ilset
April 22-25
New York
USA
2021

Proceedings of International Conference on

Life Sciences, Engineering and Technology

April 22-25, 2021 New York, USA

istes Organization

Volume 1, Pages 1-49

Proceedings of International Conference on Life Sciences, Engineering and Technology

ISBN: 978-1-952092-19-0

Editors: Richard Thripp & Ismail Sahin

Articles: 1-6

Conference: International Conference on Life Sciences, Engineering and Technology (iLSET)

Dates: April 22-25, 2021

Location: New York, USA

Conference Chair(s):

Stephen Jackowicz, University of Bridgeport, United States

Richard Thripp, University of Central Florida, United States

The submissions are subject to a double-blind peer review process by at least two reviewers with expertise in the relevant subject area. The review policy is available at the conference web page: www.ilset.net

President

Stephen Jackowicz, University of Bridgeport, United States
Richard Thripp, University of Central Florida, United States

Scientific Board

Janice Fournillier, Georgia State University, United States
Wilfried Admiraal, Leiden University, Netherlands
Elizabeth (Betsy) Kersey, University of Northern Colorado, United States
Anastasios Theodoropoulos, University of Peloponnese, Greece
Arturo Tobias Calizon, University of Perpetual Help System Dalta, Philippines
Brett Buttliere, Technical University Dresden, Germany
Cara Williams, Emirates College For Advanced Education, United Arab Emirates
Chandra Pratama Syaima, University of Lampung, Indonesia
Chris Plyley, University of the Virgin Islands, Virgin Islands
Claudiu Mereuta, Dunarea De Jos University of Galati, Romania
Dana Aizenkot, Ashkelon Academic College, Israel
El Takach Suzanne, Lebanese University, Lebanon
Farouk Bouhadiba, University of Oran 2, Algeria
Frank Angelo Pacala, Samar State University, Philippines
Hou-Chang Chiu, Fu-Jen Catholic University Hospital, Taiwan
Irena Markovska, Assen Zlatarov University, Bulgaria
Irina Andreeva, Peter The Great St. Petersburg Polytechnic University (SPBPU), Russia
Iwona Bodys-Cupak, Jagiellonian University, Poland
Jaya Bishnu Pradhan, Tribhuvan University, Nepal
Jean-Yves Gantois, ICHEC, Belgium
Kassa Mickael, Addis Ababa University, Ethiopia
Kemmanat Mingsiritham, Sukhothai Thammathirat Open University, Thailand
Kristyna Balatova, University of Zilina, Slovakia
Milan Kubiatko, Jan Evangelista Purkyně University, Czech Republic
Neide Da Fonseca Parracho Sant'anna, Colegio Pedro II, Brazil
Oguz Akturk, Necmettin Erbakan University, Turkey
Ossi Autio, Univertsity of Helsinki, Finland
Philomina Ifeanyi Onwuka, Delta State University, Nigeria
Sharif Abu Karsh, Arab American University, Palestine
Shenglei Pi, Guangzhou University, China
Siew Nyet Moi, Universiti Malaysia Sabah, Malaysia
Sindorela Doli Kryeziu, University of Gjakova, Albania
Siti Sarawati Johar, Universiti Tun Hussein Onn Malaysia, Malaysia
Sodangi Umar, Federal University Gusau, Nigeria
Tayfur Ozturk, Necmettin Erbakan University, Turkey
Theodore Chadjipadelis, Aristotle University of Thessaloniki, Greece
Tryfon Mavropalias, University of Western Macedonia, Greece
Volodymyr Sulyma, Dnipropetrovsk Medical Academy, Ukraine

Organizing Committee

Janice Fournillier, Georgia State University, United States
Wilfried Admiraal, Leiden University, Netherlands
Elizabeth (Betsy) Kersey, University of Northern Colorado, United States
Aehsan Haj Yahya, Beit-Berl College, Israel
Alaa AlDahdouh, University of Minho, Portugal
Augusto Z. Macalalag, Arcadia University, United States
Bhesh Mainali, Rider University, United States
Janez Jamsek, University of Ljubljana, Slovenia
Josiah Zachary Nyangau, Louisiana State University, United States
Kent Löfgren, Umeå University, Sweden
Laurie Murphy, Saint Joseph's College, United States
Marelbi Olmos Perez, Universidad Tecnológica de Bolívar, Colombia
Masood Badri, UAE University, United Arab Emirates
Monica Reichenberg, University of Gothenburg, Sweden
Phu Vu, The University of Nebraska at Kearney, United States
Qian Wang, Manhattan College, United States
Rachid Ait Maalem Lahcen, University of Central Florida, United States
Wei Zakharov, Purdue University, United States
Zhanat Alma Burch, Duke University, United States

Table of Contents

Integrating Common Data Analytics Tools into Non-Technical Undergraduate Curricula

Kurt Kirstein, EdD
Central Washington University, USA

Abstract: Aside from statistics courses, accessible data analytics skills are often excluded from traditional non-technical university programs. These are topics that are typically the domain of programs that focus on math, statistics and computer science. Yet the need for these skills in non-technical disciplines is changing. A rapid expansion of data-related processes in organizations of many types requires individuals who have at least a working knowledge of common analytic tools. This article briefly describes three categories of data analytics tools that can be useful for graduates in any discipline. The first category covers descriptive tools that allow students to learn what is in a data set and what meaning can be made of it. The second category of tools teaches students how to predict likely outcomes based on relationships in past data. The final category introduces students to tools that allow them to segment data into useful clusters and classes and to build meaningful associations within the data.

Keywords: Data analytics, Non-technical disciplines, University curriculum, Common data tools

Introduction

In recent years, data has had an ever-increasing impact on many aspects of organizations. Mayer-Schonberger and Cukier (2013) point out how companies are often regarding their data as a valuable commodity that can drive strategic and economic decisions. Data provides companies with opportunities to personalize their customer interactions, forecast product demand, and predict market behavior (Rosidi, 2019). And it is not just large companies with "big data" that are taking advantage of this trend. Great insights are possible even in organizations where the data sets are limited in size.

As with many industry shifts, especially those of a technical nature, the skills of the labor force can lag behind the demand for them. This has been the case with data skills (Evans, 2019; Rosidi, 2019). Much of the expertise for handling data has been in the disciplines of statistics and computer science, which are contributing to new domains such as data science or machine learning (Taddy, 2019). However, centralized expertise limits organizations who are dependent on a small number of individuals to fully recognize the value that can be derived from data. It is especially unfortunate when the data knowledge required is readily accessible to

Kirstein, K. (2021). Integrating common data analytics tools into non-technical undergraduate curricula. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 1-10), New York, USA. ISTES Organization.

organizational leaders who do not need to be skilled in statistics and computer science to derive value from their data. Too often, they just lack the training.

As the expansion of data continues, the need for data skills in non-technical disciplines is changing. A rapid expansion of data-related processes in organizations of many types requires individuals who have at least a working knowledge of common analytic tools. A collection of such tools is beginning to find its way into the curricular plans of certain disciplines. These tools are practical and useful to help leaders solve common organizational problems or capitalize on strategic opportunities and they don't require advanced technical or statistical degrees.

This remainder of this article briefly describes three categories of data analytics tools that can be useful for graduates in any discipline. The first category covers descriptive tools that allow students to learn what is in a data set and what meaning can be made of it. The second category of tools teaches students how to predict likely outcomes based on relationships in past data. The final category introduces students to tools that allow them to segment data into useful clusters and classes and to build meaningful associations within the data. Together these make up a analytics skill set that can be infused into the curriculum of university programs that traditionally have not included any focus on data.

The De-mystification of Data Analysis

For many companies, the approach to data analysis and similar skills has been to focus them largely within technical units. They have been seen as skills that reside outside of traditional non-technical disciplines. To make use of data requires the assistance of "data people" who can bring the expertise necessary to turn raw data into something meaningful. This traditional approach has been known to create the false impression that even foundational data skills are inaccessible and are best left to the technical "data people".

This is now changing and that change is largely being driven by an expansion in the availability of data. Some disciplines (e.g. business programs) are beginning to recognize the advantage of including data skills in their curriculum but the need is outpacing practice. More university programs need to adopt the inclusion of data skills. In a practical sense, there are a number of data analysis tools, not traditionally built into university-level programs, that could be helpful in addressing common scenarios that future graduates are likely to face. Additionally, it may be best if these skills are not only an elective part of a non-technical program but are tightly infused into several of the core courses that make up those programs.

Any data analysis instruction, built into university programs, should de-mystify these tools and underscore their applicability to common operational problems. When students begin to understand these tools, they can build their own data analysis skill sets leading them to higher levels of productivity and efficacy in their careers. The remainder of this article will largely reference business students and their professional concerns but this is just one example of a non-technical discipline where data skills can have a significant impact on the employability

of graduates.

A Data Analysis Skill Set

To successfully integrate data analysis into non-technical education requires a re-framing of students' understanding of data analysis. Sometimes, students approach data instruction with trepidation because they believe themselves to be "not good at math or statistics" and they are hesitant to try. In other cases, graduates have had little experience with data because their programs did not include this type of content into the core instruction of their discipline. However, non-technical students can be successful in the use of these tools when they are explained in a real-world context and used in authentic instructional scenarios. Students can learn these skills through simulated, real-world practice as they begin to build their own data analysis skill sets.

A student's analytic skill set can consist of three parts: data discovery, prediction, and segmentation. In data discovery, students learn to visualize their data, clean up any problems such as missing or flawed data, and reduce its dimensionality to narrow their focus. In prediction, students learn to use past data sets to make predictions about the future. These tools include forecasting, simple and multiple regression, and logistic regression. In segmentation, students learn to understand the relationships and associations of different parts of their data, helping them to focus their analysis on specific groups or scenarios. Such tools include clustering, classification, and association rules. In the sections that follow, each of these tools and their applicability to business professions will be discussed in more detail.

Discovering the Data

Data can come from a number of sources. Regardless of how the data are accessed, the first step in data analysis is to discover what is in the data set. This can help inform later decisions about the most appropriate direction for analysis. In some cases, part of data discovery can be duplicitous because the contents of the data file are already known, or they match the fields that were asked for when the output was requested. In many cases, however, the data analyst must take time to learn what they have. Even when the fields that make up the data file are known, there three important steps to the data discovery process: visualization, data cleansing and data reduction.

Visualization

Visualization can be the first stage of the data analysis process. It serves many purposes, but its main use is as an introduction to the scope and nature of the data. Visualization can be conducted using numeric summarizations or graphical tools. The goal is the same which is to explore data and provide an effective way to present results.

Visualization techniques are primarily used in the preprocessing portion of the data analysis process (Shmueli et al., 2018). They can help us identify clear errors in the data (e.g. customers whose age is 999), replace missing values, remove duplicate rows, and deal with other formatting or content errors that must be corrected or accounted for. Beyond ensuring a clean data set, data visualization techniques support freeform exploration for the purpose of understanding the data structure, identifying interesting patterns, and generating novel questions (Shmuli et al., 2018).

Cleaning the Data

Closely aligned with data visualization is the process of data cleansing. Two key problems with many datasets are missing data and outliers, although data duplication can be an issue as well. At this stage, the questions to be asked are: How should missing data to be handled? Do the ranges of data points make sense or are there obvious outliers? If outliers exist, what strategy should be employed to handle them? Did all data fields convert correctly (e.g. dates and units of measure)? Are there any sections of the dataset that have been duplicated?

Missing data can be a frequent problem in datasets. Should a dataset have empty fields, different applications will handle them in different ways and some data analysis tools will simply fail to run until the missing data problem is resolved. Strategies for handling missing data include row removal or value replacement (with column mean). Future data analysis needs can often drive decisions regarding the handling of missing data.

In general, there are two types of outliers. The first is obvious errors in the data. The second are legitimate but extreme values that fall well outside the expected ranges for variables. How each type is handled is up to the analyst. Removal of outliers that are clearly errors can be accomplished by deleting the impacted rows entirely or by replacing erroneous outliers with the column mean that excludes the outlier values. Extreme but legitimate values require more consideration of both their importance and their impact on the analysis. It may be best to remove extreme values that do not impact later analysis. In other cases, the inclusion of extreme values is necessary and the analyst will need to anticipate the impact on results.

Two other challenges that may be encountered at the data cleansing stage are field formatting and data duplication. It is not uncommon that numeric fields, such as currency or dates, lose their proper formats when transferred from one platform to another. This is generally easily fixed by specifying the format of the fields in a separate step. Data duplication can be found using sorting, visualization or other tools that show the number of times that values of a key variable appear in a data set. Erroneously duplicated data rows, once identified, can simply be removed.

Data Reduction

Once the data are visualized and cleaned, and the analyst is familiar with the contents of the data set, more informed decisions can be made about the how the data set will be analyzed. This may lead to data reduction or

the removal of fields that will not be used in later analysis. Unneeded columns can simply be deleted from the data set (with proper backups made) but this process should be done judiciously so that further reviews, suggested by the initial analysis, are not hampered or rendered impossible because of missing data.

In some cases, analysis of the data is simplified by removing rows of unneeded records. Regardless of method, the goal of data reduction is to simplify the dataset and the analysis. It is largely driven by the questions that the analysis of the dataset is intended to answer. It allows the analyst to focus energy and effort only on those parts of the dataset that will contribute to the analysis without requiring extra time, storage, or effort.

Learn from the Past

There is a great deal that companies can learn from old data. Information about products that customers bought, or did not buy, is often captured in datasets and this useful in determining product design, feature enhancement, and marketing strategies. Past data can tell analytsts how much customers are willing to spend, which customer segments they should be focused on, information about customer retention or attrition, the likelihood that customers will be interested in our products and how many of those products will need to be produced to precisely meet customer demand (Siegel, 2016).

Included in an analyst's data tool kit should be methods of using past data to predict future customer behavior. At a minimum, analysts should be familiar with forecasting, linear regression, and logistic regression. Each of these are described further in the sections that follow.

Forecasting and Time Series Analysis

Forecasting future customer behavior can be a key part of an organization's success. It can be one of the most important methods of ensuring that the right number of products are available when customers want them. Underproduction can lose customers who go to competitors when products are not available, and overproduction leads to waste. Product forecasts must be as accurate as possible and a key method for forecasting future sales is to include information about past sales.

Time series analysis uses past data to predict future performance by tracking key variables over intervals (months, quarters) to determine how those values have changed over time. When using a time series for forecasting, both the trend and the seasonality of past data must be considered. Line charts and other visualization tools can help the analyst see such patterns. Through this process, it is easy to determine if there an up or downward trend or if the data is flat and consistent with limited variability. It is also possible to detect patterns in the data that suggest seasonality where there are consistent and predictable peaks and valleys in the values. These questions become an important part of the strategy that can be used to determine future forecasts.

Linear Regression

Another tool that can be used for prediction and forecasting is linear regression. With regression, an analyst may be attempting to predict numeric values of an output variable by considering relationships in past data. Regression uses two sets of variables: predictor and output. In more complex models, there can be multiple predictor variables used to predict the output.

The process of building a regression model is to look at how the predictors were related to the output variable in past data. From that analysis, the model provides coefficients that define the relationship between the predictor and output variables. These coefficients can then be applied to scenarios were predictor variables are known but the output values are not. By doing so, predicted output values can be produced. The accuracy of the regression model is highly dependent on the past relationship between the predictor and output variables and on the volume of data that is used to "train" or build the model.

Logistic Regression

While we may be looking for predicted numeric values in linear regression, logistic regression allows the prediction of likelihood or probability. Similar to linear regression, the relationship between predictor and output variables is determined. Also, similar to linear regression, there can be single or multiple predictors. But with logistic regression the output is converted to a probability, ranging from 0 to 1. Logistic regression helps determine the likelihood of customer behavior by looking at the relationship between predictor variables and customer behavior in past data. A common use of logistic regression is in loan approval decisions. Predictors such as credit score and income can be used to determine the likelihood of loan repayment which drives the loan approval decision.

Data Segmentation

For a variety of purposes, many companies find it useful to segment their customers into groups based on patterns of behavior (Linhof & Berry, 2011). Customers are separated based on their purchasing patterns, income level, frequency, or a combination of similar fields. Each customer is assigned to a group according to which categories or combinations of categories they match best. Then organizations can mount targeted campaigns leading to improved returns and expanded customer relationships. The expectation is that by focusing on groups of similar customers, these efforts will be more effective than a one-size-fits-all approach.

Applying segmentation to large customer transaction databases helps companies to understand the behavior of their different customer groups including what they are buying and how much they are spending. Each segmented group of transactions becomes its own cluster. Once the transactions are segmented into clusters,

association rules can be built on each separate cluster further personalizing the marketing approach that a company can take.

The following sections will cover three tools that can be used for segmenting or associating data: clustering, classification and association rules. These are useful tools for a number of scenarios where there is benefit to understand grouping and associations that are found in a data set. All three should be part of the data analyst's skill set.

Cluster Analysis

In simple terms, clustering allows a large dataset to be broken into smaller sets based on the values of selected variables. These variables could be customer demographics, purchase frequency, or product selection. The goal is to identify subsets of customers who can be associated with each other because they have similar patterns across selected variables. Then companies can reach out to those subsets using customized approaches.

There are two common methods for clustering a large dataset. The first divides the data into a pre-determined number of subsets based on a geometric determination of distance. K-means clustering is one common method of dividing data into groups. The newly clustered records can be reviewed in both graphical and tabular formats making it easy to see which cluster each record has been assigned to. One downside of the k-means clustering process is that it forces all records to be in a limited number of k clusters and that can be complicated by outliers.

The second form of clustering is agglomerative or hierarchical clusters. Under the agglomerative clustering process, all records start out as their own cluster. Next, the two records that are most similar to each other, across a selected set of variables, are joined. The process of joining similar records continues until every record is part of a single cluster. The value of this strategy is that the entire process of joining records is captured and represented in a graphical tree structure called a dendrogram (Linhof & Berry, 2011). Similar records are on the same branches and the lengths of the branches indicate the level of similarity.

Both clustering tools help to separate customers or transaction records into smaller sections, each of which can be treated separately which helps to build stronger customer relationships. It is important to note that k-means and agglomerative clustering are only two of the many different types of clustering available to data analysts. There are many more accessible clustering methods, the use of which depends on the analyst's needs.

Classifiers

Very similar to clustering is the process of classification and, like clustering, there are many methods of classifying. One method that is easy to understand and set up is the k nearest neighbor (kNN) classification process. As was the case with k-means clustering, similarity is most often determined by distance which is

usually a geometric measure of distance between the new record and the closest k records. Each existing record has already been assigned to a class. The new records are compared to their closest neighbors and the class represented by highest number of close neighbors becomes the class of the new record.

Classification can be used to learn more information about new data. From a customer relationship standpoint, it is useful to determine how to approach new customers who do not have established buying patterns. The assumption is that customers who are similar, in many aspects, to already known customers are likely to have similar purchasing patterns and this can help define relationship strategies for unknown customers.

Association Rules

The creation of association rules, also often known as market basket analysis, allows companies to determine item sets or collections of products that have a higher likelihood of being purchased together. Through careful analysis of purchase records, confidence measures can help companies determine product sales as a function of their associations with other products. Counting the number of items of each type that are purchased and then counting associations of items purchased together allows for the determination of the likelihood that items will be bought together.

Culling through a transaction database makes it possible to create meaningful association rules leading to different approaches to product placement and marketing. Products with strong associations can be placed together in a store on a website encouraging customers to purchase both (or all) products at once. Beyond just a retail application, understanding associations between events can have far wider applications.

Accessibility of Analysis Tools

The previous sections described a number of analysis tools and a key argument of this article is that these tools are and should be accessible to a wider group of individuals who will benefit from their use in an increasingly data-driven environment. While advanced and deep-level analysis of large data sets may remain primarily the domain of a company's technical professionals, there are a number of widely accessible packages that provide access to many of the technical tools described in this article. These tools can be also included as a key part of non-technical university programs.

Microsoft Excel is, perhaps, the most widely accessible tool and its Data Analysis add-on package extends its capabilities in important ways. Certainly, data discovery, visualization, cleansing and reduction can all be done using the standard version of Excel but the newer versions and the Data Analysis add-on package provide tools such as forecasting and linear regression. Rudimentary segmentation, that resembles clustering, is also available in Excel but more advanced tools that can be used for logistic regression, clustering, classification and association rules can be accessed using open source scripting languages such as R or Python. Should analysts

have access to an advanced statistical package, such as SPSS, the use of these more advanced tools is made even easier.

Data analysis tools are useful at many levels in a company and the tools needed to conduct effective analysis are easily accessible. What may stop non-technical majors from using these tools upon graduation may have more to do with the extent to which they were, or were not, effectively covered in their university curriculum. Thus, there is an increasing need to infuse the use of these tools throughout the curriculum and to cover their use in real-world, authentic learning scenarios.

Conclusion

This article contains a brief description of analysis tools that could make up a non-technical analyst's data skill set. It is intended to describe common tools that are accessible and that should be included in a non-technical university education where they have the potential to be useful for graduates seeking employment in a data-driven economy. They are just a subset of analysis skills, yet ones that are useful and relevant to many of today's current data challenges.

But tools are just tools. The primary driver in any effort to utilize data and data analysis tools should be an analytic mindset, one in which questions are asked or opportunities are identified that drive data exploration (Rosidi, 2019). This is similar to the generation of hypotheses that drive research. Directive questions and goals should be identified at the outset of any data project as the selection of data sources and analysis methods are dependent on them. A data project without a purpose is analogous to a journey without a destination.

When planning university curricula, instructors need to consider the ongoing digital disruption that is currently impacting so many fields (Arthur, 2013). It is not going away and employees in today's companies need to adapt to the technological demands that are being placed on them. This includes data and the ability to analyze it. No longer can the responsibility for data analysis be relegated to technical departments who are likely overloaded with their own demands. Data analysis skills should and must become a standard part of the education that prepares future graduates, in many disciplines, for a data-driven marketplace.

References

Arthur, L. (2013). *Big data marketing: Engage your customers more effectively and drive value*. Wiley.

Evans, J. R. (2019). *Business analytics: Methods, models and decisions*. (3rd ed.). Pearson.

Mayer-Schonberger, V., & Cukier, K. (2013). *Big data*. Houghton Mifflin Harcourt Publishing.

Linhof, G. S., & Berry, M. J. A. (2011). *Data analysis techniques for marketing, sales, and relationship management* (3rd ed.). Wiley.

Patel, N. (2019). *10 ways data analysis can help you get a competitive edge.* Retrieved from https://neilpatel.com/blog/data-analysis/

Rosidi, N. (2019). *How I teach analytics to non-technical students.* Retrived from https://towardsdatascience.com/how-i-teach-analytics-to-non-technical-students-2db4a900f0cf

Siegel, E. (2016). *Predictive analytics: The power to predict who will click, buy, lie or die*. Wiley.

Shmueli, G., Bruce, P. C., Yahav, I., Patel, N. R., & Lichtendahl, K. C. (2018). *Data analysis for business analytics: Concepts, Techniques, and applications in R.* Wiley.

Taddy, M. (2019). *Business data science: Combining machine learning and economics to optimize, automate and accelereate business decisions.* McGraw Hill.

Machine Failure Prediction Based on Neural Network Analysis in a Steel Construction Factory

Mehmet Burçin Önay
Kilis 7 Aralık University, Turkey, https://orcid.org/0000-0002-6061-8270

Serap Ulusam Seçkiner
Gaziantep University, Turkey, https://orcid.org/0000-0002-1612-6033

Yunus Eroğlu
İskenderun Technical University, Turkey, https://orcid.org/0000-0002-8354-6783

Abstract: The improvement of condition monitoring of critical failures of steel construction processes provides the development of maintenance planning. The main aim of predictive failure is to find how the system failure is affecting the performance of steel construction processes. Based on man-hour and production amount, the maintenance planner defines the necessity for intervention on the steel construction processes. This study presents a tool regarding the operational performance of steel construction processes. The monthly real failures with man-hours and production amount are analyzed aiming at evaluating the critical failure patterns associated with specific failure modes. The findings of that failure route may be used to support the evaluation of any anomalous component operational condition that may affect steel construction processes. The artificial neural network was able to detect anomalies in steel processing. The use of worker behavior towards risk indicates that the definition of undesired performance is strongly affected by the operator's man-hour and related to the amount of production as a ton. If the scheduling of the operator would be made according to these predictions, there would be less unconformity within the maintenance staff and unexpected job rotations may not demotivate the steel construction operators.

Keywords: Artificial Neural Network, Steel Construction Manufacturing, Machine Failure Prediction

Introduction

The improvement of condition monitoring of critical failures of steel construction manufacturing processes provides the development of maintenance planning. The main aim of predictive failure is to find how the system failure is affecting the performance of steel construction manufacturing processes. This study is an experimental training operation carried on to train a neural network to perform a specific task and observe the results of the

Önay, M. B., Ulusam Seçkiner, S., & Eroğlu, Y. (2021). Machine failure prediction based on neural network analysis in a steel construction factory. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 11-21), New York, USA. ISTES Organization.

training operation and provide a discussion of the observations. The objective of this study is to create a neural network that predicts machine failure in a steel construction manufacturing factory based on a set of characteristics that are provided to the network about the work load of the factory.

In the complex manufacturing area (Thoben et al., 2017), where many elements (eg human, material and intangible resources) interact with each other, a large amount of data is collected during manufacturing processes. A computing infrastructure that is aware of the processed production data can be controlled by pre-trained Artificial Intelligence (AI) algorithms.

AI techniques are widely used to extract useful information from production data. AI techniques flow intelligence into systems to automatically learn and adapt to the changing environment using education and historical experience (Lu, 1990). In addition, the ability to use high-dimensional data, reduce complexity, improve existing data, and identify relevant process relationships are emphasized to demonstrate the applicability of AI techniques in the manufacturing industry (Wuest et al., 2016). In order to reduce the variety in the production line and improve productivity and product quality, the manufacturer has the opportunity to predict the issue with AI support. In this way, the future behavior of the production system can be approached by applying AI algorithms to the system, and this generated information can assist in decision making.

Meaningful information extracted; It provides insights into better decision-making that can help the transformation towards sustainable practices in the manufacturing industry, such as energy and resource efficiency (Lee et al., 2015), waste management (Moyne & Iskandar, 2016), predictive maintenance (PdM) (Kaiser & Gebraeel, 2009).

This study consists of five main sections: Basic Concepts, Dataset, Methodology, Results, and Discussion. The basic concepts section provides a definition of artificial neural networks, the concept of neural network training, and the various types of datasets used in the training operation. The dataset section is a description of the dataset that is used in this study (its values and columns). The methodology section describes the activities that were carried out to create the results of this paper. The results section lists the results that were generated. Lastly, the discussion section provides key information resulting from the analysis of the results.

Basic Concepts

Artificial Neural Networks

Artificial Neural Network is a solution method that mimics the function of human neural organs. Just like in a biological neural network, an artificial neural network consists of: one or more neurons, an input and an output. A neural network may consist of a single neuron, multiple neurons, or multi-layers of neurons. The components of a single neural network are: input array, weights array, a summer, a transfer function, and an output value. The network operates as follows: the input and weights arrays are multiplied to generate the weighted inputs array. The weighted inputs array is then sent to a summer, and then a transfer function is applied on the sum

value to generate an output value. The operation of multi-neuron and/or multi-layer networks is the similar to that of a single-neuron network; the difference is that in multi-neuron networks, the weighted inputs of all neurons are sent to the summer. Multi-layer networks are multiple layers of neurons, each layer consisting of multiple neurons, where the input of a layer is the output of a preceding layer; thus, there are three types of layers: input layer, hidden layers, and output layer. Artificial Neural Networks have a variety of applications in various sectors, including: Banking, Defense, Electronics, Entertainment, Manufacturing, Medical, Robotics, Telecommunications, and Transportation.

Neural Network Training

Neural networks are trained in order to generate the desired outputs and results. Training basically means adjusting the neural network weights to reduce the error in the neural network output. By reducing output errors, the neural network results become closer to the desired results. There are two types of training: supervised training and unsupervised training. In supervised training, the network's performance is either manually graded in order to indicate whether the performance is improving or not; or, the inputs and outputs are provided to the network and an error rate is calculated and propagated across the network to adjust weights and improve performance. The calculation of errors and adjustment of weights is done continuously until a stop condition is met (typically, a desired error rate or a number of iterations). In **unsupervised training**, the neural network is expected to make sense of the inputs and the outputs and adjust itself accordingly. Training starts with randomly assigned weights and then those weights are adjusted to improve the network performance.

Training Set

A training set is a collection of data instances used to train the neural network. Each data instance consists of values of all the network's inputs, coupled with a desired output value. The error value of a training set is propagated throughout the network to adjust weights and improve network performance.

Validation Set

A validation set is a collection of data instances used to validate the neural network performance. The validation set is not used to adjust weights. Instead, it is used to check whether the network is being trained to consider all possible instances, and not only those in the training set. The validation set is checked periodically during the training process, and if the validation error is unchanged or is not improving, the training process is terminated.

Test Set

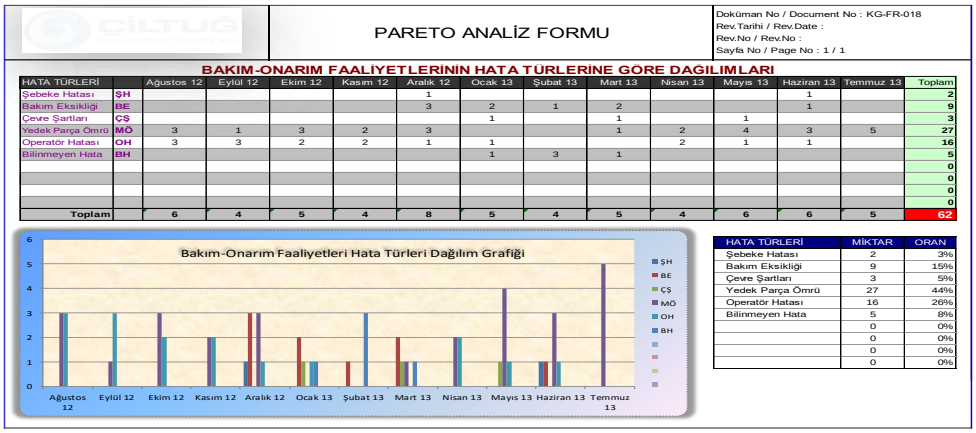
A test set is used to test the final result of the training process, i.e. the neural network in its final form after training has completed.

Dataset

The dataset used in this study was obtained from a steel construction manufacturing factory. Machine failure data were compiled by a machine technician with the aim of providing a statistical analysis and sustain the development of machine maintenance department. Man-hour and production amount (Ton) data are compiled with ERP program (Logo-Tiger Enterprise) and they are entered to the system by the responsible mechanical engineers of each project.

Output Values

Machine failure data were used as output data of our study. The data sheets are shown in Figure 1, Figure 2 and Figure 3. The values of December 2012 and July 2014 are omitted because of the values of that months were "8" and "0" respectively, and they are outliner values.



PARETO ANALİZ FORMU

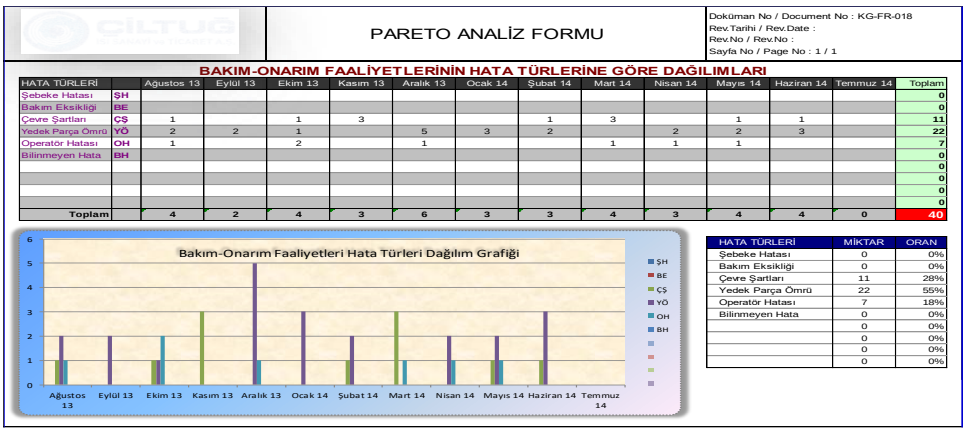
Doküman No / Document No : KG-FR-018
Rev.Tarihi / Rev.Date :
Rev.No / Rev.No :
Sayfa No / Page No : 1 / 1

BAKIM-ONARIM FAALİYETLERİNİN HATA TÜRLERİNE GÖRE DAĞILIMLARI

HATA TÜRLERİ		Ağustos 12	Eylül 12	Ekim 12	Kasım 12	Aralık 12	Ocak 13	Şubat 13	Mart 13	Nisan 13	Mayıs 13	Haziran 13	Temmuz 13	Toplam
Şebeke Hatası	ŞH					1						1		2
Bakım Eksikliği	BE					3	2	1	2			1		9
Çevre Şartları	ÇŞ						1		1		1			3
Yedek Parça Ömrü	MÖ	3	1	3	2	3			1	2	4	3	5	27
Operatör Hatası	OH	3	3	2	2	1	1			2	1	1		16
Bilinmeyen Hata	BH						1	3	1					5
														0
														0
														0
														0
Toplam		6	4	5	4	8	5	4	5	4	6	6	5	62

HATA TÜRLERİ	MİKTAR	ORAN
Şebeke Hatası	2	3%
Bakım Eksikliği	9	15%
Çevre Şartları	3	5%
Yedek Parça Ömrü	27	44%
Operatör Hatası	16	26%
Bilinmeyen Hata	5	8%
	0	0%
	0	0%
	0	0%
	0	0%

Figure 1. Machine Failure Data between August 2012-July 2013



PARETO ANALİZ FORMU

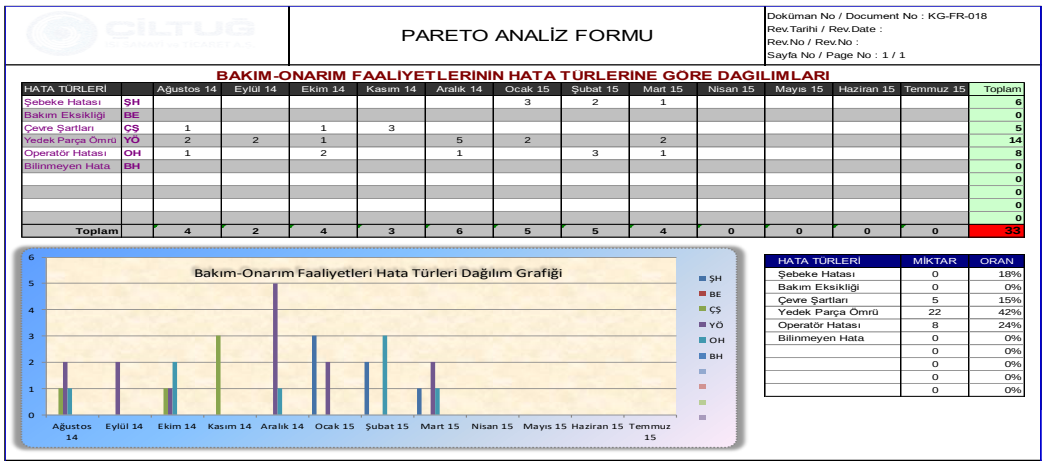
Doküman No / Document No : KG-FR-018
Rev.Tarihi / Rev.Date :
Rev.No / Rev.No :
Sayfa No / Page No : 1 / 1

BAKIM-ONARIM FAALİYETLERİNİN HATA TÜRLERİNE GÖRE DAĞILIMLARI

HATA TÜRLERİ		Ağustos 13	Eylül 13	Ekim 13	Kasım 13	Aralık 13	Ocak 14	Şubat 14	Mart 14	Nisan 14	Mayıs 14	Haziran 14	Temmuz 14	Toplam
Şebeke Hatası	ŞH													0
Bakım Eksikliği	BE													0
Çevre Şartları	ÇŞ	1		1	3			1	3		1	1		11
Yedek Parça Ömrü	YÖ	2	2	1		5	3	2		2	2	3		22
Operatör Hatası	OH	1		2		1			1	1	1			7
Bilinmeyen Hata	BH													0
														0
														0
														0
														0
Toplam		4	2	4	3	6	3	3	4	3	4	4	0	40

HATA TÜRLERİ	MİKTAR	ORAN
Şebeke Hatası	0	0%
Bakım Eksikliği	0	0%
Çevre Şartları	11	28%
Yedek Parça Ömrü	22	55%
Operatör Hatası	7	18%
Bilinmeyen Hata	0	0%
	0	0%
	0	0%
	0	0%
	0	0%

Figure 2. Machine Failure Data between August 2013-July 2014



HATA TÜRLERİ		Ağustos 14	Eylül 14	Ekim 14	Kasım 14	Aralık 14	Ocak 15	Şubat 15	Mart 15	Nisan 15	Mayıs 15	Haziran 15	Temmuz 15	Toplam
Şebeke Hatası	ŞH						3	2	1					6
Bakım Eksikliği	BE													0
Çevre Şartları	ÇŞ	1		1	3									5
Yedek Parça Ömrü	YÖ	2	2	1		5	2		2					14
Operatör Hatası	OH	1		2		1		3	1					8
Bilinmeyen Hata	BH													0
														0
														0
														0
														0
Toplam		4	2	4	3	6	5	5	4	0	0	0	0	33

HATA TÜRLERİ	MİKTAR	ORAN
Şebeke Hatası	0	18%
Bakım Eksikliği	0	0%
Çevre Şartları	5	15%
Yedek Parça Ömrü	22	42%
Operatör Hatası	8	24%
Bilinmeyen Hata	0	0%
	0	0%
	0	0%
	0	0%
	0	0%

Figure 3. Machine Failure Data between August 2014-March 2015

Input Values

Man-hour and production amount (Ton) values are used as input values of neural network. Many projects are carried out at the same time in the examined factory. So that man hour and project progress rates are defined by responsible project engineer. According to given progress rates total production amounts are calculated. At the end of each month total man-hour and produced tons are summed. So here, man-hour values are actually measured values, production amounts are multiplied values of progress rates with total product weights.

On the other hand; each project has its own degree of difficulty. Where a project can be 2 Tons with 600 man-hour requirement but another project can be 50 Tons with 500 man-hour requirement. This difference can affect the amount of machine strain. Then at the total of month, machine failures can be affected by this man-hour/Ton rates. So that; to have better result we used man-hour/Ton values as another input for our neural network.
Our data set table became as shown in the Table 1.

Table 1. Overall Dataset

#	Month	Manhour (Input-1) (AS)	Production (Input-2) (TON)	AS/TON (Input-3)	Failure (Input-4)
1	Aug-12	40.134	396	101	6
2	Sep-12	43.777	396	110	4
3	Oct-12	45.740	396	115	5
4	Nov-12	52.030	396	131	4
5	Jan-13	46.339	423	110	5
6	Feb-13	41.343	423	98	4
7	Mar-13	44.683	426	105	5
8	Apr-13	43.435	439	99	4
9	May-13	40.990	440	93	6
10	Jun-13	36.668	455	81	6
11	Jul-13	35.224	459	77	5

12	Aug-13	35.514	459	77	4
13	Sep-13	46.697	457	102	2
14	Oct-13	38.187	191	200	4
15	Nov-13	43.495	247	176	3
16	Dec-13	41.175	199	207	6
17	Jan-14	49.473	402	123	3
18	Feb-14	49.473	402	123	3
19	Mar-14	49.473	402	123	4
20	Apr-14	49.473	402	123	3
21	May-14	49.473	402	123	4
22	Jun-14	49.473	402	123	4
23	Aug-14	49.473	402	123	4
24	Sep-14	49.473	402	123	2
25	Oct-14	49.473	402	123	4
26	Nov-14	49.473	402	123	3
27	Dec-14	49.473	402	123	6
28	Jan-15	49.003	234	209	5
29	Feb-15	40.060	234	171	5
30	Mar-15	43.412	234	186	4

Methodology

The purpose of this study is to create a forecasting tool for the dataset. In order to understand the relationship between the input and output data, a regression test was created using Excel. The regression test was made to show whether creating the tool with neural networks was a viable option. The Microsoft Excel Data Analysis component was used to perform a regression analysis on the dataset.

The neural network was designed as follows: Data are separated in to two different excel files which are input and output files. Nftool of Matlab is used for neural network application. The dataset was loaded into the Matlab software applicationvia nftool. The following steps were followed in order to register the data in the Analysis Tables.

1. Open matlab
2. Create two new variables named input and target
3. From excel file we will copy the input matrix and target matrix and paste in the variables
4. Type nftool command
5. Press next button on the nftool window
6. Select input variable on the Inputs label
7. Select target varialble on the Targets label
8. Select matrix rows on the “Samples are” label
9. Press next button

10. Divide 20 samples for training, 5 samples for validation, 5 samples for testing as seen in the default values
11. Pressed next button
12. Define number of hidden neurons 10, as seen in the default values,
13. Pressed next button
14. Waited for the network creation
15. Choosed training algorithm as Levenberg-Marquardt
16. Pressed Train button
17. R and MSE values are calculated for Training, Validation and Testing sets
18. A MAPE variable is defined
19. Value A is calculated as the transpose of the forecasting matrix
20. Mape formula is defined as; MAPE=mean(abs(output-A)./output*100)

After these steps, 8 trainings applied then we saw that there are always low regression values between output and forecasting values. So we used the (t-1)'th month's value for the forecasting of the t'th month's failure number. Which means we supposed that we were at the end of a month, we were knowing the number of machine failure of that month, and we were trying to forecast the succeeding month's failure value. Subsequent to that adjustment we divided 24 samples for training (%80), 3 samples for validation (%10) and 3 samples for test (%10), and finally we increased number of hidden neurons to 50, then we had better results. And we choose the best result that we had seen.

Results

Regression Analysis

The results of the regression analysis applied on the data are in Figure 4.

Regression Statistics	
Multiple R	0,380681
R Square	0,144918
Adjusted R Square	0,085947
Standart Error	1,431198
Observations	32

Figure 4. Regression Analysis Results

Neural Network Training

In this section, the results of the activities described in the Methodology section are presented. Before the adjustment, 8 trainings are examined and their results are shown in Table 2.

Table 2. Training Results before Adjustment

Levenberg-Marquardt Algorithm Training Results Before Adjustment								
Training No	1	2	3	4	5	6	7	8
Epoch	23	9	8	16	9	9	8	6
Time	00:00:01	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00	00:00:00
Performance	27,8	7,7	12,9	3,02	14,2	19,1	8,2	9,64
Gradient	53,7	16,2	29,2	14,1	35,7	37,7	21	15,8
Mu	0,001	0,001	0,001	0,001	0,001	0,001	0,001	0,001
Validation Checks	6	6	6	6	6	6	6	6
All Regression R Value	0,68	0,5	0,51	0,59	0,3	0,36	0,24	0,04

After seeing these low R values, data is adjusted as explained in the method section, data is divided to training, validation and test sets as %80, %10 and %10 respectively, then training trials are made with 50, 60 and 70 hidden layers. The results of the model that gave the lowest MAPE value 8,15% out of about 100 trials are given between Figure 5-Figure 7. The graphs in Figure 5 and Figure 6 were generated by Matlab. Figure 7 was created in Excel.

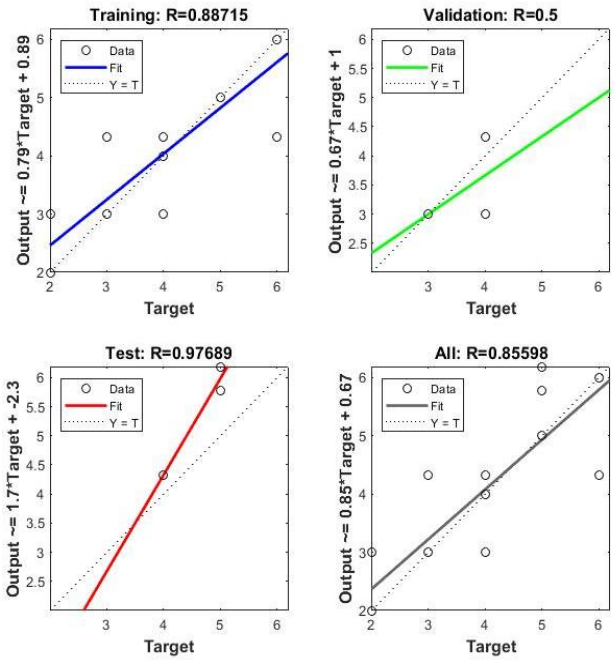


Figure 5. Regression Graphs

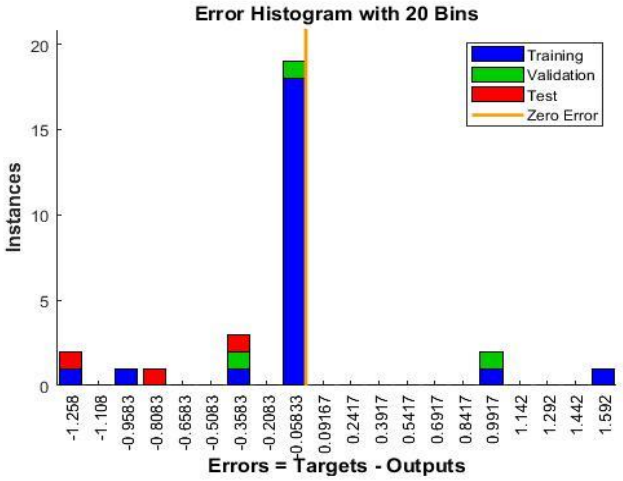


Figure 6. Error Histogram

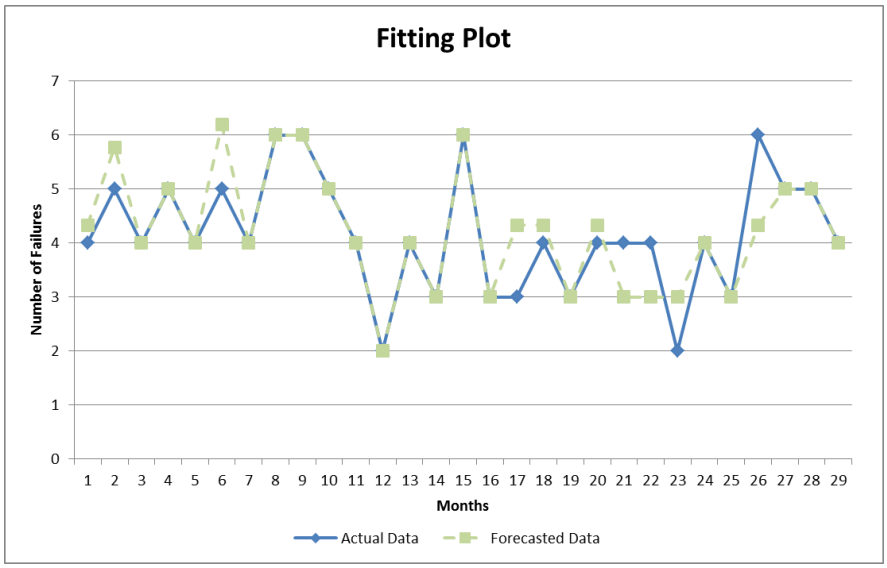


Figure 7. Fitting Plot Diagram

Conclusion and Discussion

As per the regression analysis shown in section 5.1, it can be observed that the correlation coefficient (Adjusted R Square) equals ~0.085. This means that the relationship between the dependent and independent variables (i.e. the input and output variables) is poor, therefore, designing an ANN to perform this study is a viable option.

We started to this study with 32 data. But after the elimination of two outliners 30 data remained. For the first 8 training data we had very low R values. So we adjusted the data as explained in Section 4. After this adjustment we increased the number of training set and hidden neurons. Then we made about a hundred trainings. We calculated MAPE values for good solutions in MATLAB but, the best MAPE value is calculated as 8.15%. Forecasted data of this solution is shown in Table 4. In the result of this training process R values of sets are

shown in Table 3. 19 forecast errors are very close to 0 which is shown in Figure 6.

Table 3. R Values

Sets	R Values
Training	0,887
Validation	0,5
Test	0,977
All	0,856

Table 4. Actual and Forecasted Data

Months	Actual Data	Forecasted Data
1	6	-
2	4	4,3283
3	5	5,7732
4	4	4
5	5	5
6	4	4
7	5	6,191
8	4	4
9	6	6
10	6	6
11	5	5
12	4	4
13	2	2
14	4	4
15	3	3
16	6	6
17	3	3
18	3	4,3333
19	4	4,3333
20	3	3
21	4	4,3333
22	4	3
23	4	3
24	2	3
25	4	4
26	3	3

27	6	4,3333
28	5	5
29	5	5
30	4	4
	MAPE	8,15

The proposed model can be used by managers to plan maintenance department and make more effective personal schedules. And also having a good forecast about the succeeding month's machine failure can be beneficial for production planning and budget management.

In this study we have limitations because of the number of data, If we could have more data we could use more data for validation and test sets. By this way we could have better solutions.

In the future studies we can obtain maintenance cost data and with these data we can use machine failure numbers as input and forecast maintenance cost. In additionally, we can make another study to forecast the probability of failure type by using FMEA method.

References

Kaiser, K. A., & Gebraeel, N. Z. (2009). Predictive Maintenance Management Using Sensor-Based Degradation Models. *Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans*, *39*(4), 840–849. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4914831

Lee, J., Bagheri, B., & Kao, H. A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, *3*, 18–23. https://doi.org/10.1016/j.mfglet.2014.12.001

Lu, S. C. Y. (1990). Machine learning approaches to knowledge synthesis and integration tasks for advanced engineering automation. *Computers in Industry*, *15*(1–2), 105–120. https://doi.org/10.1016/0166-3615(90)90088-7

Moyne, J., & Iskandar, J. (2016). *Big Data Analytics for Smart Manufacturing: Case Studies in Semiconductor Manufacturing*. https://doi.org/10.3390/pr5030039

Thoben, K. D., Wiesner, S. A., & Wuest, T. (2017). "Industrie 4.0" and smart manufacturing-a review of research issues and application examples. *International Journal of Automation Technology*, *11*(1), 4–16. https://doi.org/10.20965/ijat.2017.p0004

Wuest, T., Weimer, D., Irgens, C., & Thoben, K.-D. (2016). Machine learning in manufacturing: advantages, challenges, and applications. *Production & Manufacturing Research*, *4*(1), 23–45. https://doi.org/10.1080/21693277.2016.1192517

Tumor Dynamics under Immunotherapy: A Time Delay Modeling Approach

Samantha Condo

Northeastern Illinois University, USA, https://orcid.org/0000-0001-6914-8829

Emma Turian

Northeastern Illinois University, USA, https://orcid.org/0000-0003-0718-4298

Abstract: Cancer remains a leading cause of death worldwide. Recent research into novel approaches to treatment has suggested that immunotherapy, may be a promising strategy. Quantitative models simulating the dynamics of tumor-immune system interaction, can facilitate both basic and clinical research efforts aimed at better understanding the impact of immunotherapy in the management of the disease. Previous studies of these dynamics used Michaelis-Menten dynamics to describe effector cells evolution. The goal of this study, is to propose an evolutionary two population model based on a logistic growth with a dynamic carrying capacity of the effector cells with discrete delays, in order to assess their impact on tumor dynamics. Our model suggests that tumor dormancy exists for certain parameters describing effector cell stimulation and innactivation. Further system analysis allows us to arrive at critical conditions for tumor boundness and its equilibrium. By employing a two-fold sensitivity analysis, we determine the most influential parameters that describe the model, and use this information to prioritize our findings. Our analysis suggests that under certain conditions it is possible not only to control tumor growth but also have tumor size reduced due to the response of the immune system.

Keywords: Immunotherapy, Dynamic Carrying Capacity, Discrete Delay, Monte Carlo Simulations

Introduction

Immunotherapy is often used along radiation therapy (RT) and chemotherapy to treat tumors. Immunity is the state of protection against foreign pathogens or substances (i.e., antigens), and it can be divided into categories, such as [American Cancer Society 2019]:

a) Checkpoint inhibitors are drugs that help the immune system recognize and fight cancer cells.

b) Cytokines are proteins that help stimulate the immune system attack cancer cells.

c) Vaccines against cancer which can be preventive or therapeutic.

d) Monoclonal antibodies are man-made versions of immune system protein design to attack a specific

Kostromina E., Eremin P., Veremeev A., Putova M., Kudryashova I., & Gilmutdinova I. (2021). Comparative analysis of cell-based products obtained with different systems for isolation of stromal vascular fraction from human adipose tissue. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 36-42), New York, USA. ISTES Organization.

part of a cancer cell.

The immune system in humans includes two main components, the innate and adaptive, that protect our bodies from pathogens. The innate system acts very quickly and it represents the first line of defense against foreign substances. This system includes the natural killer (NK) cells which specialize in first identifying cells with changes in their surface and then attacking this surface with cell toxins. The adaptive immune system consists of T lymphocites, B lymphocites, and antibodies, and destroys germs that could not be killed by the innate system. Since this system requires identification of such germs, and thus it takes longer to have an effect. This type of immune response remembers these germs next time we get sick by the same disease. It is known that it might take a few days for this immune system to attack foreign cells [IQWiG 2020], or longer (up to 7 days) [Mahlbacher 2019]. The other aspect of the immune response is the fact that the presence of cancer cells activates the immune cells. Effector cells are short-lived activated cells that defend our body in an immune response [Encyclopaedia Britannica 2019].

Quantitative models simulating the dynamics of tumor-immune system interaction, can facilitate both basic and clinical research efforts aimed at better understanding the impact of immunotherapy in the management of the disease. Previous modeling of tumor-immune interaction employs the Michaelis-Menten kinetics term to describe the rate-limiting recruitment of the effector cells, [de Pillis 2014]. In order to incorporate the effector cells response to cancer described earlier, our work replaces the Michaelis-Menten growth term for the effector cells by a logistic type that includes a dynamic carrying capacity coupled with a discrete delay that measures the timely response of the effector cell population. These two choices for the dynamic carrying capacity and discrete delay incorporate the following reasoning. The choice for the Verhulst logistic term with dynamic carrying capacity is explained by the fact that the presence of cancer cells activates the immune cells. Thus, the immune cells should reach a carrying capacity proportional to cancer cells concentration at time t. We assume that the delay, τ, represents the time from the formation of the effector cells until there is an effect on the cancer cell poplation. This is based on Hutchinson's (1948) model on the reproduction of eggs in a single species. Since reproduction does not happen instaneously, the egg formation happens τ number of time units before hatching [Ruan 2006].

We propose a two-population model that includes the cancer cell concentration, C, and the effector cell population, E. The goal of this proposed model is to assess the impact of such effector cell dynamic on cancer cell evolution.

We analyze our proposed sytem by looking at its properties. System analysis suggests the existence of critical conditions for tumor boundness and its equilibrium. Using numerical simulations with parameter values from experimental literature by Kuznetsov and Knott [Kuznetsov 2001], we evaluate the dynamic of the system. Numerical simulations suggest that tumor dormancy exists for certain parameters describing effector cell stimulations and discrete delays. Our two-fold sensitivity that includes parameter perturbation and Monte Carlo simulations suggests that parameters related to effector cells evolution produce the lowest cancer cell

concentration. By finding the most influential parameters that describe the model, we are then able to use this information to prioritize our findings. Our hope is that such findings could assist oncologists in their search to finding a path towards establishing the potential of immunotherapy effect on cancer dynamic.

Our paper is organized as follows. The first part describes the evolution of tumor-effector interaction starting with earlier models up until more recent approaches. Then, we descibe our model and its properties followed by numerical simulation results. Next, we descibe our findings from the two-fold sensitivity analysis, followed by a discussion and a conclusion section.

Earlier Two Population Models of Tumor-Immune Response

The simplest model for the logistic growth of two competing species [Wodarz 2014] is represented by Equations $0(\mathrm{a})$:

$$\begin{cases} \frac{dX}{dt} = a_1 X\left(1 - \frac{X}{M}\right) - a_1 r_{21}\left(\frac{XY}{M}\right), \\ \frac{dY}{dt} = a_2 Y\left(1 - \frac{Y}{M}\right) - a_2 r_{12}\left(\frac{XY}{M}\right), \end{cases}$$

Equations $0(a)$.

where, a_1 and a_2 are the respective growth rates for the X and Y populations, respectively; and r_{21} and r_{12} are the proportions of respective cell types that first bound and are later destroyed. In this model the X and Y populations are both growing logisticaly in the absence of the other population. These equations are given different growth rates, but they are structured to grow and die at the same rate. For the cancer and effector cells population this represents an oversimplified model because these two types do not behave in the same way.
In all the following models, $C = C(t)$ and $E = E(t)$ denote the concentrations of cancer cells and effector cells at time t, respectively, and M, the logistic carrying capacity for cancer cells concentration.

The simplest predator-prey used to model the effector-immune interaction [de Pillis 2014], where tumor is the prey and the effector cells are the predators is described by Equations $0(b)$.

$$\begin{cases} \frac{dC}{dt} = a\left(1 - \frac{C}{M}\right)C - b_1 CE, \\ \frac{dE}{dt} = -\delta E + b_2 CE. \end{cases}$$

Equations $0(b)$.

In the model described by Equations $0(b)$ only the tumor grows logistically at a rate a and carrying capacity M, in the absence of the effector cells. The effector cells will die in the absence of cancer cells at a rate described by δ. The interactions between the two types of cells are harmful to the tumor and beneficial to the effector cells.

These "mass-action" terms, represented by b_1CE and b_2CE, use the assumption that the encounters between cancer and effector cells are proportional to the product of their populations. Tumor cells are killed by the effector cells by having their membrane being damaged by a protein called perforin or by the start of apoptosis by another protein called FasL. Since the effector cells produce these proteins in a limited amount, the interaction between the tumor and the effector cells lowers the ability of effector cells to kill cancer cells. This is the reason for later introducing a negative "mass-action" term in the effector cells evolution equation [de Pillis 2014].

A more realistic model is described by Equations $0(c)$:

$$\begin{cases} \frac{dC}{dt} = a\left(1 - \frac{C}{M}\right)C - b_1CE, \\ \frac{dE}{dt} = \alpha + f(C,E) - b_2CE - \delta E. \end{cases}$$

Equations $0(c)$.

System of Equations $0(c)$, [Kuznetsov 2001, dePillis 2014] represents a reduced system from Kuznetsov and Taylor's [Kuznetsov 1994] study on tumor regrowth and immunotherapy. In the absence of effector cells, tumor also grows logistically at a rate a with a carrying capacity denoted by M. This model addresses some of the issues from the simplest model (see Equations $0(b)$). For example, the interaction term for the immune cells is now set to be decreasing. Since cancer cells stimulate the production of immune cells, the arrival of the effector cells at the tumor site is described by a rate limiting recruitment term in the form of Michaelis-Menten function $f(C, E)$. Effector cells exist even in the absence of cancer cells and are produced in the bone marrow. Afterwards, they travel throughout our bodies and search for harmful cells. The constant source rate of effector cells is decribed by α. Since not all effector cells will have access to the tumor, these cells will grow and die on their own at a rate δ.

Mathematical Model with Hutchinson's Type Discrete Delay

We build our model from the system presented in Kuznetsov and Knott [Kuznetsov 2001], which represents a reduced version of the one described in Kuznetsov et al. [Kuznetsov 1994]. We change the Kuznetsov and Knott [Kuznetsov 2001] model by replacing the Michaelis-Menten rate-limiting term in the effector cells equation by a Verhulst logistic term with dynamic carrying capacity and Hutchinson's delay.

The evolution of the effector and cancer cells is described by the system of Equations (1) where C, and E denote the concentrations of cancer cells and effector cells, respectively.

The assumptions are as follows:

a) In the absence of the effector cells, cancer cells grow according with the logistic law, at an intrinsic

rate, a, and tumor carrying capacity (TCC), M.

b) The interactions between cancer cells and effector cells are harmful to cancer and effector cells.

c) The effector cell population has a constant source rate, α.

d) The delayed effect of the effector cells is described by Hutchinson's type of logistic growth with delay.

e) Since the presence of tumor cells stimulates the response of the effector cells, effector cells attain a time dependent carrying capacity proportional to the concentration of the cancer cells, described by $C(t)/\nu$.

f) The effector cells are dying at a rate described by δ.

$$\begin{cases} \dfrac{dC(t)}{dt} = a\left(1 - \dfrac{C(t)}{M}\right)C(t) - b_1 C(t)\,E(t), \\ \dfrac{dE(t)}{dt} = \alpha + \beta E(t)\left(1 - \dfrac{\nu E(t-\tau)}{C(t)}\right) - b_2 C(t)\,E(t) - \delta E(t). \end{cases}$$

Equations 1. Evolution Equations for the Cancer and Effector Cells

To analyze our system of Equations 1, we use parameter values from experimental literature by Kuznetsov and Knott [Kuznetsov 2001]. The experiment uses data from chimeric mice injected into the spleen with BCL_1 lymphoma tumor cells. The mean tumor cells from four groups of mice were recorded. We use the initial tumor cells concentration, $C(0) = 0.5$ millions of cells per mouse from the first two groups, and the initial effector cells concentration, $E(0) = 0.3$ millions killer cells present at the time of injection. The values for the remaining parameters are outlined in Table 1.

The effectiveness at time t of the effector cells depends on the concentration of these effector cells at t- τ . Since cancer cells stimulate the production of effector cells, and cancer cells reach a carrying capacity, M, then the effector cells reach a carrying capacity proportional to the size of cancer concentration at time t, namely, C(t)/ ν. Choices for τ and ν are as follows. Based on the findings presented in the previous section, we analyze the system of equations using two extreme delays: $\tau = 3$, and $\tau = 8$. The choice for the carrying capacity for the effector cells (CCE= C(t)/ ν) and its factor, ν, is explained by the four cases (see Figures 1-4) illustrating different carrying capacity factors for the effector cells relative to cancer cells. Figures 1-4 show the time evolution of the cancer, C, and effector, E, cell concentrations for $\tau = 8$. These figures suggest that dynamic carrying capacity for the effector cells needs to be at least twice that of cancer cells at any time, t, for the tumor to reach a limiting carrying capacity. Similar outcome plots are obtained by using $\tau = 3$ days (plots not shown here). Thus, we employ for our numerical simulations the bifurcation value, $\nu = 0.5$.

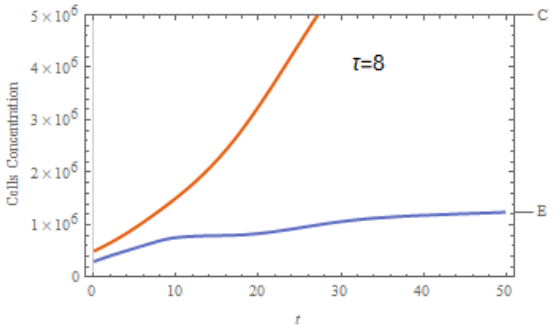


Figure 1. CCE same as C(t).

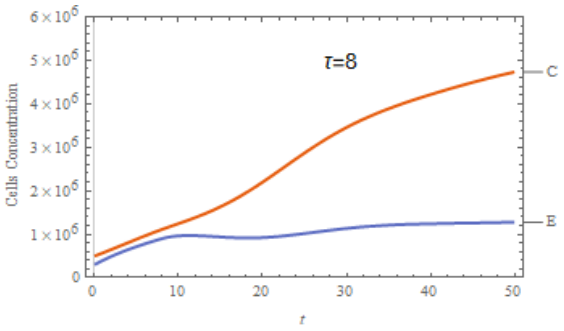


Figure 2. CCE is twice C(t).

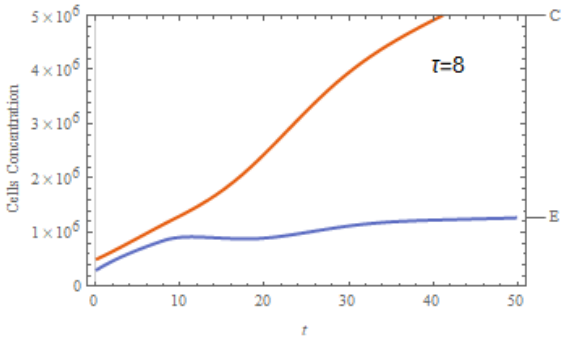


Figure 3. CCE is 5/3 times C(t).

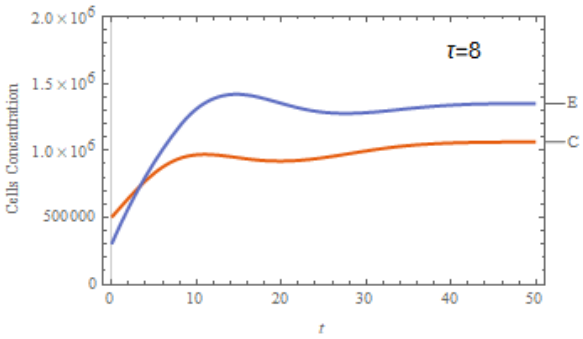


Figure 4. CCE is 10 times C(t).

Table 1. Paramters used in Numerical Simulations [Kuznetsov 2001]

Parameter	Value
a	**0.1877 (1/days)**
M	**$1/(0.00188 \times 10^{-6})$ (cells)**
b_1	**0.13845×10^{-6} (1/days×cells)**
α	**0.177×10^{6} (cells/days)**
β	**0.525 (1/days)**
b_2	**2.4966×10^{-10} (1/days×cells)**
δ	**0.59 (1/days)**

Analysis

Nonnegativity and Boundedness of the Solutions for Equations (1).

We show that the set $(C(t), E(t))$ is positive, [Tarfulea 2018].

Lemma 1:

Let $(C(t), E(t))$ be a continuous solution set of system of Equations 1, and $C_0, E_0 : (-\infty, 0] \rightarrow (0, \infty)$ be continuous initial conditions. Assume that $C_0, E_0 > 0$, for a t $\in (-\infty, 0]$. Then $C(t) \in [0, \infty)$, $E(t) \in (0, \infty)$, for $t \in (0, \infty)$.

Proof:

From continuity of the initial conditions, $C(t), E(t)$ are positive in a δt_+ (right side neighborhood of the initial $t = 0$).

Suppose there exists t_1 outside δt_+ with t_1>0, such that $C(t_1) = 0$. Then from the first equation from the system of Equations (1), $\frac{dC}{dt}_{|t=t_1} = 0$. Thus, since C reaches $C = 0$, and it is constant at $t = t_1$, $C(t) \in [0, \infty)$.

Suppose there exists t_2 outside δt_+ with t_2>0, such that $E(t_2) = 0$ first time. Then, $\frac{dE}{dt}_{|t=t_2} = \alpha > 0$. Thus E is increasing at $t = t_2$. This contradicts E increasing, $E > 0$ for $t < t_2$. Therefore, $E(t) \in (0, \infty)$. ■

Similar argument can be made for a piecewise continuous set $(C(t), E(t))$.

Gronwall Inequality:

For g and C real valued continuous functions in I (interval starting at a), if C is differentiable on the interior of I, and satisfies $C'(t) \leq g(t)C(t)$ for t $\in$ interior of I, then $C(t) \leq C(a)e^{\int_a^t g(s)ds}$.

Notation: Let c, and e be the antiderivatives of C, and E, respectively. We denote: $\frac{\Delta c}{\Delta t} = \frac{c(t)-c(0)}{t-0}$, and $\frac{\Delta e}{\Delta t} = \frac{e(t)-e(0)}{t-0}$.

Lemma 2.

Let $C(t)$ and $E(t)$, the solutions of Equations 1, be integrable functions. If the relative (to mitosis/lysis rates) growth rate of the effector cells concentration $\frac{1}{a/b_1}\frac{\Delta e}{\Delta t} > 1 - \frac{1}{M}\frac{\Delta c}{\Delta t}$, then $C(t) \leq C(0)$ for any time, t.

Proof:

From the first equation from system of Equations 1, by using the Gronwall inequality, for $I = (0, \infty)$, $t \in$ interior of I, and $g(s) = a - \frac{aC(s)}{M} - b_1 E(s)$,

$C(t) \leq C(0)\, e^{\int_0^t \left(a - \frac{aC(s)}{M} - b_1 E(s)\right) ds} = C(0) e^{at - \left(\frac{a}{M}\int_0^t C(s)ds + b_1 \int_0^t E(s)ds\right)}$.

Assume $\frac{1}{M}\frac{\Delta c}{\Delta t} + \frac{1}{a/b_1}\frac{\Delta e}{\Delta t} > 1$.

Then the expression: $at - \left(\frac{a}{M}\int_0^t C(s)\, ds + b_1 \int_0^t E(s)\, ds\right) < 0$.

This implies $C(t) \leq C(0)$ for any time, $t \in$ interior of I. ■

Classical Gronwall Inequality [Tarfulea 2018]:

If function $E:[0,T]\to\mathbb{R}$ and $E'(t)\le h(t)+g(t)E(t)$ with g continuous and h locally integrable, $E(t)\le E(0)e^{G(t)}+\int_0^t e^{G(t)-G(s)}h(s)ds$ for $G(t):=\int_0^t g(s)\,ds$.

In Lemma 3, we show that E, the solution from Equation 1 is bounded from above, for a special case.

Lemma 3.

Let $C(t)$ and $E(t)$, be the solutions of Equations 1. Let $E:[0,T]\to\mathbb{R}_{>0}$ and $E'(t)\le\alpha-\delta E(t)$. If $(t-\tau)\ge\frac{C(t)}{\nu}$, meaning that the production of effector cells at a time $t-\tau$ exceeds that of their carrying capacity at time t, then $E(t)\le E(0)+\frac{\alpha}{\delta}$.

Proof:

Assume $E(t-\tau)\ge\frac{C(t)}{\nu}$. From the second equation of system of Equations 1,

$$\frac{dE}{dt}=\alpha+\beta E(t)\left(1-\frac{\nu E(t-\tau)}{C(t)}\right)-b_2C(t)E(t)-\delta E(t)\le\alpha-\delta E(t).$$

Using classical Gronwall inequality with $h(t)=\alpha$ and $g(t)=-\delta$, we have

$$E(t)\le E(0)e^{-\delta t}+\frac{\alpha}{\delta}-\frac{\alpha}{\delta}e^{-\delta t}\le E(0)+\frac{\alpha}{\delta}\text{ for any } t\in[0,T].$$ ■

Lemma 4.

There are no periodic solutions to Equations 1, with $\tau=0$, in $D=(\mathbb{R}^+,\mathbb{R}^+)$.

Proof:

Let $\dot{X}=(\dot{C},\ \dot{E})$. Define the auxiliary differentiable, real-valued function $h(X)=\frac{1}{CE}$. The following divergence

$$\nabla\cdot(h\dot{X})=-\frac{aC^2E+\alpha MC+\beta\nu ME^2}{MC^2E^2}<0.$$

Based on Bendixon-Dulac Theorem, the system of Equations 1 has no closed orbits lying entirely in D. ■

Linear Analysis

We employ the following non-dimensional groupings:

$$\hat{C}=\frac{C}{M},\hat{E}=\frac{\nu E}{M},\hat{t}=at,\hat{\tau}=a\tau,\hat{a}=\frac{b_1M}{\nu a},\hat{b}=\frac{b_2M}{\nu a},\hat{\alpha}=\frac{a\nu}{aM},\hat{\beta}=\frac{\beta}{a},\hat{\delta}=\frac{\delta}{a}.$$

Here, ν is a nondimensional factor. Nondimensionalization of the term including delay is described in Murray [Murray 2002].

The dimensionless version of system of Equations 1 is:

$$\begin{cases} \dfrac{d\hat{C}}{d\hat{t}} = \hat{C}(1-\hat{C}) - \hat{a}\hat{C}\hat{E}, \\ \dfrac{d\hat{E}}{d\hat{t}} = \hat{\alpha} + \hat{\beta}\hat{E}\left(1 - \dfrac{\hat{E}(\hat{t}-\hat{\tau})}{\hat{C}}\right) - \nu\hat{b}\hat{C}\hat{E} - \hat{\delta}\hat{E}. \end{cases}$$

Equations 2. Dimensionless evolution equations.

Lemma 5.

There exists an approximate tumor free equilibrium for system of Equations 2, with $\tau = 0$:

$$\begin{cases} \hat{C}^* \cong 0, \\ \hat{E}^* \cong \sqrt{\dfrac{\hat{\alpha}}{\gamma}}, \end{cases}$$

where $\gamma = \frac{\hat{\beta}}{\hat{c}}$.

Proof:

By taking $\left(\frac{d\hat{C}}{d\hat{t}}, \frac{d\hat{E}}{d\hat{t}}\right)=(0,0)$, we find the solution

$$\begin{cases} \hat{C}^* \cong 0, \\ \hat{E}^* \cong \dfrac{\hat{\beta} - \hat{\delta} \pm \sqrt{\hat{\beta}^2 + 4\hat{\alpha}\gamma - 2\hat{\beta}\hat{\delta} + \hat{\delta}^2}}{2\gamma}. \end{cases}$$

Since $\gamma \gg (\hat{\beta} - \hat{\delta})^2$, $\hat{E}^* \cong \frac{\hat{\beta}-\hat{\delta}}{2\gamma} + \sqrt{\frac{\hat{\alpha}}{\gamma}} \cong \sqrt{\frac{\hat{\alpha}}{\gamma}}$.

We linearize about the approximate tumor free steady state by setting:

$$\begin{cases} x(t) = \hat{C}(t) - 0, \\ y(t) = \hat{E}(t) - \sqrt{\dfrac{\hat{\alpha}}{\gamma}}. \end{cases}$$

Let A be the Jacobian matrix at the steady state from Lemma 5. Since the eigenvalues from $|A - \lambda I| = 0$, are $\lambda_1 = 1 > 0$, and $\lambda_2 = \hat{\beta} - \hat{\delta} - 2\sqrt{\hat{\alpha}\gamma}$. For $\hat{\beta} - \hat{\delta} < 2\sqrt{\hat{\alpha}\gamma}$ the equilibrium $(\hat{C}^*, \hat{E}^*)$ is a saddle point.

Results

Numerical Simulation Results

Figures 5 and 6 show the two trajectories for the cancer and effector cell concentrations for the first 200 days. The model suggests that a stable level of effector cells leads to tumor equilibrium. An increased delay in the effect of the effector cells (from 3 to 8 days) means a slighter higher overall cancer cell concentration for the first 50 days (with the exception of a small dip in cancer cell concentration following an oscillatory effector cell concentration). Also, the model suggests that an increased delay in the effect of the effector cells leads to an oscillatory behavior of the effector cells during the first 50 days.

This experimental parameter values suggest that for the treatment to be effective (the number of cancer cells levels off) the effector cells stimulations need to occur at intervals in the order of two to three times the respective time delay (τ) value. This can be visualized in Figure 6 by checking the start of each E stimulation marked by the respective arrows.

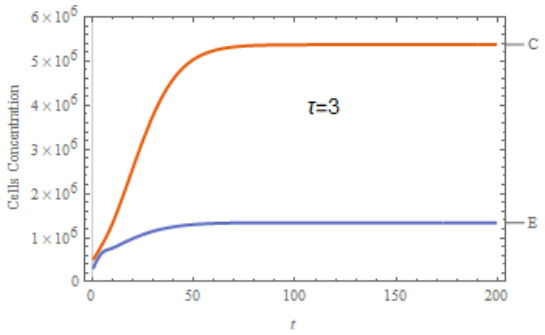


Figure 5. C and E as functions of time, $\tau = 3$.

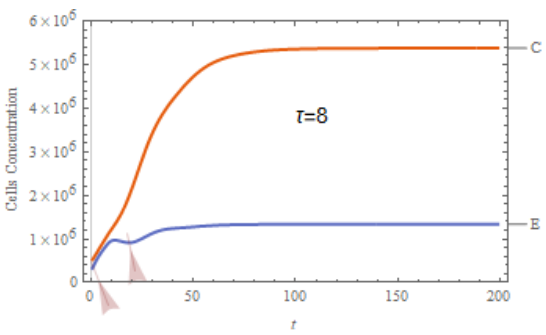


Figure 6. C and E as functions of time, $\tau = 8$.

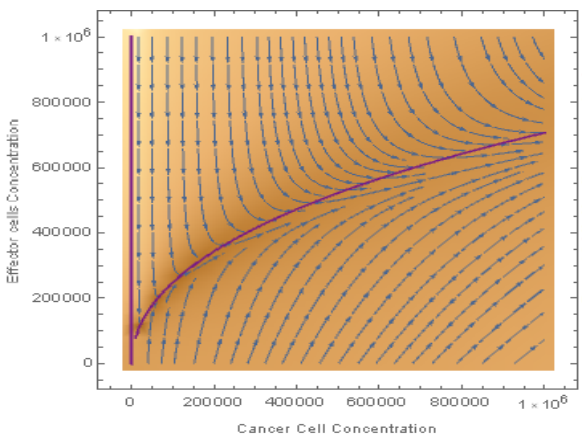


Figure 7. Stream plot with nullclines.

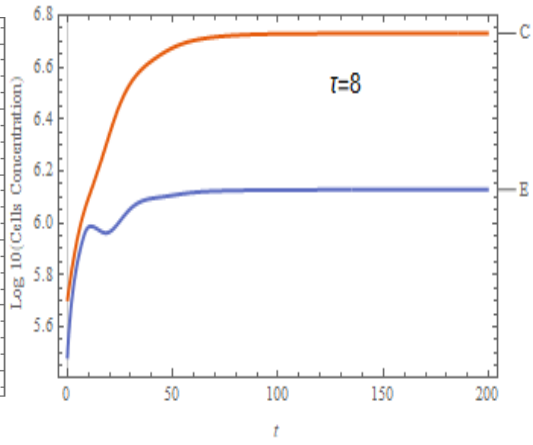


Figure 8. Log(Cell Concentration).

Figures 7 represents the stream plot of the vector field, superimposed on a background that is based on a logarithm of the field magnitude, and Figure 8 represents the Log_{10}(Cell Concentrations) as functions of time. The nullclines (shown in purple) from Figure 7 suggest the existence of an approximate tumor free equilibrium.

Sensitivity Analysis Results

In order to establish which parameters are the most influential (critical to the dynamics of cancer) for the system of Equations 1, we perform a sensitivity analysis for a delay $\tau = 3$, using two different methods: Parameter perturbation and Monte-Carlo simulations. The former records the cancer cell concentration for when parameters increase and decrease. The details of such illustrations are outlined in [dePillis 2014]. For the latter method we choose to illustrate the coefficient of variation (CV) and the mean of the lowest tumor concentration values.

Parameters Perturbation

Parameters are perturbed from their values by 5%, and the tumor cells concentration at the bifurcation $t = 50$ days is recorded. Figure 9 illustrates the cancer cells concentrations for the respective increase or decrease in parameter values. The figure suggests that the death rate of the effector cells, δ (marked in red), has the largest influence on cancer concentration since its increase produces the largest tumors, and its decrease leads to smallest tumors, among all existing parameter changes. This parameter is closely followed by the rate of effector population increase, β, whose changes has opposites effects to the cancer cells concentration.

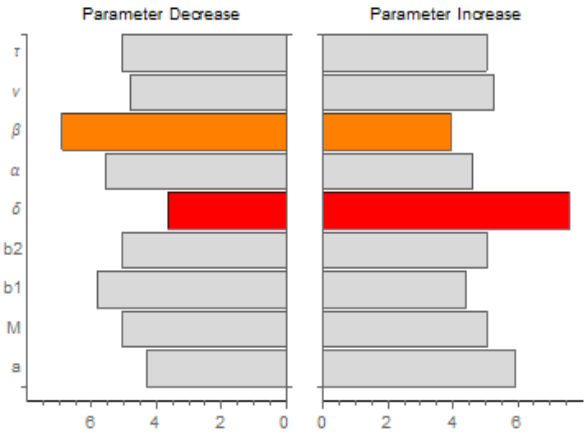


Figure 9. Cancer cell concentration $\times 10^6$ at $t = 50$ days

Monte Carlo simulations

Next we perform a sensitivity analysis using Monte Carlo sampling from an uniform distribution to establish the most influential parameter for tumor cell concentration. Each parameter was perturbed by 5% using random samples of size 10^3 and tumor cell concentration was computed at $t = 50$ days. The results for the mean, median, standard deviation, coefficient of variation (CV), and the mean of the lowest 100 values of the obtained cancer cells concentration values are shown in Table 2. An illustration showing the CV and the mean of the

lowest 100 cancer cells concentration values for each parameter from system of Equations 1 is shown in Figure 10. This figure suggests that again δ followed by β provide the largest CV cancer cells concentration values and the lowest tumor size.

Table 2. Monte Carlo Simulation Results

Values in this table represent cancer cells concentration values for the respective parameter changes					
Parameter	Mean	Median	Standard Deviation	Coefficient of Variation (CV)	Mean of the lowest
a	5×10^6	5×10^6	466779	0.0924	4.4×10^6
M	5×10^6	5×10^6	3676.34	0.0007	5×10^6
b_1	5×10^6	5×10^6	414367	0.0817	4.47×10^6
τ	5×10^6	5×10^6	5124.29	0.0010	5×10^6
α	5×10^6	5×10^6	281224	0.056	4.6×10^6
β	5.16×10^6	5×10^6	831858	0.1614	4×10^6
ν	5×10^6	5×10^6	136419	0.027	4.8×10^6
b_2	5×10^6	5×10^6	1894.81	0.0004	5×10^6
δ	5.2×10^6	5×10^6	1.1×10^6	0.2187	3.74×10^6

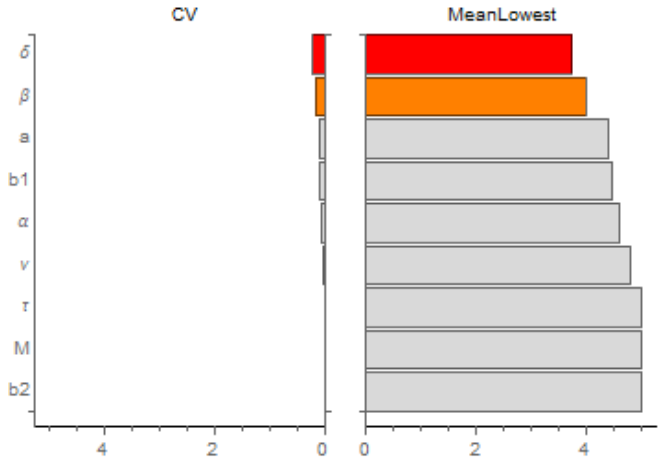


Figure 10. CV and mean (lowest cancer cell concentration)

Discussion

We analyze the dynamic of cancer during immunotherapy using a two-population model, that includes the cancer and the effector cells concentration. We propose that our model includes a logistic type rate-limiting term with Hutchinson's type of delay for the effector cells and a dynamic carrying capacity. The drawback of such

model is that it is difficult to analyze its steady states even on a nondimensionalized version of the original model, since the tumor free steady state cannot be found analytically due to the presence of a singularity at this point. Using parameter values from experimental literature, our model suggests that for the treatment to be effective effector cell stimulation needs to occur at a frequency in the order of around 2τ to 3τ days. Our double sensitivity analysis suggests that the parameters δ (death rate of the effector cells) followed by β (creation rate of the effector cells) are the most influential parameters. Linear analysis suggests that tumor free equilibrium can be achieved according to influential parameters.

Conclusion

Since the lowest amount of cancer cell concentration can be achieved through decreased effector cells death rate and increased arrival of these effector cells, we assume that our system suggests that under these conditions immuno treatments are effective for any type of cancer cells (regardless their mitosis rate and carrying capacity).

Acknowledgements

The corresponding author is grateful to Krista Lotesto and to all her students from the Mathematical Modeling for Cancer Risk Assessment course for their active input to all class discussions related to this topic.

References

The American Cancer Society medical and editorial content team. (2019, December 27). *How Immunotherapy is Used to Treat Cancer*. Cancer. https://www.cancer.org/treatment/treatments-and-side-effects/treatment-types/immunotherapy/what-is-immunotherapy.html#written_by

The Editors of the Encyclopaedia Britannica. (2019, October 22). *Effector cell*. Britannica. https://www.britannica.com/science/effector-cell

Institute for Quality and Efficiency in Health Care (IQWiG). (2020, July 30). *The innate and adaptive immune systems*. NCBI. https://www.ncbi.nlm.nih.gov/books/NBK279396/

Kuznetsov, V. A., Makalkin, I. A. & Taylor, M. A., Perelson, A. S. (1994). Nonlinear Dynamics of Immunogenic Tumors: Parameter Estimation and Global Bifurcation Analysis. *Bulletin of Mathematical Biology*, *56*(2), 295-321.

Kuznetsov, V. A., & Knott, G. D. (2001). Modeling Tumor Regrowth and Immunotherapy. *Mathematical and Computer Modelling*, *33*, 1275-1287.

Mahlbacher, G. E., Reihmer, K. C., & Frieboes, H. B. (2019). Mathematical Modeling of Tumor-Immune Cell Interactions. *J. Theor Biol.* 469, 47-60.

Murray, J. D. (2002). *Mathematical Biology I: An Introduction* (3rd ed.). S. S. Antman, L. Sirovich, J.E.

Marsden, S. Wiggins (Eds.). Springer-Verlag Berlin Heidelberg.

de Pillis, L. G., Radunskaya, A. E. (2014). Modeling Tumor-Immune Dynamics. In A. Eladdadi et al (Eds.), *Mathematical Models of Tumor-Immune System Dynamics* (pp. 59-108). New York, NY: Springer Science+Business Media.

Ruan, S. (2006). Delay Differential Equations in Single Species Dynamics. O. Arino et al. (Eds.), *Delay Differential Equations and Applications* (pp. 477-517), Springer, Berlin.

Tarfulea, N. E. (2018). A Mathematical Model of HIV Infection with Cellular and Immune Delays. *Appl. Math. Inf. Sci. 12*(5), 917-921.

Wodarz, D., Komarova, N. L. (2014). Two-species competition dynamics. *DYNAMICS of CANCER Mathematical Foundations of Oncology* (pp. 47-56). Singapore. World Scientific Publishing Co. Pte. Ltd.

Comparative Analysis of Cell-Based Products Obtained with Different Systems for Isolation of Stromal Vascular Fraction from Human Adipose Tissue

Elena Kostromina
National Medical Research Center for Rehabilitation and Balneology, Russian Federation,
http://orcid.org/0000-0002-9728-7938

Petr Eremin
National Medical Research Center for Rehabilitation and Balneology, Russian Federation,
https://orcid.org/0000-0001-8832-8470

Alexey Veremeev
Gamaleya National Research Center of Epidemiology and Microbiology, Russian Federation,
https://orcid.org/0000-0001-9946-1015

Maria Putova
Loginov Moscow Clinical Scientific Center, Russian Federation,
https://orcid.org/0000-0002-0825-7148

Irina Kudryashova
National Medical Research Center for Rehabilitation and Balneology, Russian Federation,
https://orcid.org/0000-0003-0261-7955

Ilmira Gilmutdinova
National Medical Research Center for Rehabilitation and Balneology, Russian Federation,
http://orcid.org/0000-0001-6743-2615

Abstract: For the wide application of adipose tissue-derived cell products in clinical practice, it is important to develop and implement new devices to automate and standardize the procedure of stromal vascular fraction (SVF) isolation. This work aimed to comparatively assess cell products obtained from human adipose tissue with two different systems based on mechanical and enzymatic isolation methods. A recently developed "ESVIEF System" (JoinTechCell LLC, RF) was used for enzymatic isolation of SVF. In the mechanical method, the SVF was obtained from adipose tissue by a double syringe system. Adipose tissue specimens were

Kostromina E., Eremin P., Veremeev A., Putova M., Kudryashova I., & Gilmutdinova I. (2021). Comparative analysis of cell-based products obtained with different systems for isolation of stromal vascular fraction from human adipose tissue. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 36-42), New York, USA. ISTES Organization.

obtained from 14 healthy donors undergoing a liposuction procedure in a plastic surgery clinic. The use of these two systems showed different results in terms of the yield, viability, and phenotype of the nucleated cells. Compared with the enzymatic method, the mechanical system produced products with a higher volume of residual oil, number of destroyed cells, and connective tissue particles. The cell product obtained by the enzymatic method was characterized by a higher yield of nucleated cells/ml of fat compared with the mechanical system as well as higher viability of the cells. Therefore, our data demonstrate that the method of adipose tissue treatment has a significant impact on the characteristics of the SVF cells obtained.

Keywords: Adipose tissue, Isolation systems, Stromal vascular fraction, Stem cell technology, Regenerative medicine

Introduction

Over the last decade, the use of cell-based therapy has become widespread in rehabilitation and regenerative medicine (McKay et al., 2019; Zakrzewski et al., 2019). Adipose tissue-derived stromal vascular fraction (SVF) is effective in treating a wide range of diseases, including cardiovascular and lung diseases, disorders of the musculoskeletal system, neurodegenerative diseases, and many others (Bora & Majumdar, 2017; Veremeev et al., 2016). For the wide application of cell-based products in clinical practice, it is important to develop and implement new devices to automate and standardize the procedure for the isolation of the SVF.

For a long time, the isolation of the SVF cells from adipose tissue (AT) has been performed using the enzymatic manual method developed by Zuk and coauthors (Zhu et al., 2013; Zuk et al., 2001). Despite the widespread use of manual methods, they are characterized by high time and organizational costs, as well as possible influence of human factor on the results, which is considered a critical problem for the implementation of GMP (good manufacturing practice) and GTP (good tissue practice) standards. In this regard, technologies for the isolation of the SVF from AT are developing towards the automation of the procedure and standardization of protocols using automated/semi-automated and/or closed isolation systems (SundarRaj et al., 2015). One such device is the recently developed by JoinTechCell LLC (RF) semi-automatic “ESVIEF System”, based on the enzymatic method of SVF cell isolation.

In addition to obtaining cell products using enzymes, technologies based on the mechanical processing of AT have also found applications (Condé-Green et al., 2016). However, the use of different methods of fat processing can lead to a variable yield of cells in the final product. The work aimed to perform a comparative evaluation of cell products obtained from human AT using two different systems based on mechanical and enzymatic isolation methods.

Materials and Methods

Adipose tissue specimens were obtained from 14 healthy donors (women aged 25-49 years), who underwent liposuction surgery at the Clinic of Plastic Surgery. Biomaterial donation was carried out with written informed consent of the patients. Lipoaspirate samples were harvested by syringe liposuction in the area of the anterior and lateral abdominal walls. The clinic established standard inclusion and exclusion criteria for liposuction. The obtained biomaterial was transported in a sterile thermal container and processed within 4 hours after the liposuction procedure. All tissue specimens were processed by both methods (mechanical and enzymatic) for all comparative studies of yield, viability, phenotype of the nucleated cells, and other parameters.

Characteristics of Two Systems used for the Isolation of SVF from Adipose Tissue

In the mechanical method, a double syringe system (Arthrex GmbH, Germany) was used to obtain the cell product from AT. The larger syringe (volume of 15 ml) is used for taking in lipoaspirate, the smaller syringe (volume of 6 ml) is intended for withdrawing fractions from the larger syringe after centrifugation. Two adapters with mesh diameters of 2.4 μm and 1.4 μm are used to disperse AT and obtain cell fractions. The maximum volume of fat tissue that can be processed at once is 15 ml.

Using the system for mechanical isolation (SM), SVF cells were obtained from AT as previously described (Stevens, 2018). Lipoaspirate was transferred into a double-chamber syringe using an adapter and centrifuged at 2500 rpm for 4 min, followed by separation into an aqueous fraction (lower layer), AT (middle layer), and oil (upper layer). An inner syringe was used to extract the upper layer (oil), and the aqueous fraction was removed through the syringe inlet. Fat tissue was then mechanically processed (dispersed) by pumping tissue from one syringe into another through a metal mesh adapter with a mesh diameter of 2.4 μm 5 times, then through an adapter with a mesh diameter of 1.4 μm 30 times. After subsequent centrifugation (4 minutes at 2500 rpm), the upper fraction (residual oil) was collected using the inner syringe, and a cell fraction of 2 ml was taken from the outer syringe.

The system for enzymatic isolation (SE) of the SVF, consists of a separator - a disposable sealed transparent plastic cylindrical container with two chambers located one above the other and separated by a nylon filter with a pore size of 100 μm, and an enzyme, Collagenase NB6 GMP Grade (Nordmark Arzneimittel GmbH & Co. KG, Germany). There are six labeled isolated channels on the sidewalls of the container with connectors for component input and output. The maximum amount of AT to be treated at once is 200 ml, the minimum volume is at least 50 ml.

Isolation of the SVF using the SE system was performed according to the manufacture's protocol. The procedure of adipose tissue processing included the following steps: adding lipoaspirate (50-120 ml) to the separator through channels 1-4, washing it twice with Hartmann's solution and centrifuging at 300 g for 8 min,

enzymatic treatment of AT with a collagenase solution in a 1: 1 ratio. (30 minutes at 37°C and stirring speed 100-120 rpm). After 2-fold washing and centrifugation for 10 min at 300 g, a cell suspension was taken from the lower chamber of the separator through the channel marked "SVF". The volume of the obtained SVF was 5 ml due to the design features of the separator.

Preparation of Cytology Samples

Thin-layer cytological samples were prepared from fixed cell material by the ThinPrep® liquid cytology method (Hoda et al., 2017). The samples were prepared according to the "liquid" method recommended by the manufacturer using a Cyto-Tek and Thinprep cytocentrifuge (Hologic). Fixed cell preparations were stained using the Papanicolaou staining technique (Gill, 2013).

Primary Culture of Adipose-derived Stem Cells (ASC)

Cell culture procedures were carried out accordingly to the standard protocol for culturing multipotent mesenchymal stromal cells (Freshney, 2015). Cells were plated at a density of 105 cells/cm2 on a T25 culture flask (Corning, USA) and cultured in DMEM with low glucose supplemented with 10% FBS, 2 mM L-glutamine, 100 units/ml penicillin, and 100 µg/ml streptomycin (StemCellTechnology, USA). The study of the regenerative potential was carried out on the 3rd passaging of ASC culture.

To assess the colony-forming efficiency, ASCs were seeded at a clonal density of 1×104 cells on the surface of 35 mm diameter Petri dishes and cultured for 14 days in a growth medium with elevated serum content (20% FBS) according to the standard technique. The number of colonies (> 50 cells), their types were determined as described elsewhere (Freshney, 2015).

Microscopy

Cells were visualised using a Leica DM IL LED inverted microscope (Leica Microsystems, Germany). Assessment of cell count and viability were performed as previously described (Gilmutdinova et al., 2021) using a hemocytometer (Hausser Scientific, Horshman, USA) and 0.1% trypan blue solution (Life Technologies, Eugene, USA). Thin-layer cytological preparations were examined under a Leica DM1000 microscope (Leica Microsystems, Germany).

Statistical Analysis

The data are presented as means ± SD. Statistical analysis was carried out using Student's t-test. The level of statistical significance was set at $P < 0.05$.

Results and Discussion

Two systems for isolating the SVF from human AT based on mechanical and enzymatic extraction methods were tested. To compare the efficiency of these two systems, the obtained cell products were evaluated according to several parameters: the number of nucleated cells in the preparation, their viability, phenotype, and other parameters. The results showed that the method of adipose tissue treatment (enzymatic or mechanical) has a significant impact on the characteristics of the obtained SVF cells (Table 1).

Compared with the enzymatic method, the mechanical system produced products with a higher relative volume of residual oil. Microscopic examination showed a significant difference in the number of isolated cells/ml of AT/lipoaspirate as well as cell viability when using the SM and SE systems. It should be noted that destroyed cells and connective tissue particles were detected in the samples obtained from AT using the SM system.

Table 1. Comparative Characterization of Cell Products obtained with Two Systems Utilizing Enzymatic (SE) and Mechanical (SM) Isolation of SVF from Adipose Tissue

Parameters	SE	SM
Total volume of residual oil, %	18.21±7.10	25.42±14.40
Total number of nucleated cells/1 ml of fat ($x10^6$)	0.96±0,45	0.21±0.20
Cell viability, %	86.25±4.15	35.50±18.30
Total number of cells/1 ml of lipoaspirate ($x10^6$)	0.71±0,32	0.13±0,11
Number of specimens	14	12

The results of the cytological examination showed the presence of fibroblasts, endothelial cells, single lipophages, and adipocytes in all preparations obtained from AT using both systems (Fig. 1).

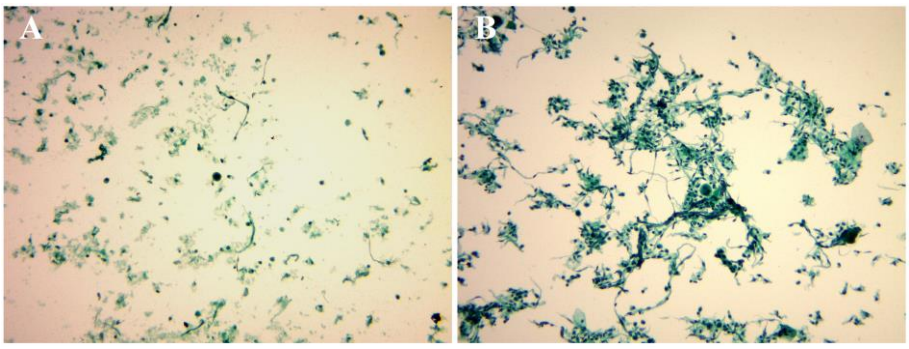


Figure 1. Cytological Staining of Adipose Tissue-derived SVF obtained with the SM (A) and SE (B) Systems. Papanicolaou staining. x100.

Compared with the enzymatic method, the percentage of ASCs in the cell product obtained with the SM system was very low. In the enzymatic method, the study of the ASC colony formation efficiency showed a predominance of dense colonies (55.5±3.5%) over the proportion of mixed (29.2±3.8%) and diffuse colonies

(15.3±2.7%), which also indicated a high regenerative potential of this cell culture.

In mechanical isolation, there was a problem in some cases with the transfer of AT to the double syringe system. As fat tissue can aggregate into small conglomerates, this made it very difficult to transfer the lipoaspirate into the syringe according to the manufacturer's method (AT could not pass through the adapter with a mesh diameter of 2.4 μm). In these cases, it was necessary to transfer the AT directly through the removed syringe piston. Therefore, the use of two systems utilizing mechanical and enzymatic procedures for isolating SVF from AT showed different results in terms of the yield, viability, proliferation potential, and phenotype of the SVF cells.

Conclusion

The method of adipose tissue processing (enzymatic or mechanical) has a significant impact on the characteristics of the SVF obtained. When isolated mechanically, the obtained cell product is characterized by a low yield of nucleated cells, large amounts of residual oil, and degraded connective tissue in the final product. In enzymatic isolation, the cell product is characterized by a high yield of nucleated cells similar to the manual enzymatic method, and a high proliferative potential of ASC.

References

Bora, P., & Majumdar, A. S. (2017). Adipose tissue-derived stromal vascular fraction in regenerative medicine: a brief review on biology and translation. *Stem Cell Res Ther*, *8*(1), 145. https://doi.org/10.1186/s13287-017-0598-y

Condé-Green, A., Kotamarti, V. S., Sherman, L. S., Keith, J. D., Lee, E. S., Granick, M. S., & Rameshwar, P. (2016). Shift toward Mechanical Isolation of Adipose-derived Stromal Vascular Fraction: Review of Upcoming Techniques. *Plast Reconstr Surg Glob Open*, *4*(9), e1017. https://doi.org/10.1097/gox.0000000000001017

Freshney, R. I. (2015). *Culture of animal cells: a manual of basic technique and specialized applications*. John Wiley & Sons.

Gill, G. W. (2013). Papanicolaou Stain. In *Cytopreparation: Principles & Practice* (pp. 143-189). Springer New York. https://doi.org/10.1007/978-1-4614-4933-1_10

Gilmutdinova, I. R., Kostromina, E., Yakupova, R. D., & Eremin, P. S. (2021). Development of nanostructured bioplastic material for wound healing. *European Journal of Translational Myology*. https://doi.org/10.4081/ejtm.2020.9388

Hoda, R. S., VandenBussche, C., & Hoda, S. A. (2017). Liquid-Based Specimen Collection, Preparation, and Morphology. In R. S. Hoda, C. VandenBussche, & S. A. Hoda (Eds.), *Diagnostic Liquid-Based Cytology* (pp. 1-12). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-53905-7_1

McKay, J., Frantzen, K., Vercruyssen, N., Hafsi, K., Opitz, T., Davis, A., & Murrell, W. (2019). Rehabilitation following regenerative medicine treatment for knee osteoarthritis-current concept review. *J Clin Orthop Trauma*, *10*(1), 59-66. https://doi.org/10.1016/j.jcot.2018.10.018

Stevens, H. (2018). ACA-Technik: „stromal vascular fraction“, „platelet-rich plasma“ und Mikrofett zur körpereigenen Regeneration und Hautverjüngung. *Journal für Ästhetische Chirurgie*, *12*. https://doi.org/10.1007/s12631-018-0151-6

SundarRaj, S., Deshmukh, A., Priya, N., Krishnan, V. S., Cherat, M., & Majumdar, A. S. (2015). Development of a System and Method for Automated Isolation of Stromal Vascular Fraction from Adipose Tissue Lipoaspirate. *Stem Cells Int*, *2015*, 109353. https://doi.org/10.1155/2015/109353

Veremeev, A., Bolgarin, R., Petkova, M., Katz, N., & Nesterenko, V. (2016). Adipose derived stromal vascular fraction as an alternative source of cells for the regenerative medicine. *Genes & Cells*, *11*(1), 35-42.

Zakrzewski, W., Dobrzyński, M., Szymonowicz, M., & Rybak, Z. (2019). Stem cells: past, present, and future. *Stem Cell Res Ther*, *10*(1), 68. https://doi.org/10.1186/s13287-019-1165-5

Zhu, M., Heydarkhan-Hagvall, S., Hedrick, M., Benhaim, P., & Zuk, P. (2013). Manual isolation of adipose-derived stem cells from human lipoaspirates. *J Vis Exp*(79), e50585. https://doi.org/10.3791/50585

Zuk, P. A., Zhu, M., Mizuno, H., Huang, J., Futrell, J. W., Katz, A. J., Benhaim, P., Lorenz, H. P., & Hedrick, M. H. (2001). Multilineage Cells from Human Adipose Tissue: Implications for Cell-Based Therapies. *Tissue Engineering*, *7*(2), 211-228. https://doi.org/10.1089/107632701300062859

Copper (II) Bromide as an Efficient Catalyst for the Selective Protection and Deprotection of Alcohols

Roufia Mezaache
Université de Batna, Algeria

Yénimégué Albert Dembelé
Université de Bamako, Mali

Yann Bikard
Université de Strasbourg, France

Jean-Marc Weibel
Université de Strasbourg, France

Aurélien Blanc
Université de Strasbourg, France

Patrick Pale
Université de Strasbourg, France

Abstract: The synthesis of highly functionalized molecules usually requires several steps dealing with the protection and deprotection of those functional groups [1, 2]. The choice of protecting groups is often critical for synthesis success, especially for the total synthesis of complex natural products and analogs [2, 3]. Benzyl type protecting groups are among the most commonly used, due to their deprotection conditions orthogonal to other protecting and functional groups [1-3], and they have been applied to the protection of alcohols, thiols, amines, and acids [1, 2]. Nevertheless, their introduction is not always simple due to the basic or acid condition required [2], in order to solve this problem, we recently described a convenient and efficient method based on Copper bromide catalyst (CuBr2) has been developed for the protection of Primary and secondary alcohols with bis(4-methoxyphenyl) méthanol (BMPMOH) in good yield using CuBr2 as catalyst in acetonitrile at room temperature [4]. Deprotection could easily be achieved using the same catalyst but in ethanol. Both Cu-catalyzed protection and deprotection were orthogonal to other methods and fully compatible with other functional groups. The mildness of these protection and deprotection methods as well as their selectivity render them very useful tools for total synthesis.

Mezaache, R., Dembelé, Y. A., Bikard, Y., Weibel, J. M., Blanc, A., & Pale, P. (2021). Copper (II) bromide as an efficient catalyst for the selective protection and deprotection of alcohols. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 43-44), New York, USA. ISTES Organization.

Keywords: alcohols, ethers, protection, déprotection, BMPMOH, CuBr2, CH3CN.

References

1. Greene, T. W.; Wuts, P. G. M. *Protective Groups in Organic Synthesis*, 3rd ed.; J. Wiley & Sons: New York, NY, 1999.
2. Kocienski, P. J. *Protecting groups*, 3rd ed .; G. Thieme: Stuttgart, New York, NY, 2004.
3. Nicolaou, K. C.; Snyder, S. A *Classics in Total Synthesis* II; Wiley-VCH: Weinheim, 2003.
4. Mezaache, R. ; Dembelé, Y. A. ; Bikard, Y. ; Weibel, J.M. ; Blanc, A. ; Pale, P. *Tetrahedron Letters*, 2009, 50, 7322–7326.

Sorting Real Numbers into a Linked List on the PRAM Model

Yijie Han
University of Missouri at Kansas City, USA

Pruthvi Kasani
University of Missouri at Kansas City, USA

Abstract: We study the sorting of real numbers into a linked list on the PRAM (Parallel Random-Access Machine) model. There are EREW (Exclusive Read Exclusive Write) PRAM, the CREW (Concurrent Read Exclusive Write) PRAM and the CRCW (Concurrent Read Concurrent Write) PRAM. In CRCW PRAM, multiple processors can read or write on a memory cell in a step. Since CRCW allows multiple processors to read or write on a single memory cell, there are some arbitrary schemes designed to perform the actions. Hence, there are Common CRCW, Priority CRCW and Arbitrary CRCW PRAM Models. The algorithm proposed in the paper runs on the Common CRCW PRAM Model and thus run on the Arbitrary and Priority CRCW Model. We show that n real numbers can be sorted into a linked list in constant time using n^2 processors. Previously, n numbers can be sorted into a linked list using n^2 processors in O(loglogn) time.

Keywords: parallel algorithms, parallel sorting, CRCW (Concurrent Read Concurrent Write), EREW (Exclusive Read Exclusive Write), CREW (Concurrent Read Exclusive Write).

Introduction

In this paper we study parallel sorting of real numbers into a linked list. The computation model we used for our algorithm is the PRAM (Parallel Random Access Machine) [9]. There are EREW (Exclusive Read Exclusive Write) PRAM, the CREW (Concurrent Read Exclusive Write) PRAM and the CRCW (Concurrent Read Concurrent Write) PRAM [9]. Each of these sub models differentiates themselves on how the memory is shared among all the processors in PRAM. On the EREW PRAM, at any step no more than one processor can either read or write on a memory cell. On the CREW PRAM, one or more processors can simultaneously read a memory cells in a step but no more than one processor can write a memory cell in a step. Whereas in CRCW PRAM, multiple processors can read or write on a memory cell in a step. Since CRCW allows multiple processors to read or write on a single memory cell, there are some arbitrary schemes designed to perform the actions. On the Priority CRCW PRAM, the processor having the highest priority wins the write on the memory cell among the processors writing to the memory cell. The priority can be the index of the processor. On the Arbitrary CRCW PRAM an arbitrary processor wins among the processors to write on the memory cell. On the

Han, Y. & Kasani, P. (2021). Sorting real numbers into a linked list on the PRAM model. In R. Thripp & I. Sahin (Eds.), *Proceedings of ILSET 2021-- International Conference on Life Sciences, Engineering and Technology* (pp. 45-49), New York, USA. ISTES Organization.

Common CRCW PRAM when multiple processors write the same memory cell in one step, they must write the same value and that value is written into the memory cell. Among all the CRCW PRAM, Priority CRCW is the strongest model, Arbitrary CRCW PRAM is weaker than the Priority CRCW PRAM, and Common CRCW PRAM is the weakest among the three. In this paper we will use the Common CRCW PRAM. Because our algorithm runs on the Common CRCW PRAM and thus they can run on the Arbitrary and Priority CRCW PRAM.

Let Tp be the time complexity of a parallel algorithm using p processors. Let T1 be the time complexity of the best serial algorithm for the same problem. Then pTp ≥T1. When pTp=T1 then this parallel algorithm is an optimal parallel algorithm.

When we have a TP time algorithm using P processors, then when we use p processors the time can be expressed or translated as TPP/p+TP.

A parallel algorithm for a problem of size n using polynomial number processors (i.e., nc processors for a constant c) and running in polylog time (i.e., O(logcn) time for a constant c) is regarded as belong to the NC class [3], where NC is Nick's class.

Researchers in parallel algorithm field are working to achieve NC algorithms and fast and efficient parallel algorithms.

In this paper, we will study sorting real numbers into a linked list in constant time using n2 processors. Previously it is known that n real numbers can be sorted into a linked list in O(loglogn) (constant time) using n2 (n3) processors [5,7,8].

It is known that sorting n real numbers into an array takes at least Ω(logn/loglogn) time on the CRCW PRAM with polynomial number of processors [1]. It takes at least Ω(loglogn) time if we are to sort them into a padded array [4]. However, if we are going to sort them into a linked list, we show here that it can be done in constant time. Thus, the lower bound of Ω(logn/loglogn) [1] and the lower bound of Ω(loglogn) [4] are really the lower bound for arranging numbers in an array instead of the lower bound of "sorting" them.

There are results before for sorting integers into a linked list [2, 6]. It is known there that n integers in {0, 1, …, m-1} can be sorted into a linked list in constant time using nlogm processors. m here cannot be bounded by functions of n. Except our earlier results for sorting real numbers into a linked list [6,10,11] we do not know other results for parallel sorting real numbers into a linked list and we do not know earlier results of sorting real numbers in constant time.

Sorting Real Numbers into a Linked List

We assume that the n input real numbers are distinct. This can be achieved by replacing every real number a by a pair (a, i) where i is the index of the number a in the input array.

Firstly, let us discuss about the algorithm on how to sort the real numbers in linked list using constant time using n3 processors. Let us say, A[0.....n-1] be the input array of n real numbers and we have n3 processors to achieve constant time.

Assign n processors to each element of the array to compare it with the other elements in the array. It will write as 1 for the elements that it greater than the given element and 0 for the elements if it is less than it. For example, we have the given input array elements as 4,2,5,1,6,3,9. Let us pick an element 5 from the array. As said above, it marks 1 to the elements greater than 5 and 0 for the ones lesser than 5. So, the output is 0,0,0,0,1,0,1. We use the n2 processors to the elements marked as 1 and find the smallest number among them (i.e., 6) in constant time [10,11] and link it to the element 5. So, here we have 6 and 9 out of which 6 is the minimum. So, 6 is linked to 5. This process is executed in parallel to all the elements in the array, and we get the final sorted linked list of elements. This algorithm can be done in constant time using n3 processors.

Now, let us show the algorithm on sorting the real numbers into a linked list using n2 processors in O(loglogn) time on the Common CRCW PRAM. This algorithm is like the above algorithm where we assign n processors to compare a number to the rest of the elements in the array. Now, we need to compute the minimum of n numbers using n processors. This can be done in O(loglogn) time [10,11]. Let us say A[0....n-1] be the input array of n real numbers. As above, the comparison task of comparing one element A[i] to other elements takes constant time. Now, we need to find the minimum of elements in A that are larger than A[i]. Let us say m is the minimum element. Now, for each element in A[i] we will copy it into a new array Ai. This usually takes constant time. We now compare A[i] with every element Ai[j] in Ai. If A[i] ≥ Ai[j] then we will do Ai[j] =MAX. Then we will find the minimum element Ai[k] in Ai. This takes constant time using n1+ε processors (or O(loglogn) time with n processors) for Ai [10.11]. For all i=0, 1… n-1, this takes constant time with n2+ε processors (or O(loglogn) time with n2 processors). If Ai[k] is the smallest element larger than A[i]. Thus, we can make a link from A[k] to A [i].

Now we show our new algorithm which allows to sort n real numbers into a linked list in constant time with n2 processors. We divide the input numbers into groups. So, now each group has n1/2 numbers. Assign n3/2 processors for each group. So now the total number of processors to do this will be n1/2 x n3/2 processors which is n2 processors. We already know that building a sorted linked list with n3/2 processors of n1/2 numbers takes constant time. Now we have groups with sorted linked lists. Since we have n1/2 groups there will be O(n) pairs of groups in total. Let us assign n processors for every pair of groups. So, we require n processors x O(n) pairs which is O(n2) processors total. So, for every number in a group, we can use n1/2 processors. So, we

require O(n) processors for each group. Now, let us say we have a number A in group 1. It finds the smallest number B larger than it in group 2 by comparing with every number in group 2 and using the sorted linked list already built for group 2. This process is repeated for all the pairs of groups like group 1, group 3 and group 1, group 4 etc. We find smallest numbers larger than A. In general, if we do it in parallel each number find n1/2 smallest numbers larger than it. Each number then uses n processors to find the minimum among these n1/2 smallest numbers in constant time [10,11]. So, in total the proposed algorithm uses n2 processors to sort the n real numbers in a linked list in constant time.

Theorem

Theorem 1. n real numbers can be sorted into a linked list in constant time using n2 processors on the Common CRCW PRAM.

We have been able to optimize the existing algorithms with less number processors and time. Earlier, we had algorithms like sorting of n real numbers into a linked list in constant time using n^3 processors and sorting of n real numbers into a linked list in O(loglogn) time using n^2 processors [5, 7, 8].

Conclusions

We discussed about sorting n real numbers into a linked list using n2 processors in constant time. This algorithm is more effective than the ones that require n3 processors to sort into a linked list in constant time and n2 processors to sort into a linked list in O(loglogn) time. We have followed the approach to assign the processors by dividing the given input numbers into groups. The most interesting part of this algorithm is that we were able to sort the n real numbers into the linked list by decreasing the number of processors from n3 to n2 and by achieving this in constant time.

It looks to us that reducing the number of processors further while still achieving constant time is not trivial. A plausible way of doing this is to convert real numbers into integers while utilizing advantage integers bring to sorting. We have not been achieved further results along this direction.

References

[1]. P. Beame, J. Hastad. Optimal bounds for decision problems on the CRCW PRAM. Proc. 1987 ACM Symp. On Theory of Computing (STOC'1987), 83-93(1987).

[2]. P.C.P. Bhatt, K. Diks, T. Hagerup, V.C. Prasad, T. Radzik, S. Saxena. Improved deterministic parallel integer sorting. Information and Computation, 94, 29-47(1991).

[3]. S. A. Cook. Towards a complexity theory of synchronous parallel computation. L' Enseignement Mathématique, 27, 99-124(1981).

[4]. T. Goldberg, U. Zwick. Optimal deterministic approximate parallel prefix sums and their applications. Proc. 3rd. Israel Symp. On Theory and Computing Systems, 220-228(1995).

[5]. N. Goyal. An arbitrary CRCW PRAM algorithm for sorting integers into the linked list and chaining on a trie. Master's Thesis. University of Missouri at Kansas City. 2020.

[6]. T. Hagerup. Towards optimal parallel bucket sorting. Information and Computation. 75, 39-51(1987).

[7]. Y. Han, N. Goyal, H. Koganti. Sort integers into a linked list. Computer and Information Science. Vol. 13, No.1, 51-57(2020).

[8]. Y. Han, T. Sreevalli. Parallel merging and sorting on linked list. International Journal of Computer and Information Technology (IJCIT). Vol. 10, No. 2, 92-95(March 2021).

[9]. R. M. Karp, V. Ramachandran, Parallel algorithms for shared-memory machines. In Handbook of Theoretical Computer Science (Vol. A): Algorithms and Complexity, J. van Leeuwen, Ed., New York, NY: Elsevier, 869-941(1991).

[10]. C. P. Kruskal. Searching, merging, and sorting in parallel computation. IEEE Trans. Comput., C-32, 942-946(1983).

[11]. L. G. Valiant. Parallelism in comparison problems. SIAM J. on Computing, Vol. 4. No. 3, 348-355(1975).