnosis Issues for Analgesic Drug Development", Special Articles, 95, 215-239.
- [17] Pramod K., Tahir M. A., Charoo N. A., Ansari S. H., Ali J., (2016), "Pharmaceutical product development: A quality by design approach", International Journal of Pharmaceutical Investigation, 6 (3), 129-138.
- [18] Hijma H., J., Groeneveld J. G., (2021), "Analgesic drug development: proof-of-mechanism and proof-of-concept in early phase clinical studies", Medicine in Drug Discovery, 10, 100-183.
- [19] Rao V. A., Kim J. J., Patael D. S., Rajins K., Esutoll C. R., (2020), "A Comprehensive Scientific Survey of Excipients Used in Currently Marketed", Therapeutic Biological Drug Products, Pharmaceutical Research, 37 (10).
- [20] Allen V. L., (2008), "Dosage forms designs and developments", Clinical Therapeutics, 30 (11), 2102-2111.
- [21] Pehlivanov I., (2020), "Self-emulsifying drug delivery systems (sedds): excipient selection, formulation and characterization", Journal of IMAB Annual Proceeding, 26 (3), 3226-3233
- [22] Hemant H. A., (2013), "Fast release paracetamol tablets", GlaxoSmithKline Consumer Healthcare UK IP Ltd, US8461210B2.

- [23] Golightly L. K., Sumoliinske S. S., Bennett M. L., Iii E. W., Rumalock B. H., (1988), "Pharmaceutical excipients adverse effects associated with inactive ingredients in drug products", In Concepts in Toxicology Review Medical Toxicology, 3 (12), 36-58.
- [24] Singh V. K., Chaudhary A., (2016), "Ayurvedic natural excipients: an advance option for modern medicaments", International Journal of Pharmaceutical Sciences and Research, 7 (12), 4743-4755.
- [25] Pifferi G., Sanitorro P., Pedraani M., (1999), "Quality and functionality of excipients", Il Farmaco, 54 (1), 1-14.
- [26] Wu Y., Levons J., Narang A. S., Raghavan K., Rao V. M., (2011), "Reactive impurities in excipients: profiling, identification and mitigation of drug-excipient incompatibility", AAPS PharmSciTech, 12 (4), 5-24.
- [27] Grimm W., (1993), "Storage conditions for stability testing in the ec. japan and usa; the most important market for drug products", Drug Development and Industrial Pharmacy, 19 (20), 2795-2830.
- [28] Web 7, (2022), <https://q1scientific.com/ich-conditions-pharm/>, (Date of Access: 02/05/2022).
- [29] Vykontam U., Potnuri N., (2019), "Stability studies of pharmaceutical products", world journal of pharmaceutical research, 8 (8), 479-492.
- [30] Bannigan P., Aldoghi M., Bao Z., Häse F., (2021), "Machine learning directed drug formulation development", Advanced Drug Delivery Reviews, 175, 113-806.
- [31] Rantanen J., Khinast J., (2015), "The future of pharmaceutical manufacturing sciences", Journal of Pharm Sciences, 104 (11), 3612-3638.
- [32] Lavecchia A., (2015), "Machine-learning approaches in drug discovery: Methods and applications", In Drug Discovery Today, 20 (3), 318-331.

- [33] Sridhara R., Johnson J. R, Justice R., Keegan P., Chakravarty A., (2010), "Review of oncology and hematology drug product approvals at the us food and drug administration", *Journal of the National Cancer Institute*, 102 (4), 230-243.
- [34] Schmitz G. P., Aldrich C., Gouws F. S., (1999), "ANN-DT: An algorithm for extraction of decision trees from artificial neural networks", In *IEEE Transactions on Neural Networks*, 10 (6), 1-21.
- [35] Caramella C., Colombo P., Conte U., Ferrari F., Gazzaniga A., Lamanna A., Peppas N. A., (1988), "A physical analysis of the phenomenon of tablet disintegration", In *International Journal of Pharmaceutics*, 44 (3), 20-58.
- [36] Murase H., Nagashima H., Yonezaki S., (2009), "Application of a generalized additive model (GAM) to reveal relationships between environmental factors and distributions of pelagic fish and krill: a case study in Sendai Bay", *ICES Journal of Marine Science*, 66 (6), 1417-1424.
- [37] Svetnik V., Liaw A., Tong C., (2003), "Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling", *Journal of Chemical Information and Computer Sciences*, 43 (6), 1947-1958.

BIOGRAPHY

Esila DOLDUR successfully completed the Chemical Engineering Department of Kocaeli University, which she started in 2015, in 2019. She started her master's education at Gebze Technical University, Institute of Science, Department of Chemical Engineering in 2020.

She has been working as an assistant specialist in R&D Pharmaceutical Technology Development in a pharmaceutical company since 2020.



APPENDICES

Appendix A: Publications Within the Scope of Thesis Study

Yücel O., Doldur E., (2022), “Application of artificial neural networks in development of an analgesic solid dosage form and prediction of analysis results”, Ege 7th International Conference on Applied Sciences, 176, İzmir, Turkey, December 24-25.

