ISBN 978-2-38236-154-2

LIVRE DE LYON

livredelyon.com
livredelyon
livredelyon
livredelyon

CURRENT ACADEMIC STUDIES IN SCIENCE AND MATHEMATICS SCIENCES

Current Academic Studies in SCIENCE AND MATHEMATICS SCIENCES

Editors
Prof. Dr. D. Esra Yıldız
Assoc. Prof. Dr. Esma Yıldız Özkan

Natural Sciences

LIVRE DE LYON
2021

CURRENT ACADEMIC STUDIES IN SCIENCE AND MATHEMATICS SCIENCES-II

EDITORS

PROF. DR. D. ESRA YILDIZ

ASSOC. PROF. DR. ESMA YILDIZ ÖZKAN

LIVRE DE LYON

Lyon 2021

Editors • Prof. Dr. D. Esra Yıldız • ORCID: 0000-0003-2212-199X
Assoc. Prof. Dr. Esma Yıldız Özkan • ORCID: 0000-0003-1283-833X
Cover Design • Mirajul Kayal
Layout • Mirajul Kayal
First Published • May 2021, Lyon

ISBN: 978-2-38236-154-2

Publisher • Livre de Lyon
Address • 37 rue marietton, 69009, Lyon France
website • http://www.livredelyon.com
e-mail • livredelyon@gmail.com

PREFACE

The natural purpose of the humanity is to make the life easier. The scientists serve to this purpose by studying unceasingly.

This book is based on current studies in science and mathematics and consists of seven sections with academic articles presented as a reflection of the scientific studies.

The editors would like to express their endless thanks to all the authors of the chapters contributed to the implementation of this book, to the referees who shared their valuable opinions and to the Livre de Lyon Publishing House which pioneered the publication of this book.

Prof. Dr. D. Esra Yıldız

Assoc. Prof. Dr. Esma Yıldız Özkan

CONTENTS

CHAPTER I

KLOPMAN INDEX: A NEW DESCRIPTOR USED IN MCET FOR 3D/4D-QSAR MODELING

Burçin Türkmenoğlu[1] & Yahya Güzel[2]

[1](Asst. Prof. Dr.), Erzincan Binali Yıldırım University,
e-mail: burcin.turkmenoglu@erzincan.edu.tr
ORCID: 0000-0002-5770-0847

[2](Prof. Dr.), Erciyes University, e-mail: yguzel@erciyes.edu.tr
ORCID: 0000-0002-4698-1866

1. Introduction

It is now widely accepted that the choice of descriptors is important in potential drug molecules, as it would indicate that they have some, if not all, of the pharmacophore (Segall, Beresford, Gola, Hawksley, & Tarbit, 2006). Considering pharmacophore properties in the drug discovery process can reduce drug development costs and help identify drug candidates easily (Obrezanova, Csanyi, Gola, & Segall, 2007). Modern demand for drug discovery due to pharmacophore has led to the trend of developing effective descriptors for better understanding of the interaction when new data are available (Cartmell, Enoch, Krstajic, & Leahy, 2005; Zhang, Golbraikh, Oloff, Kohn, & Tropsha, 2006). The purpose of such descriptors is to save scientists' time, explore the possibility of more realistic modeling, and create the "Quantitative Structure-Activity Relationship" (QSAR) model that even non-experts can easily understand (Obrezanova et al., 2007). Local Reactive Descriptor (LRD) such as electron density, electric field, and electrostatic potential have been used as various three-dimensional (3D) molecular similarity indexes. Accordingly, activities were estimated based on the overlapping of analog domains of all components in a 3D grid (Bender & Glen, 2004; Kubinyi, 1998a, 1998b; Willett, 2000).

In this review article, we discuss the application of a new LRD as the Klopman Index (KI) (Kizilcan, Turkmenoglu, & Guzel, 2020) to predict the interaction between the Ligand-Receptor (L-R) (Tokat, Turkmenoglu, Guzel, & Kizilcan, 2019). We determine the parameters at each interaction

point of the pharmacophore, which is the 3-dimensional geometric structure and electronic properties on the active side of the receptor, with the LRD values of the ligand.

Since KI shows both Coulomb and covalent interaction, which are "Hard and Soft Acids and Bases" (HSAB), they can give better results than other LRDs that provide a single interaction. For the pharmacophore of various compounds, we compare the statistical results obtained from the proposed 3D-QSAR model using the KI descriptor and other LRDs in our own homemade Molecular Conformer Electron Topological (MCET) software.

2. Overview of The Klopman Index

We introduce a new descriptor, the KI, for the QSAR modeling of pharmacophore properties (Klopman, 1968; Salem, 1968). This descriptor helps us to understand the interaction between L-R more comprehensively and realistically as it is based on the HSAB approach, which offers the most general interaction properties of atoms. Although the electronic properties of atoms as a LRD are included in many studies, KI (Guzel, Aslan, Turkmenoglu, & Su, 2018) has not yet been widely used in QSAR modeling (Obrezanova et al., 2007). In our previous studies, this descriptor was used as a new one to calculate activity in multiple data sets to create a 3D/4D-QSAR model (Guzel et al., 2018; Kizilcan et al., 2020; Tokat et al., 2019; Turkmenoglu & Guzel, 2018; Turkmenoglu, Guzel, Su, & Kizilcan, 2020). An advantage of our descriptor, unlike Hopfinger's work (Hopfinger et al., 1997), which explains the L-R interaction mechanism the type of atoms such as C, N, O or their interaction pharmacophore elements (IPEs i.e., atom types) any type, nonpolar, polar-positive charge, polar-negative charge, hydrogen bond acceptor, hydrogen bond donor and aromatic does not require the knowledge of different types such as (Patel et al., 2014) . KI can represent IPEs used in many studies as a single descriptor. KI reflects both the atomic charge and the LRD interaction with the quantum chemical properties of the atoms, belonging to atomic coefficients.

This descriptor solves the desired LRD problem in QSAR modeling, which serves the emergence of pharmacophore, with the following features:

•Its most important feature reflects both effects of HSAB simultaneously.

•It is suitable for modeling the interaction of atoms regardless of their type.

•Over-training is avoided by using coefficients in the HOMO of the receptor and the LUMO of the ligand (or vice versa) through simplified KI instead of using all orbitals of the ligand and the receptor.

•This descriptor has a natural ability to select key interaction points. It considers not only atomic charges at each point, but also the coefficients in the frontier orbitals.

KI has provided a better understanding of many chemical reactions whose mechanism cannot be explained. It was also used as one of the simplest and most effective interactions involving multiple orbitals. KI is sufficient to produce a model that addresses all the interaction properties of the pharmacophore, and since it does not require any subjective input from the user, this descriptor is perfect for (Obrezanova et al., 2007) demonstrating a true interaction.

A potential disadvantage of KI is that it composes pharmacophore models that are difficult to interpret. Because the pharmacophore has both Coulombic (ionic additive) and orbital (covalent additive) interactions at each point of the pharmacophore, it is difficult to infer the effect of the descriptor on the observed activity or property (Obrezanova et al., 2007). This difficulty depends on the nature of modeling electronic values that are not identical at each point of interaction, since the electrostatic and covalent interaction values at these points can contribute opposite (attraction or repulsion) and different (more or less) to each other. The contribution of both interactions at a point is different from ligand to ligand. For each point of interaction on the receptor side, the parameters for both types of interaction must be calculated. So, another disadvantage in using KI is that it can be computationally costly. Calculation of parameters for both types of interactions at a point involves two M-dimensional vectors of the electronic values in the respective atoms of the ligands using two types of NxM-dimensional matrix, where N is a set of compounds in the training set (Obrezanova et al., 2007) and M is the number of pharmacophore points. Such a matrix solution is provided by the Levenberg-Marquardt algorithm developed in MCET (Kizilcan et al., 2020; Turkmenoglu & Guzel, 2018).

In this review, we tried to explain how the interaction energy between L-R is calculated by using KI for regression problems by referring to our previous articles. This energy is made by determining the parameters of the receptor side by taking part in the model with a non-linear contribution (Guzel et al., 2018). We consider the regression results according to the function optimized with the global and local gradient descent optimization techniques in the Levenberg-Marquardt algorithm. In a new technique we use for the HSAB interaction at each point, we determine the values of two parameters belonging to both atomic coefficients and atomic charges in the Levenberg-Marquardt algorithm with a nested loop.

It is necessary to align all molecules on the common core structure to make the molecules geometrically and electronically fit according to the receptor

(Kizilcan et al., 2020). A conformer that best coincides with all atoms of the template is chosen as representative of other conformers in its molecule. The maximum tolerance value coincides with the bond length and atoms in similar positions are collected in the same cluster. A clustering method attempts to define relationships and suggest a categorization of data without the need to predefine categories. Many clustering schemes have been proposed, and some have even been shown to allow structural research to be accelerated (Jorgensen, Groves, & Hammer, 2017; Sorensen, Jorgensen, Bruix, & Hammer, 2018). Subsets created in large data sets are handled stochastically with Genetic Algorithm (GA). One of the best of several stochastic data subsets is determined by the best rank in statistical results. There is a best statistical result for each LRD, but among them it is possible to choose the descriptor type with the most perfect. In addition to KI, which is one of the descriptor types that create various 3D-QSAR models, other LRDs are used. To show the best descriptor, the results of q^2 and r^2 obtained from the two data sets of training and test sets were compared, respectively.

3. Method

3.1. Klopman Index (KI)

KI is an effective descriptor in L-R interaction as HSAB and is used as a new LRD. The results of the models created with KI, which have both atomic charge and coefficient properties of these descriptors, are compared with the atomic charges, HOMO/LUMO coefficients and Fukui index (Fukui, 1982), which are used as common LRD descriptors. While HOMO/LUMO approach and Fukui index lead to more covalent interaction, Mulliken, Electrostatic and Natural atomic charges lead to ionic interaction.

The Klopman-Salem equation is generally as follows (Fleming, 1976; Klopman, 1968; Salem, 1968):

$$\Delta E = (-\sum_{a,b}(q_a + q_b)\beta_{ab}S_{ab}) + \sum_{k<l}\left(\frac{Q_k Q_l}{\varepsilon R_{kl}}\right) + \left(\sum_{r}^{occ}\sum_{s}^{unocc} - \sum_{s}^{occ}\sum_{r}^{unocc} \frac{2\left(\sum_{a,b} c_{ra}\, c_{sb}\beta_{ab}\right)^2}{E_r - E_s}\right) \quad \textbf{Equation 1}$$

q_a: Electron population in the a-atomic orbitals,

β_{ab} and S_{ab}: Resonance and overlap integrals between a and b atomic orbitals,

Q_k: The total charge on the k-atom,

ε: Local dielectric constant,

R_{kl}: The distance between the nuclei of the k and l atoms,

c_{ra}: Coefficient of atomic orbital a in molecular orbital r,

E_r: Energy level of the r-orbital of the molecule.

The first term in Equation 1 describes the steric effect with the closed shell repulsion of electron-occupied molecular orbitals of two molecules. The second term is the ionic contribution of the electrostatically Coulombic repulsion or attraction of the corresponding atoms between two molecules. The third term is the covalent contribution between two molecular orbitals with all possible interactions, occupied and not. In the third term, orbitals with a close difference in Molecular Orbital (MO) energies in the denominator make a great contribution.

Equation 1 is arranged as in Equation 2 below, neglecting the first term, which contains all the polarity of the solvent and various dissolution and insoluble energies depending on the electron population in the atomic orbitals. This arrangement also includes each m interactions at a total of M interaction points between L-R.

$$\Delta E = \sum_{m=1}^{M} \left[\frac{Q_{nuc,m} Q_{elec,m}}{\varepsilon R_{(nuc-elec),m}} + \left(\sum_{r}^{occ} \sum_{s}^{unocc} - \sum_{s}^{occ} \sum_{r}^{unocc} \frac{2\left(\sum_{a,b} c_{ra,m}\, c_{sb,m} \beta_{ab,m}\right)^2}{E_r - E_s} \right) \right] \quad \textbf{Equation 2}$$

Ignoring the occupied-unoccupied interactions in Equation 2, Equation 3 has been simplified, such as the "simplified Klopman Index", mostly in the form of only the highest filled and lowest empty molecular orbitals (HOMO-LUMO interaction) of the reactants (Fleming, 1976; Klopman, 1968; Salem, 1968).

$$\Delta E = \sum_{m=1}^{M} \left[\frac{Q_{nuc,m} Q_{elec,m}}{\varepsilon R_{(nuc-elec),m}} + \frac{2\left(c_{nuc,m}\, c_{elec,m} \beta_m\right)^2}{E_{HOMO(nuc)} - E_{LUMO(elec)}} \right] \quad \textbf{Equation 3}$$

The relative contributions of the first (Coulombic term-ionic) and the second (frontier orbital term-covalent) characterize the bonding. Here m indicates the number of interactions at a point, while M indicates the total point of interaction. cm is the total coefficient of atomic orbital types at the corresponding point m in the frontier orbital. For HOMO-LUMO orbitals, the greater or lesser difference between MO energies highlights ionic or covalent contributions, respectively.

If the numerical results of both terms in Equation 3 are negative, they contribute to the attraction force. Therefore, it should be kept in mind that $Q_{nuc} < 0$ for the first term being negative and $E_{LUMO} > E_{HOMO}$ for the second term being negative. On the other hand, high energy HOMO and low energy LUMO serve to be soft nucleophile and soft electrophile, respectively, while the reverse is hard nucleophile and hard electrophile. This hard and soft electrophile (acid) and nucleophile (base) play an important role in explaining the HSAB chemical interactions (Fukui,

1982). There are hard-hard interactions governed by the ionic term and soft-soft interactions governed by the covalent term (Pearson, 1997).

3.2. Steps to Discover a Pharmacophore

In this section,

- **a)** Creating the core structure from the atoms of the selected template,
- **b)** Aligning core atoms with molecules and overlapping the remaining atoms,
- **c)** Determining the best one among the superimposed molecular conformer,
- **d)** Creating clusters from stacked atoms,
- **e)** Trying to discover pharmacophore in subsets by applying GA,
- **f)** Setting the parameter constants of the pharmacophore for the receptor side,
- **g)** To predict the best possible pharmacophore, result according to the selected LRD type.

3.2.a) The lowest conformer of the most active molecule in the least number of conformers in a series of molecules is chosen as the template. With the combination of atoms on the template, core structures are formed from their geometric and electronic properties. Core structures are formed by the combination of different numbers of atoms ranging from 3 to 6, containing function atoms. If a 6-atom structure of the core could form, then the 3-, 4- and 5-atom subgroups of the 6-atoms may not be considered. Especially since the possibility of having more than 6 common structures in the molecular set with different skeletons is reduced, more combinations are not needed, thus reducing the calculation process. There is at least one functional atom (such as O, N, S or C in C = O) within the core structures. Only one of the different core structures can be an initial part of the pharmacophore responsible for activity.

By means of the core structure, the ligand can initially approach the receptor with approximately the same electronic and geometric properties. By pulling the Cartesian coordinate values of the first three atoms in the core structures to the origin, the coordinate values of the remaining atoms can be rearranged according to the core structure. The key-lock interaction between the L-R can be with the core structure, and oriented atoms can contribute null/positive/negative to the effect of this core structure.

3.2.b) The main advantages of aligning molecules:

- ✓ Aligning the core structure can be a good start to arranging molecules with respect to the receptor.
- ✓ It is possible for all atoms of the molecules to be optimally aligned.

- ✓ With maximum overlap in all molecules, a safe superimposition can occur. Even molecules with very different skeletal structures can have atoms coincident at points that are not visible or noticed.
- ✓ With respect to the receptor, the most compatible structure of the molecule can be chosen from the conformer that allows the most atoms to overlap.
- ✓ Corresponding ligand atoms to contact the interaction points of the receptor can be in the same cluster, making clustering optimum and safe.

The major disadvantage of aligning molecules is that misdirected atoms and one or more molecules with aggregations that form because of this significantly distort the calculation results. As with anything, a small mistake made at the start can lead to more errors in the next. In determining a common core structure, care is taken to ensure that the electronic tolerance (ET) and geometric tolerance (GT) values are as small as allowed. By default, GT values start with a value of 0.2 or 0.3 Angstrom at the beginning and are produced with values of 0.1 increments up to 0.5 Angstroms (the sum of ± 0.5 is 1 and 1Å is approximately the length of an H-bond). On the other hand, ET values (for atomic charge) are produced in 0.05 increments starting from 0.05 atomic load unit up to 0.15. Accordingly, the C-atom in CH_3 and the O-atom in O-H should not be in the same cluster in the core due to their different charges. By clustering around a common core structure, molecules adapt to the 3D structure and electronic values of the receptor in space. If the structure of the receptor is not known, this core structure is considered stochastic for a virtual receptor, on the other hand, if the receptor structure is known, a single core structure compatible with the receptor should be selected. To prevent any structure of the molecule under consideration from becoming a pseudo-core structure, attention is paid to the potential of the remaining atoms to coincide with the template, and the maximum number of cores that provide overlap is selected.

3.2.c) Superimposed molecules with each other according to all atoms of the template is the most reliable way for molecules to behave similarly to the receptor. Among the conformers, it is the best superimposed one that provides the most overlap with the template and can be chosen as the conformation of the molecule to represent the molecule. On the other hand, at least one of the conformers of the active molecule must contain both electronic and geometric values of the core structure (Tokat et al., 2019).

3.2.d) Clusters are formed from overlapping atoms remaining from the common core structure (Tokat et al., 2019). Geometric similarities are considered, not electronic similarities, for atoms to be contained in a cluster. A cluster of atoms is placed in the cluster with a maximum

difference of 1Å to the Cartesian coordinate values of the reference atom. Clusters of atoms above a specified limit number form a cluster. Many clusters can have different positions in space besides the core structure and can contact and interact with the receptor.

The atoms in the core, which have geometric and electronic common values in all the molecules studied, form the starting part of clustering. Regardless of whether their electronic properties are the same from the remaining oriented atoms, clustering with atoms in similar positions continues. In the meantime, it is not checked whether all molecules contain atoms in the cluster, but the number of atoms in the cluster must be above a limit value for aggregation. Thus, heaps with a small number of atoms are not included in the aggregation, reducing unnecessary detail, and saving computation time.

3.2.e. Finding or adjusting parameters with GA is the most important tool in determining pharmacophore. GA is used to reveal the 3D structure of the optimum subset within a model. The correct estimation of the features is provided by calculating the parameters depending on the GA. After possible combinations between sets, a model is tried to be established by creating examples among all descriptors. Among them, the optimal values of the parameters for each sample are found and it is understood with GA whether the studied sample has a better value than the previous sample and the rejection/acceptance option is operated. In short, the idea of GA is as follows. With a combination of the first three atoms of the core structure, the descriptor group is created, then the GA is operated until the best result is obtained with a fourth descriptor add/remove. After selecting the optimized sample with the fourth descriptor, a similar attempt is made to find the fifth descriptor and similar operations are repeated to add the next descriptors. Some instances of low probability created by previous descriptor spaces are replaced with new instances with higher probability values. At the end of the iterative process (Obrezanova et al., 2007), an optimum subset from the descriptor space with interaction points corresponding to higher probability values; means pharmacophore. The most important advantage of this approach is that; A search is made among the clusters in the training set of the analyzed series. In addition, instead of "catching up" the local minimum probability, a general situation with multiple minima is tried to be found. The major disadvantage is that if a directed atom of a molecule is classified in an incorrect cluster due to the insufficient GT value, the GA's performance is partially lost. Since the GT value used is the same for all clusters, it may not work with the same efficiency in all of them. This situation may cause some clusters to be erroneous and hence GA's inadequacy in choosing sub-clusters. Under these conditions, the living subset of clusters with the greatest probability can be considered optimal.

3.2.f. By applying GA with acceptance-rejection in subsets, it can be predicted what a pharmacophore might be. The subset is enlarged and matured by adding a new set with Advanced Variable Selection, starting from the core structure. Although GA works well with a large pool of descriptors (Obrezanova et al., 2007), it is preferable to define the structure-activity/structure-properties relationship with a subset with a small number of elements. Therefore, if the calculation from the pharmacophore property provided by the newly added cluster caused little improvement from the sub-cluster, the maturation of the sub-cluster is meaningless and unnecessary growth. This means that the scenario discussed is unnecessary. Identifying and selecting relevant subsets from a large collection with very different coordinate values is difficult. Therefore, adding a new subset to a scripted subset is an important step in QSAR modeling. In addition, since the number of clusters must be optimum, the addition of new clusters is not needed as enough variables used in the calculation improves the expressiveness of the proposed model. A common approach for adding a new set to the subset is forward variable selection. The selection of the new interaction set is handled according to the best optimization selection and carried out by minimizing the number of descriptors.

3.3. The dependence of the activity on the descriptor vector

Let us denote the data in the training set with D = $\{\mathbf{y}, \mathbf{X}\}$, where the vector $\mathbf{y} = \{y^{(n)}\}_{n=1}^{N}$ are the observed property (activity) values, the matrix $\mathbf{X} = \{\mathbf{x}^{(n)}\}_{n=1}^{N}$ and vector $\mathbf{x}^{(n)} = \{x^{(m)}\}_{m=1}^{M'}$ (Obrezanova et al., 2007) are LRD descriptors.

Since the model is based on the pharmacophore representing the 3D interaction, the X matrix formed by the LRD vector that plays a role in the 3D interaction points is used. Where N is a set of compounds in the training set (Obrezanova et al., 2007), while M' is the total amount of sets with different positions. It is almost impossible for all these sets to take part in the function. Therefore, the most efficient subset GA is considered stochastic, and someone with M number of subsets with the best statistical result is chosen and the vector $\mathbf{x}^{(n)} = \{x^{(m)}\}_{m=1}^{M}$ consists of sets of M.

Given the descriptor vector x as the subset, we want to construct a model y(x) that predicts the property of a molecule (Obrezanova et al., 2007). Since the values of X in N, M' dimension (similarly N, M size subsets) are different from each other for each LRD type, the calculated y(x) will also be different.

The probable pharmacophore model according to the one that gives the best result for the y(x) calculated from each selected LRD type is to be predicted from the N, M dimensional X matrix. Although there is only one type of X

matrix for atomic charge or coefficients, separate X matrices of both types are formed for both atomic charge and coefficient in KI. The calculated parameters of the receptor side according to this index are also of two types. Instead of considering a single type of interaction (ionic or covalent) separately in the model, both types of interaction are used together and simultaneously. Accordingly, a more stable and general calculation is made by considering a detailed interaction. To highlight this difference, the question arising from the HSAB interactions that arise is that; "How ionic is the interaction in each molecule, how covalent is the amount at each interaction point the same, and how can we suggest the HOMO (or LUMO) energy level of the receptor side in particular?" The answer depends on the difference in energy levels of the HOMO of one of the ligands and receptor sides and the LUMO of the other. The smaller this difference, the more covalent contribution occurs. Even if the receptor structure is known, it is very difficult to precisely determine the HOMO/LUMO frontier orbital energy levels of the receptor with a structure consisting of many atoms. Therefore, the frontier orbital energy value of the receptor can be taken from the average value in the studied molecule series, considering the receptors whose structure is unknown. In the denominator of the second term of Equation 3, frontier orbital energy levels are used as HOMO (or LUMO) of the receptor and LUMO (or HOMO) water for the ligand, respectively, as constant and variable. On the other hand, although the parameters in both types of interaction (ionic and covalent) on the receptor side are constant, the LRD value on the ligand side varies from molecule to molecule (Tokat et al., 2019). Considering HSAB, which has all these details in MCET, both types of interaction and pharmacophore are suggested. Therefore, the prediction function is expected to be smooth with a real 3D model and fit the observed data as much as possible.

Setting the parameters is predicted in the model set up in the training set through the "Leave One Out-Cross Validation" (LOO-CV) (Evsukoff, Fairbairn, Faria, Silvoso, & Toledo, 2006) and can be accurately controlled in the molecules in the external test set (Turkmenoglu et al., 2020). Although the interaction energy arising from parameters and LRD values affects the activity with a non-linear equation, these constants are determined very well with the Levenberg-Marquardt algorithm. Feature and function values are closely followed by regression analysis. LRD values on the ligand side are variable in each molecule, while the adjusted parameters (for atomic charges; Kappa-κ and for atomic coefficients; Xi-ξ) on the receptor side in a proposed subset are constant. Equation 4, where the abbreviation is made by writing the parameters of the receptor side, is arranged according to the ligand being nucleophile and the receptor being electrophile. If the ligand is electrophile and the receptor is nucleophile, the reverse arrangement should be made. Which is electrophile and which

is nucleophile, is decided by the lowest LUMO and highest HOMO energy values, respectively.

$$\Delta E = \sum_{m=1}^{M} [Q_{nuc} * \kappa + \frac{c_{nuc}^2 * \xi}{E_{HOMO(nuc)} - E_{LUMO(elec)}}]_m \quad \textbf{Equation 4}$$

Accordingly, the activity of the molecule is different from each other because of the different LRD cluster vectors of the molecule. If the LRD set vectors of the two molecules are similar across the index, the function values corresponding to each index are also similar. This means that the constants we can adjust between the smoothness of y and the quality (Obrezanova et al., 2007) of fit mentioned above, each LRD type used, and each subset proposed are computed. Even if there are different subsets for each LRD type, the small number of elements in the subsets and the size of q^2 for the training set and r^2 for the test set (Obrezanova et al., 2007) are important.

$$Y^{(n)}(x) = A_l * e^{\left(-\frac{\Delta E_n}{RT}\right)} / \left(e^{\left(-\frac{\Delta E_l}{RT}\right)}\right) \text{ or } Y^{(n)}(x) = A_l * e^{\left(-\frac{(\Delta E_n - \Delta E_l)}{RT}\right)} \quad \textbf{Equation 5}$$

The ΔE_n in Equation 5 was calculated by using the LRD values in the M subset of the n-numbered ligand and the parameter values on the receptor side in Equation 4. The activity value of the n-number molecule depends on the activity of the l-reference molecule and the nonlinear interaction energy of both. Here, the interaction energy of the l-reference molecule is ΔE_l, while the interaction energy of n-molecule is ΔE_n. The parameters are adjusted according to the minimum sum of squares of the differences between the observed and predicted values (Turkmenoglu & Guzel, 2018) of $Y^{(n)}(x)$ in Equation 5. By using the parameters and the x 'vector descriptor given for one of the external test molecules together, the most interaction energy can be calculated and the activity value of the molecule under consideration, {y, X}, can be predicted as in Equation 5. Optimum splitting of the sets is possible according to the similarities of the x vector descriptors and y(x) activities in the training and test sets (Turkmenoglu et al., 2020). Accordingly, four of the five molecules with similar x vector descriptors and y(x) activities are included in the training set and one in the test set (Turkmenoglu et al., 2020).

The prediction of multiple new molecule activities in test sets relative to the training set is a direct extension of using the parameters calculated in Equation 4. If the standard deviation determined for both sets is too large, the descriptor area of the new molecules will indicate that it is far outside the descriptor area covered by the training data (Obrezanova et al., 2007). In such a case, an unreliable guess has been made. An important point about this variance is that, as can be seen from Equation 4, the close values of observed and predicted properties for a molecule in the test set are

tightly linked to their descriptive values as well as adjustable parameter constants. The close values of the observed and predicted properties for a molecule in the test set depend on the adjustable parameter constants. Since the interaction at all points of the pharmacophore is simultaneous, the parameters are adjusted simultaneously. The next question is that both Kappa and Xi parameters are linked to each other simultaneously at each point, specific to KI, and how to calculate it with a special algorithm. On the other hand, the units of both types of parameters must be such that; From the first and second terms in Equation 4, the energy unit on the left side of the equation should come out.

3.4. Algorithms in MCET

3.4.1. Align

a) Suggest LRD type.

b) Choose the lowest energy conformer of the simplest, least atoms, and most active molecule as template.

c) Keep the electronic and geometric values of the core structure that results from the combination of atoms in the template.

d) Place the first three atoms of the core structure in the template (0_1,0_1,0_1 zero-point, origin; x_2,0_2,0_2 and x_3, y_3,0_3) at the coordinate center and arrange the new x-, y-, z-coordinates of the remaining atoms.

e) Generate both electronic tolerance (ET) and geometric tolerance (GT) values.

f) Align the core structure in the conformers of the molecules according to the template based on ET and GT values, keep the new coordinates of the remaining atoms.

g) Sort the core structures according to the number of molecules they contain (in descending order).

h) Repeat 3.4.1-d.

3.4.2. Superimposition

a) Examine the prominent I-core structures. (i = 1; i<I)

b) Generate Geometric Tolerance (GT) values. (ds=0.1; ds<1.0; ++0.1, ds- maximum= 1.0 Angstrom).

c) Examine the j-th conformer of the n-th molecule containing the i-st core. (n=1; n<N, j=1; j<=J)

d) The conformer, overlap the template on the scale of GT values, relative to the core atoms.

e) Select the conformer that best matches the template as the structure of the molecule.

f) Repeat 3.4.2-c.

3.4.3. Clustering Analysis

a) Create M' clusters from atom positions in and around the core structure of the template.

b) Use a chosen conformer structure for each of the N-molecules to represent the molecule. (n = 1, n<N).

c) Include conformer atoms in each of the m-set in M' according to their GT values. (m = 1; m<M').

d) Repeat 3.4.3-b.

e) If the number of molecules containing atoms in the m-cluster under study is not above the critical value, exclude the m-th cluster from the total set M' and reduce the number of M' by one (M' = M'-1).

f) Place each n-th molecule with atoms of similar space structure as the template, in the same m-th cluster index order.

g) Generate X (N, M') matrices for the i-th core and ds tolerance scale examined.

h) Repeat 3.4.2-b until you determine the tolerance value ds that serves optimal clustering.

3.4.4. Application of GA

a) Initially set the M-number positions of the atoms in the core as a subset.

b) Add the m-th cluster outside the M subset ($m \notin M$, $m \in M'$ and $M \cup M'$) to be $m \in M$ (m=M; m<M').

c) Simultaneously calculate the receptor-side parameters of the subset of M-numbers according to the values of the dependent and independent variables {y, X} of the molecules.

d) Make sure to provide the best correlation (q^2) value in the training set to generate the final M subset with the GA.

e) Calculate the correlation (r^2) value in the excluded test set using the parameters of the training model.

f) Repeat 3.4.4-b until you find the cluster that will give the best results.

g) Show that the best model has the best regression (q^2 and r^2, respectively) in both training and test sets.

h) Repeat 3.4.2-a until you find the core structure that mediates the best align and superimposition.

i) Repeat 3.4.1-a until decided on the best LRD.

4. Conclusion

To determine the pharmacophore in 3D space, the clusters formed at the positions of the atoms of the template were considered as local interaction areas and the electronic values of the atoms here were used as LRD. The

molecules were aligned with the receptor by aligning in common core structures, and the remaining and overlapping clusters of atoms were considered in the same cluster and in the same category. Electronic values at the receptor side of the geometry of the subset that can be formed between sets were calculated as parameters. Geometric and electronic values on the active side of the receptor are used in the models as the counterpart of the pharmacophore. The descriptors at each point belonging to the proposed subset are not individually but considered together (Obrezanova et al., 2007). Activities were calculated using the determined parameters on the receptor side and the independent variable LRD descriptors on the ligand side in the model. Two types of parameter values were calculated for HSAB interactions, both Coulombic and covalent, used in the newly developed KI descriptor as LRD.

Many trials are required for individual LRDs, as there may be many pharmacophore candidates for each LRD type. One way out of this confusion; The most important development is to use KI, which includes all the P_{+}/P_{-} ionic, H-donor/acceptor hydrogen bond, and covalent/non-covalent interactions. Thus, it is possible to avoid the use of different LRD types by considering real 3D interactions with quantum values for each type of atom. In this way, an opportunity to comment on a model that will show the HSAB interaction between L-R to obtain a most meaningful pharmacophore structure was provided.

References

Bender, A., & Glen, R. C. (2004). Molecular similarity: a key technique in molecular informatics. *Organic & Biomolecular Chemistry, 2*(22), 3204-3218. doi:10.1039/b409813g

Cartmell, J., Enoch, S., Krstajic, D., & Leahy, D. E. (2005). Automated QSPR through competitive workflow. *Journal of Computer-Aided Molecular Design, 19*(11), 821-833. doi:10.1007/s10822-005-9029-8

Evsukoff, A. G., Fairbairn, E. M. R., Faria, E. F., Silvoso, M. M., & Toledo, R. D. (2006). Modeling adiabatic temperature rise during concrete hydration: A data mining approach. *Computers & Structures, 84*(31-32), 2351-2362. doi:10.1016/j.compstruc.2006.08.049

Fleming, I. (1976). *Frontier Orbitals and Organic Chemical Reactions*. Chichester, UK: Wiley.

Fukui, K. (1982). Role of Frontier Orbitals in Chemical-Reactions. *Science, 218*(4574), 747-754. doi:DOI 10.1126/science.218.4574.747

Guzel, Y., Aslan, E., Turkmenoglu, B., & Su, E. M. (2018). 4D-QSAR Studies Using a New Descriptor of the Klopman Index:

Antibacterial Activities of Sulfone Derivatives Containing 1, 3, 4-Oxadiazole Moiety Based on MCET Model. *Current Computer-Aided Drug Design, 14*(3), 207-220. doi:10.2174/1573409914666180514093543

Hopfinger, A. J., Wang, S., Tokarski, J. S., Jin, B. Q., Albuquerque, M., Madhav, P. J., & Duraiswami, C. (1997). Construction of 3D-QSAR models using the 4D-QSAR analysis formalism. *Journal of the American Chemical Society, 119*(43), 10509-10524. doi:DOI 10.1021/ja9718937

Jorgensen, M. S., Groves, M. N., & Hammer, B. (2017). Combining Evolutionary Algorithms with Clustering toward Rational Global Structure Optimization at the Atomic Scale. *Journal of Chemical Theory and Computation, 13*(3), 1486-1493. doi:10.1021/acs.jctc.6b01119

Kizilcan, D. S., Turkmenoglu, B., & Guzel, Y. (2020). The use of the Klopman index as a new descriptor for pharmacophore analysis on strong aromatase inhibitor flavonoids against estrogen-dependent breast cancer. *Structural Chemistry, 31*(4), 1339-1351. doi:10.1007/s11224-020-01498-9

Klopman, G. (1968). Chemical reactivity and the concept of charge- and frontier-controlled reactions. *Journal of the American Chemical Society, 90*(2), 223-234.

Kubinyi, H. (1998a). [Molecular similarity. 1. Chemical structure and biological action]. *Pharm Unserer Zeit, 27*(3), 92-106. doi:10.1002/pauz.19980270305

Kubinyi, H. (1998b). [Molecular similarity. 2. The structural basis of drug design]. *Pharm Unserer Zeit, 27*(4), 158-172. doi:10.1002/pauz.19980270406

Obrezanova, O., Csanyi, G., Gola, J. M. R., & Segall, M. D. (2007). Gaussian processes: A method for automatic QSAR Modeling of ADME properties. *Journal of Chemical Information and Modeling, 47*(5), 1847-1857. doi:10.1021/ci7000633

Patel, H. M., Noolvi, M. N., Sharma, P., Jaiswal, V., Bansal, S., Lohan, S., . . . Bhardwaj, V. (2014). Quantitative structure-activity relationship (QSAR) studies as strategic approach in drug discovery. *Medicinal Chemistry Research, 23*(12), 4991-5007. doi:10.1007/s00044-014-1072-3

Pearson, R. G. (1997). The HSAB Principle. In (pp. 1-27): Wiley.

Salem, L. (1968). Intermolecular orbital theory of the interaction between conjugated systems. I. General theory. *Journal of the American Chemical Society, 90*(3), 543-552.

Segall, M. D., Beresford, A. P., Gola, J. M. R., Hawksley, D., & Tarbit, M. H. (2006). Focus on success: using a probabilistic approach to achieve an optimal balance of compound properties in drug

discovery. *Expert Opinion on Drug Metabolism & Toxicology, 2*(2), 325-337. doi:10.1517/17425255.2.2.325

Sorensen, K. H., Jorgensen, M. S., Bruix, A., & Hammer, B. (2018). Accelerating atomic structure search with cluster regularization. *Journal of Chemical Physics, 148*(24). doi:Artn 241734 10.1063/1.5023671

Tokat, T. A., Turkmenoglu, B., Guzel, Y., & Kizilcan, D. S. (2019). Investigation of 3D pharmacophore of N-benzyl benzamide molecules of melanogenesis inhibitors using a new descriptor Klopman index: uncertainties in model. *Journal of Molecular Modeling, 25*(8). doi:ARTN 247 10.1007/s00894-019-4120-6

Turkmenoglu, B., & Guzel, Y. (2018). Molecular docking and 4D-QSAR studies of metastatic cancer inhibitor thiazoles. *Computational Biology and Chemistry, 76*, 327-337. doi:10.1016/j.compbiolchem.2018.07.003

Turkmenoglu, B., Guzel, Y., Su, E. M., & Kizilcan, D. S. (2020). Investigation of inhibitory activity of monoamine oxidase A with 4D-QSAR using Fukui indices identifier. *Materials Today Communications, 25*. doi:ARTN 101583 10.1016/j.mtcomm.2020.101583

Willett, P. (2000). Chemoinformatics - similarity and diversity in chemical libraries. *Current Opinion in Biotechnology, 11*(1), 85-88. doi:Doi 10.1016/S0958-1669(99)00059-2

Yilmaz, H., Boz, M., Turkmenoglu, B., & Guzel, Y. (2014). Pharmacophore and Functional Group Identification of 4,4 '-dihydroxydiphenylmethane as Bisphenol-A (BSA) Derivative. *Tropical Journal of Pharmaceutical Research, 13*(1), 117-126. doi:10.4314/tjpr.v13i1.17

Yilmaz, H., Guzel, Y., Onal, Z., Altiparmak, G., & Kocakaya, S. O. (2011). 4D-QSAR Study of p56(lck) Protein Tyrosine Kinase Inhibitory Activity of Flavonoid Derivatives Using MCET Method. *Bulletin of the Korean Chemical Society, 32*(12), 4352-4360. doi:10.5012/bkcs.2011.32.12.4352

Zhang, S. X., Golbraikh, A., Oloff, S., Kohn, H., & Tropsha, A. (2006). A novel automated lazy learning QSAR (ALL-QSAR) approach: Method development, applications, and virtual screening of chemical databases using validated ALL-QSAR models. *Journal of Chemical Information and Modeling, 46*(5), 1984-1995. doi:10.1021/ci060132x

CHAPTER II

MOLECULAR MARKER TECHNOLOGIES IN WHEAT

Dogan Ilhan

(Asst. Prof. Dr.), Kafkas University, e-mail:doganilhan@kafkas.edu.tr
ORCID: 0000-0003-2805-1638

1. Introduction

Molecular marker technology has become very popular in the evaluation of agricultural characters associated with environmental abiotic and biotic stresses, especially yield increase and resilience, in the wheat genome in the last two decades (Kiszonas and Morris, 2018). Molecular markers are defined as the reference DNA sequences associated with the gene or gene region in the genome (Staub et al. 1996). These markers are nucleotide sequences used to reveal polymorphic DNA sequences in all organisms. The most important feature of these sequences is that they are polymorphic and have the ability to produce high quality genotypes and can be found in large amounts in the genome (Semang et al. 2006). Molecular markers are highly reliable compared to morphological markers based on observable characters and biochemical markers based on protein. They are numerous, not impressed by the environment, can be easily observed in the process of plant growth, and do not create inter-locus interactions. Therefore, DNA markers are the best tools for the evaluation of plant materials in breeding studies (Ovesna et al. 2002). DNA-based molecular markers are versatile tools used in taxonomy, physiology, embryology, genetic engineering, etc. areas (Schlotterer, 2004). Along with Polymerase Chain Reaction (PCR) discovery, gene labeling, genetic mapping, determination of agriculturally important genes based on map, selection, molecular relationship, diversity studies in the wheat genome became easier (Joshi et al. 2000). In addition to genotype identification in wheat species and their relatives (Henry, 2001), molecular markers are also used to detect differences in certain loci at the species or population level. The basis in these methods is the PCR amplification of a DNA fragment with known nucleotide sequence (Williams et al. 1990). Molecular marker methods, which are used extensively in the wheat genome today, can be grouped as Traditional marker systems and Next Generation marker technologies.

In this chapter, it is aimed to summarize the molecular marker classes and related literature studies that have been used in wheat for the last 20 years.

2. Classification of Molecular Markers in Wheat

Molecular marker techniques popularly preferred in the wheat genome for different purposes in recent years, Hybridization-based techniques; RFLP (Restriction Fragment Length Polymorphism) (Blanco et al. 1998) and DarT (Diversity Arrays Technology) (Echeverry et al. 2015), PCR-based techniques; RAPD (Randomly Amplified Polymorphic DNA) (Cui et al. 2014), AFLP (Amplified Fragment Length Polymorphism) (Parker et al. 1999), SSR (Simple Sequence Repeats) (Roder et al. 1995), SNP (Single Nucleotide Polymorphism) (Jin et al. 2016), STS (Sequence Tagged Sites) (Huang et al. 2006), CAPS (Cleaved Amplified Polymorphic Sequence) (Hanif et al. 2016), SRAP (Sequence-Related Amplified Polymorphism) and TRAP (Target-Region Amplified Polymorphism) (Sun et al. 2008), EST (Express Squence Tag) (Zhao et al. 2010), iPBS (inter-Primer Binding Site) (Arystanbekkyzy et al. 2019), IRAP (Inter-Retrotransposon Amplified Polymorphism) and REMAP (Retrotransposon-Microsatellite Amplified Polymorphism) (Carvalho et al. 2010), SSAP (Sequence-Specific Amplification Polymorphism) (Queen et al. 2004), SCAR (Sequence Characterized Amplifed Region) (Liu et al. 1999), ISSR (Inter Simple Sequence Repeats) (Nagaoka and Ogihara, 1997) and SCoT (Start Codon Targeted) (Hamidi et al. 2014) (Fig. 1, Table 1). Among these markers, the RFLP technique is not preferred much today. Reason; the use of radioactive material is that it is expensive and time consuming and also high DNA requirement (Grover and Sharma, 2015). SNP markers are preferred in genetic studies because they can detect hidden polymorphisms that cannot be determined by other methods, are densely found in the genome, and also have a low mutation level (Duran et al. 2009). When we compare molecular marker techniques with each other, we can see that they have advantages and disadvantages. These aspects have been discussed in detail in many sources (Idrees and Irshad, 2014). However, in recent years, we see that SSR and SNP techniques are widely used because of high polymorphic character. SSR markers, which are highly polymorphic, are more preferred (Durand et al. 2010). In studies in which diversity analyzes are made in wheat; RAPD, RFLP, SSR and AFLP markers were compared. According to the results, it was determined that AFLP and SSR markers were the most efficient markers according to the resulting polymorphism values. In addition, it is stated that SSR markers are more advantageous and useful because AFLPs are expensive (Parker et al. 2002). EST (Labeled Expressed Sequences), which has been widely used in recent years, is a marker technique created as a result of sequence analyses of cDNAs (complementary DNAs) corresponding to all or a certain part of different mRNAs. It is also known as random sequence analysis of cDNAs. ESTs are used for determining SSRs as well as used in wheat mapping research (Ellis and Burke, 2007). Genomic SSRs can be selected in order to perform genetic diversity, gene labeling and genomic mapping analyses in the wheat genome (Prasad et al. 2000; Roy et al. 1999).

In many studies in which genetic relationships between wheat varieties are determined, it has been revealed that SSRs are very useful tools that can be used by gene banks in determining the variation of genetic resources (Dreisigacker et al. 2005). In last 10 years, molecular marker techniques used in parallel with the development of DNA sequencing technology have also improved. Based on this technology, chips have been developed for SNPs known as single nucleotide polymorphisms in the genome. SNP arrays (SNP chips) technology has been preferred to define Quantitatif Trait Loci to determine wheat quality (Jin et al. 2016). KASP (Kompetitive-Allele Specific PCR) marker system, which is a technology focused on detecting SNPs and indel (insertion / deletion) mutations in wheat genomes, has also been used in wheat populations to evaluate many agricultural characters (Rasheed et al. 2016). Primary use of sequencing based technology in wheat genome has been provided by Genotyping by Sequencing (GBS) technique for agricultural characters (Poland et al. 2012). In the following process, technologies such as single molecule sequencing (SMS) and nanopore have been developed, but their informative use for the wheat genome is not yet in question (Kiszonas and Morris, 2018) (Table 1). It is understood that new generation marker technologies are more effective than traditional marker systems (Kumar et al. 2021).

It is possible to classify the molecular markers studied in the wheat genome more generally under 2 groups; Traditional marker systems (i) and Next Generation marker technologies (ii). Traditional marker systems; While hybridization-based and PCR-based are divided into 2 groups, New Generation marker technologies can also be classified under 2 groups based on Hybridization and Sequencing (Fig. 1).

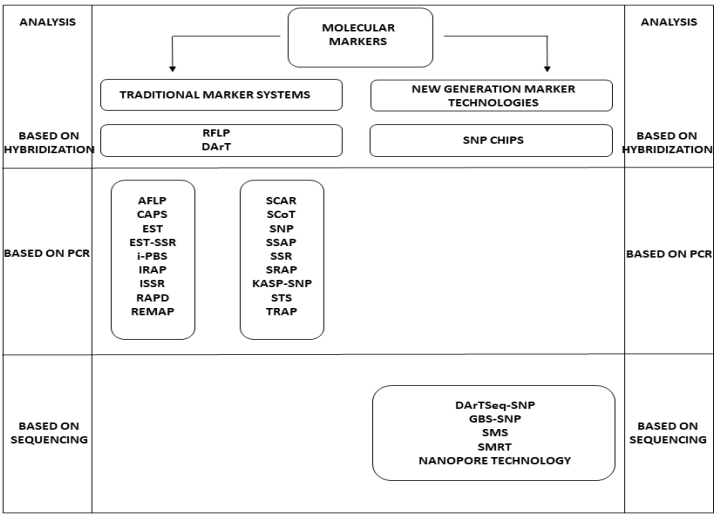


Figure 1. Traditional and Next Generation Molecular Marker Technologies Used in the Wheat Genome (Kiszonas and Morris, 2018; Kumar et al. 2021)

Table 1. Molecular Marker Technologies and Usages with References in Wheat (Kiszonas and Morris, 2018; Kumar et al. 2021)

Marker	Usage of Marker	Reference
AFLP	Mapping loci	Parker et al. 1999
CAPS	Grain weight and yield	Hanif et al. 2016
DArT	QTLs in quality related properties	Echeverry-Solarte et al. 2015
EST	QTLs for grain protein and and flour protein	Zhao et al. 2010
EST-SSR	QTL mapping for yield properties	Li et al. 2007
I-PBS	Phylogenetic and taxonomic relationship	Arystanbekkyzy et al. 2019
IRAP	Genetic variability	Carvalho et al. 2010
ISSR	Comparison with other marker systems	Nagaoka and Ogihara, 1997
RFLP	Flour SDS sedimentation	Blanco et al. 1998
RAPD	Linkage and QTL maps for grain yield-related properties	Cui et al. 2014
REMAP	Genetic variability	Carvalho et al. 2010
SCAR	Powdery mildew resistance gain with the *Pm21* gene	Liu et al. 1999
SCoT	Comparative efficiency of functional gene-based markers	Hamidi et al. 2014
SNP	QTL maps for quality parameters using Genome wide analyses	Jin et al. 2016
SSAP	Linkage use and genetic diversity analyses	Queen et al. 2004
SSR	Density and chromosomal locations of microsatellites in the genome	Roder et al. 1995
SRAP	QTLs for protein and starch traits	Sun et al. 2008
KASP-SNP	Genes underpinning key economic traits	Rasheed et al. 2016
STS	Polyphenol oxidase enzyme activity	Sun et al. 2005
TRAP	QTLs for protein and starch traits	Sun et al. 2008
SNP CHIPS	Mapping of *Lr49* gene	Nsabiyera et al. 2020
DArTSeq-SNP	Resistance to rust diseases	Chhetri et al. 2016
GBS-SNP	High-density genetic maps	Poland et al. 2012
SMS	Sequencing DNA	Heather and Chain, 2016
SMRT	Sequencing DNA	Heather and Chain, 2016
NANOPORE TECHNOLOGY	Wheat virüs identification	Fellers et al. 2019

3. An Overview of Literature Studies

Until today, there are many literature in which molecular markers are preferred in the wheat genome, especially in the last 20 years. These studies are mostly aimed at determining the origin and evolutionary process of wheat, nutritional value, genetic diversity of the lines used, molecular affinity, segregation of populations and genetic maps. Since there are many studies in the literature, studies conducted with marker systems are exemplified in this section as a summary.

Genetic and physical maps of wheat were created using RFLP markers. With this study in wheat, populations that separate populations on physical maps up to 2000 RFLP loci (Gupta et al. 1999; Hussain and Qamar, 2007). AFLPs have been used for genetic diversity, phylogenetic and mapping strategies in wheat (Burkhamer et al. 1998; Parker et al. 1999; Bohn et al. 1999). DArT markers have gained popular use, allowing the development of multiple QTLs for wheat quality (Suprayogi et al. 2009).

Using RFLPs in a wheat population popular for flour viscosity, mapping, and quality studies, QTLs and unrelated marker discovered for flour viscosity at 1AS, 2A, 2B, 2DL, 3BL positions of chromosomes (Udall et al. 1999).

In a study examining the genetic similarity between triticale varieties and common wheat, SSR, AFLP and RAPD markers were used. 14 bread wheat (*Triticum aestivum* L.emend. Fiori & Paol. And *Triticum durum* Desf.) and triticale cultures were used. 31 RAPD, 25 SSR, and 6 AFLP markers were selected in the study. Polymorphic band numbers were found as 39.7 for AFLP, 12.7 for RAPD and 4.0 for SSR. However, PIC (Polymorphism Information Content) value was determined at the highest SSRs with a value of 0.13-0.86. In the evaluation made in terms of genetic similarity, it was emphasized that dendrograms obtained from SSR are more reliable than AFLP and RAPD. It has been stated that most of the varieties can be determined with SSR markers (Garg et al. 2001).

In another study, phylogenetic and taxonomic relationships of wild and cultivated Turkish emmer wheat populations were investigated with ipbs-rtetrotransposons markers. In the study, the genetic diversity of a total of 38 wheat varieties including 29 wild emmer, 4 cultivated emmer (*Triticum turgidum* ssp *dicoccoides*) and 5 durum wheat (*T. turgidum* ssp. *durum*) was investigated. Polymorphism average was found to be 87.85%. Polymorphic information content (PIC) was also determined as 0.660. GenAlEx v6.5 was used to calculate gene diversity parameter. In addition, Shannon's information index was investigated for effective allele number. Gene diversity was determined to be 0.489, the average number of effective alleles to be 1.96, and these values were indicated as indicative of a high level of variation. In the dendrogram drawn by UPGMA method, it was determined that all genotypes were separated according to their geographical locations. In the clustering

study between 3 species, durum wheat was clearly separated from wild emmer and culture emmer. It is stated that the results revealed here are also supported by the main coordinate analysis. Genotypes are divided into two populations according to Bayesian-based clustering. Genetic distance between wheat varieties was calculated using R statistical software and Jaccord coefficient. The average genetic distance has been determined as 0.398. As a result of the study, it was emphasized that iPBS-retrotransposons are effective in investigating diversity and phylogenetic relationships (Arystanbekkyzy et al. 2019).

The genetic diversity parameters were investigated for wild emmer populations originating from Israel and Turkey. In the study, 25 wild emmer populations (200 individuals) were analyzed with 1000 SNP markers. They stated that SNP markers reveal the genetic diversity at a moderate level because they show biallity. Polymorphism rate was determined as 80.6%. PIC value was calculated as 0.153 and gene diversity as 0.184. STRUCTURE analysis was performed and the optimal K value was determined as 2. In other words, they reported that wild emmer wheats are genetically divided into two groups. They also reported that clearly separated from the rest of the case of Turkey. Cluster analysis was conducted in the study and they stated that wild emmer wheat was grouped in parallel with its geographical distribution. They explained that there are compatible correlations between marker loci and ecogeographic conditions. They reported that natural selection has an crucial importance in the adaptability of gene regions related to consistency in emmer wheat. They also stated that SNP markers are selective for emmer wheat populations to detect genetic diversity. They also reported that genetic diversity may be associated with ecogeographic conditions (Ren et al. 2013).

In another study, more than 300 wheat varieties were evaluated with 70 KASP markers for certain economic genes. Although the KASP system was effective compared to phenotypes and STS markers in the study in which quality parameters, core texture, mixograph parameters, yellow pigment and polyphenol oxidase were included, further studies were required (Rasheed et al. 2016).

In a study conducted with Italian emmer wheat, agronomic, quality and definition of emmer wheat (*Triticum dicoccon*) was made. Molecular analysis of 20 emmer wheat has been done in the study. 6 EST-SSR markers, 13 SSR markers and 6 ISSR markers were used to evaluate genetic variations. It was stated that the protein content was high in all genotypes and its average was found to be 19.2%. 118 molecular loci were identified in the study. The expected heterozygosity at all loci is 0.25. If only SSR and EST-SSRs are taken into account, the expected heterozygosity reaches 51%. 71 specific alleles were identified. In the study, it was explained that SSRs have an effective discrimination power in advanced breeding lines, modern varieties and growing area for emmer genotypes. It was stated that all molecular

markers were used in the dendrogram drawn and a clustering in two distinct groups was observed. The populations showed quite a lot of genetic differentation. The researchers stated that the reason for this may be the high variation and cross pollination. As a result, they stated that the observed large variation is important in increasing the use of emmer, transferring rich properties to durum wheat and using it in breeding programs. In addition, it was emphasized that emmer wheats show a significant genetic variability for both agromorphological and molecular characteristics. In recent years, it is stated that the cultivation area of emmer wheat will increase due to its nutritious properties and increased interest in organic agriculture (Pagnotta et al. 2009).

The correlation of eco-genetic conditions of wild emmer (*Triticum dicoccoides*) wheat with EST-SSR marker diversity was investigated. In the study, 15 wild emmer populations and 25 EST-SSR markers were analyzed. A total of 92 EST-SSR alleles were identified and the number of alleles ranged from 1 to 7. The average number of alleles was also found to be 3.68. In addition, the number of alleles in the B genome was 3.88 and the PIC value was more than those in the A genome (3.38). Genetic similarity coefficient is between 0.189-0.966. Clustering analysis was performed in the study and all genotypes were clustered in 4 main groups in the dendrogram drawn. Genetic distance value (GD) among wild emmer populations varied to 0.112-0.672 and the average GD value was calculated as 0.406. They reported that populations farther away from each other had lower genetic distances and adjacent populations had higher genetic distances. According to the Mantel test results, they reported that the genetic distance results were independent of the geographic region. As a result of the study, they reported that there is significant genetic variation in EST-SSR loci of *T. dicoccoides* populations originating from Israel and that they are associated with ecological factors, albeit very little. They also stated that EST-SSR marker variation is compatible with natural selection and is affected by various factors (Dong et al. 2009).

In another study, for the genetic diversity of *aestivum* wheat, It was tried to be determined by using morphological characters and SSR markers. Seven wheats belonging to the *T. aestivum* L. species were used. Five varieties of wheat were obtained from the Agricultural Research Center (ARC), Giza and Egypt. Two wheat varieties are selected, 48 SSR markers and 9 morphological characters were used to determine the genetic diversity of wheat. A total of 48 alleles were identified and the average of alleles was 3.2. The number of alleles varies between 2-7. The most alleles were detected in the Xgwm 437 (91-123 bp) marker. Polymorphism information content (PIC) was determined between 0.548-0.816. Gene diversity varied from 0.08 to 0.95 and the mean value was calculated as 0.72. Cluster dendrogram was created using UPGMA to evaluate the relationship between wheat varieties. Two main clusters were

observed in the dendrogram. While there are three species in one of the clusters, the other cluster is divided into sub-clusters. Serious genetic diversity was found among seven wheat. With the dendrogram, the ability of SSRs to detect a great deal of genetic variation has been demonstrated, even in genotypes with narrow genotype pools. As a result, it has been reported that SSRs can be selected successfully in genetic diversity analyses (Salem et al. 2008).

Wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) in a study to describe the size of genetic variation of populations, collected 91 samples were used in the south of Turkey. The change of chloroplast DNA in plants taken from two natural habitats was investigated. Allelic differentation in 24 SSR marker loci in the organel genome was examined and they reported that variation was detected at 15 SSR loci. The average number of alleles was determined to be 2.17 (ranging from 1-4). Estimated diversity index averages were found to be 0.28 and 0.29 for the two populations respectively. According to their results, although the distance between the two populations was only 13 km, a very clear genetic differentiation was observed. A clear separation between the 2 populations was observed as a result of PCoA. In addition, 2 main groups were observed in the dendrogram drawn. A total of 23 chloroplast haplotypes have been identified. They noted that haplotypes for each population show uneven microgeographical distribution. According to the results of the analysis, the researchers stated that even in small emmer populations, genetic diversity is higher than other wheat. As a result; Turkey's southern wild emmer wheat, within and between populations indicated that they have very high genetic diversity level (Shizuka et al. 2015).

Genotyping by Sequencing (GBS) method was used to reveal the evolutionary process of tetraploid cultivated wheats. In the study, 1,172,469 SNP were determined in the tetraploid wheat population consisting of 189 wild and cultivated wheat. Most SNPs were mapped to the A and B genomes. All of the populations have been selected to cover their entire geographic distribution. Principal component analyzes (PCA) separated wild emmer (*T. turgidum* subsp. *dicoccoides*) from culture emmer (*T. turgidum* subsp. *dicoccum*) and bare wheats. They stated that only a few populations had abnormal clustering in the PCA analysis, which may be due to misclassification. They noted that the GBS method can provide good information about molecular similarities among wheat species or subspecies. As a result of the PCA analysis, they found that tetraploid cultivated wheats had the closest relationship to wild group from the northern Fertile Crescent, coincide with the outputs of the molecular analyses on wheat. In addition, according to the PCA results, bare wheats clustered separately with the culture emmers. They reported that the diversity is higher in wild group (*T. turgidum* subsp. *dicoccoides* and *T. timopheevii* subsp. *armeniacum*) compared to culture wheat groups for genetic diversity. 189 populations resulted in a total of 605 million readings and

achieved 6,962,191 tags with three readings per tag. The proportion of heterozygous areas for 186 populations is 7.46%, with an average minor allele frequency of 0.133. STRUCTURE analysis was performed to examine the population structure and K = 4 for 186 populations. According to PCA, NJ and STRUCTURE analyzes, they noted that wild emmer wheats formed a homogeneous population. As a result of the study, they stated that determining SNP with GBS provides very accurate information for determining the molecular relations between wild and cultivated tetraploid wheat species and subspecies. They also reported that GBS is a fast way to obtain diversity data from the entire genome (Oliveria et al. 2020).

High density Infinium arrays were successfully implemented in wheat genome for whole genome SNP genotyping. The International Wheat SNP Working Group (IWSWG) worked with Illumina to develop the Infinium 9K and 90K iSelect SNP genotyping sequences (Cavanagh et al. 2013; Wang et al. 2014). 7504 SNPs were determined from 9K iSelect SNP sequences and a wheat genetic map of an average of 1.9 ± 1.0 SNP / cM was created (Cavanagh et al. 2013). The 90K iSelect SNP sequence was preferred to map 46,977 SNPs in the wheat genome (Wang et al. 2014; Liu et al. 2016). High density wheat genotyping SNP sequences have been introduced on the Affymetrix Axiom platform. By using SNP chip technology, 119 additional QTLs were determined on 20 chromosomes to determine wheat quality (Jin et al. 2016). A 35K SNP genotyping sequence was produced on the Affymetrix GeneTitan platform (Allen et al. 2017). Subsequently, the Wheat660K SNP series, developed by the Chinese Academy of Agricultural Sciences and synthesized by Affymetrix Axiom, contributed to the creation of high-density genetic maps by offering wide mapping strategies in wheat (Cui et al. 2017).

4. CONCLUSION

The cultivation area and the number of cultivated varieties of the most cultivated grains in the world are increasing day by day. In addition to classical breeding techniques, it has become possible to develop new varieties against both agricultural characteristics and abiotic and biotic stress factors thanks to modern molecular tools developed today. Thanks to molecular marker techniques and sequencing technologies, which have been used and developed extensively in the last two decades, notable improvements have been made in wheat genomics. In the light of this basic information, in this chapter, the classification of molecular marker technologies that have been popular and still developing for the last 20 years has been made and a summary is given with the support of the literature. Traditional marker techniques have left their place in New Generation Marker systems based on Sequencing, which are newer and more modern technologies in the last 5 years. By means of these technologies, it has become easier to carry out basic studies such as genetic diversity, molecular relations, genetic mapping, modern breeding and population dynamics in wheat genomes.

References

Allen, A. M., Winfield, M. O., Burridge, A. J., Downie, R. C., Benbow, H. R., Barker, G. L., et al. (2017). Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum* L.). Plant Biotechnol J., *15*, 390–401.

Arystanbekkyzy, M., Nadeem, M. A., Aktas, H., Yeken, M. Z., Zencirci, N., Nawaz, M. A., et al. (2019). Phylogenetic and taxonomic relationship of turkish wild and cultivated emmer (*Triticum turgidum* ssp. *dicoccoides*) revealed by iPBSretrotransposons markers. Int J Agric Biol. https://doi.org/10.17957/IJAB/15.0876

Blanco, A., Bellomo, M. P., Lotti, C., Maniglio, T., Pasqualone, A., Simeone, R., . . . Di Fonzo, N. (1998). Genetic mapping of sedimentation volume across environments using recombinant inbred lines of durum wheat. Plant Breeding, *117*, 413–417. https://doi.org/10.1111/j.1439 0523.1998.tb01965.x

Bohn, M., Utz, H. F., Melchinger, A. E. (1999). Genetic similarities among winter wheat cultivars determined on the basis of RFLPs, AFLPs, and SSRs and their use for predicting progeny variance. Crop Sci., *39*, 228–237.

Burkhamer, R. L., Lanning, S. P., Martens, R. J., Martin, J. M., Talbert, L. E. (1998). Predicting progeny variance from parental divergence in hard red spring wheat. Crop Sci., *38* (1), 243–248.

Carvalho, A., Guedes-Pinto, H., Martins-Lopes, P., Lima-Brito, J. (2010). Genetic variability of old Portuguese bread wheat cultivars assayed by IRAP and REMAP markers. Ann Appl Biol., *3*, 337–345.

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G. L., Akhunova, A., See, D., Bai, G., Pumphrey, M., Tomar, L., Wong, D., Kong, S., Reynolds, M., da Silva, M. L., Bockelman, H., Talbert, L., Anderson, J. A., Dreisigacker, S., Baenziger, S., Carter, A., Korzun, V., Morrell, P. L., Dubcovsky, J., Morell, M. K., Sorrells, M. E., Hayden, M. J., Akhunov, E. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc Natl Acad Sci., USA *110* (20), 8057–8062.

Chhetri, M., Bansal, U., Toor, A., Lagudah, E., Bariana, H. (2016). Genomic regions conferring resistance to rust diseases of wheat in a W195/BTSS mapping population. Euphytica, doi:10.1007/s10681-016-1640-3

Cui, F., Zhao, C. H., Ding, A. M., Li, J., Wang, L., Li, X. F., Bao, Y. G., Li, J. M., Wang, H. G. (2014). Construction of an integrative linkage map

and QTL mapping of grain yield-related traits using three related wheat RIL populations. Theor Appl Genet., *127*, 659–675.

Cui, F., Zhang, N., Fan, X. L., Zhang, W., Zhao, C. H., Yang, L. J., Pan, R. Q., Chen, M., Han, J., Zhao, X. Q., Ji, J. (2017). Utilization of a Wheat 660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. Sci Rep., *7* (1), 3788.

Dong, P., Wei, Y-M., Chen, G-Y., Li, W., Wang, J-R., Nevo, E. and Zheng, Y-L. (2009). EST-SSR diversity correlated with ecological and genetic factors of wild emmer wheat in İsrael. Hereditas, *146*, 1-10.

Dreisigacker, S., Zhang, P., Warburton, M. L., Skovmand, B., Hoisington, D. and Melchinger, A. E. (2005). Genetic Diversity Among and with in CIMMYT Wheat Landrace Accessions Investigated with SSRs and Implications for Plant Genetic Resources Management. Crop Science, *45*, 65-661.

Duran, C., Appleby, N., Vardy, M., Imelfort, M., Edwards, D. (2009). Single nucleotide polymorphism discovery in barley using autoSNPdb. Plant Biotechnology Journal, *7*, 326-333.

Durand, J., Bodenes, C., Chancerel, E., Frigerio, J. M., Vendramin, G., Sebastiani, F., Buonamici, A., Gailing, O., Koelewijn, H. P., Villani, F., Mattioni, C., Cherubini, M., Goicoechea, P. G., Herran, A., Ikaran, Z., Cabane, C., Ueno, S., Alberto, F., Dumoulin, P. Y., Guichoux, E., de Daruvar, A., Kremer, A., Plomion, C. (2010). A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics, *11*, 570.

Echeverry-Solarte, M., Kumar, A., Kianian, S., Simsek, S., Alamri, M. S., Mantovani, E. E., . . . Mergoum, M. (2015). New QTL alleles for quality-related traits in spring wheat revealed by RIL population derived from supernumerary9 non-supernumerary spikelet genotypes. Theoretical and Applied Genetics, *128*, 893–912. https://doi.org/10.1007/s00122-015-2478-0

Ellis, J. R. and Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. Heredity, *99*, 125-132.

Fellers, J. P., Webb, C., Fellers, M. C., Shoup Rupp, J., De Wolf, E. (2019). Wheat Virus Identification Within Infected Tissue Using Nanopore Sequencing Technology. Plant Dis., *103*, 2199–2203.

Garg, M., Sukhwinder, S., Baldev, S., Kuldeep, S., Dhaliwal, H. S., Singh, S., Singh, B., Singh, K. (2001). Estimates of Genetic Similarities and DNA Fingerprinting of Wheat (*Triticum* Species) and Triticale Cultivars Using Molecular Markers. Indian Journal of Agricultural Sciences, *71*, 438-443.

Gupta, P. K., Varshney, R. K., Sharma, P. C., Ramesh, B. (1999). Molecular markers and their applications in wheat breeding. Plant Breed *118* (5), 369–390.

Grover, A., Sharma, P. C. (2015). Development and use of molecular markers: past and present. Critical Reviews in Biotechnology, *36* (2), 1-13.

Hamidi, H., Talebi, R., Keshavarzi, F. (2014). Comparative efficiency of functional gene-based markers, start codon targeted polymorphism (SCoT) and conserved DNA-derived polymorphism (CDDP) with ISSR markers for diagnostic fingerprinting in wheat (*Triticum aestivum* L.). Cereal Res Commun., *42* (4), 558–567.

Hanif, M., Gao, F., Liu, J., Wen, W., Zhang, Y., Rasheed, A., Xia, X., He, Z., Cao, S. (2016). *TaTGW6-A1*, an ortholog of rice *TGW6*, is associated with grain weight and yield in bread wheat. Mol Breeding, *36* (1), 1.

Heather, J. M., and Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. Genomics, *107*, 1–8. https://doi.org/ 10.1016/j.ygeno. 2015.11.003

Henry, R. J. (2001). Plant genotyping-The DNA fingerprinting of plants. CABI Publishing, UK.

Huang, X. Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D. G., Noll, J. S., . . . Brown, P. D. (2006). Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). Theoretical and Applied Genetics, *113*, 753–766. https://doi.org/10.1007/s00122-006-0346-7

Hussain, S. S., Qamar, R. (2007). Wheat genomics: challenges and alternative strategies. Proc Pakistan Acad Sci., *44*, 305–327.

Idrees, M. and Irshad, M. (2014). Molecular markers in plants for analysis of genetic diversity: a review. Eur. Acad. Res., *2*, 1513–1540.

Jin, H., Wen, W., Liu, J., Zhai, S., Zhang, Y., Yan, J., . . . He, Z. (2016). Genome-wide QTL mapping for wheat processing quality parameters in a Gaocheng 8901/Zhoumai 16 recombinant inbred line population. Frontiers in Plant Science, *7*, 1032.

Joshi, S. P., Gupta,V. S., Aggarwal, R. K., Ranjekar, P. K., Brar, D. S. (2000). Genetic diversity and phylogenetic relationship as revealed by inter simple sequence repeat (ISSR) polymorphism in the genus *Oryza*. Theor. Appl. Genet., *100*, 1311–1320.

Kiszonas, A. M. and Morris, C. F. (2018). Wheat breeding for quality: a historical review. Cereal Chem., *95*, 17–34.

Kumar, S., Kumar, M., Mir, R. R., Kumar, R. and Kumar, S. (2021). Advances in Molecular Markers and Their Use in Genetic Improvement of Wheat. Physiological, Molecular, and Genetic Perspectives of Wheat Improvement, https://doi.org/10.1007/978-3-030-59577-7_8

Li, S., Jia, J., Wei, X., Zhang, X., Li, L., Chen, H., Fan, Y., Sun, H., Zhao, X., Lei, T., Xu, Y., Jiang, F., Wang, H., Li, L. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. Mol Breed., *20*, 167–178.

Liu, Z., Sun, Q., Ni, Z., and Yang, T. (1999). Development of SCAR markers linked to the Pm21 gene conferring resistance to powdery mildew in common wheat. Plant Breed., *118*, 215–219.

Liu, S., Assanga, S. O., Dhakal, S., Gu, X., Tan, C. T., Yang, Y. et a.l (2016). Validation of chromosomal locations of 90K array single nucleotide polymorphisms in US wheat. Crop Sci., *56* (1), 364–373.

Nagaoko, T., Ogihara, Y. (1997). Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. Theor Appl Genet., *94*, 597—602.

Nsabiyera, V., Baranwal, D., Qureshi, N., Kay, P., Forrest, K., Valárik, M., et al. (2020). Fine mapping of Lr49 using 90K SNP chip array and flow sorted chromosome sequencing in wheat. Front. Plant Sci., 10:1787:1787. doi: 10.3389/ fpls.2019.01787

Oliveria, H. R., Jacocks, L., Czajkowska, B. I., Kennedy, S. L., Brown, T.A., (2020). Multiregional origins of the domesticated tetraploid wheats. PLOS ONE, *15* (1): e0227148.

Ovesna, J., Polakova, K. and Leisova, L. (2002). DNA analyses and their Applications in Plant Breeding. Czech Journal Genet. Plant Breed., *38* (1), 29–40.

Pagnotta, M. A., Mondini, L., Codianni, P. and Fares, C. (2009). Agronomical, quality and molecular characterization of twenty Italian emmer wheat (*Triticum dicoccon*) accessions. Genet Resour Crop Evol., *56*, 299-310.

Parker, G. D., Chalmers, K. J., Rathjen, A. J., and Langridge, P. (1999). Mapping loci associated with milling yield in wheat (*Triticum aestivum* L.). Molecular Breeding, *5*, 561–568. https://doi.org/10.1023/A:10096780 23431

Parker, G. D., Fox, P. N., Langridge, P., Chalmers, K., Whan, B. and Ganter, P. (2002). Genetic diversity within Australian wheat breeding programs based on molecular and pedigree data. Euphytica, *124*, 293-306.

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using

a novel two-enzyme genotyping-by-sequencing approach. PLoS ONE, *7*, e32253. https://doi.org/10.1371/ journal.pone.0032253

Prasad, M., Varshney, R. K., Roy, J. K., Balyan, H. S., Gupta, P. K. (2000). Theuse of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. Theor. Appl. Genet., *100*, 584–592.

Queen, R. A., Gribbon, B. M., James, C., Jack, P., Flavell, A. J. (2004). Retrotransposon based molecular markers for linkage and genetic diversity analysis in wheat. Mol Genet Genomics, *271*, 91–97.

Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., . . . Mather, D. E. (2016). Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. Theoretical and Applied Genetics, *129*, 1843–1860. https://doi.org/10.1007/s00122-016-2743-x

Ren, J., Chen, L., Sun, D., You, F. M., Wang, J., Peng, Y., Nevo, E., Beiles, A., Sun, D., Luo, M. C., Peng, J. (2013). SNP-revealed genetic diversity in wild emmer wheat correlates with ecological factors. BMC Evolutionary Biology, *13* (1), 169.

Roder, M. S., Plaschke, J., Konig, S. U., Borner, A., Sorrels, M. E. (1995). Abundance, variability and chromosomal location of microsatellite in wheat. Mol Gen Genet., *246*, 327–333

Roy, J. K., Prasad, M., Varshney, R. K., Balyan, H. S., Blake, T. K., Dhaliwal, H. S., Singh, H., Edwards, K. J., Gupta, P. K. (1999). Identification of a microsatellite on chromosomes 6B and a STS on 7D of bread wheat showing an association with preharvest sprouting tolerance. Theor. Appl. Genetic., *99*, 336–340.

Salem, K. F. M., El-Zanaty, A. M. and Esmail, R. M. (2008). Assessing Wheat (*Triticum aestivum* L.) Genetic Diversity Using Morphological Characters and Microsatallite Markers. Word Journal of Agricultural Sciences, *4* (5), 538-544.

Schlotterer, C. (2004). The evolution of molecular markers-just a matter of fashion Nat. Rev. Genet., *5*, 63-69.

Semang, K., Bjornstad, A. and Ndjiondjop, M. N. (2006). An Overview of Molecular Markıır Methods for Plants. African Journal of Biotechnology *5* (25), 2540-2568.

Shizuka, T., Mori, N., Ozkan, H. and Ohta, S. (2015). Chloroplast DNA haplotype variation within two natural populations of wild emmer wheat (*Triticum turgidum* ssp. *dicoccoides*) in southern Turkey. Biotechnology&Biotechnological Equipment, *29* (3), 423-430.

Staub, J. E., Serquen, F. C. and Gubta, M. (1996). Genetic markers, map construction, and their application in plant breeding. HortScience., *31*, 729-740.

Sun, D. J., He, Z. H., Xia, X. C., Zhang, L. P., Morris, C. F., Appels, R., Ma, W. J., Wang, H. (2005). A novel STS marker for polyphenol oxidase activity in bread wheat. Mol Breed., *16*, 209–218. doi:10.1007/ s11032-005-6618-0

Sun, H., Lu, J., Fan, Y., Zhao, Y., Kong, F., Li, R., . . . Li, S. (2008). Quantitative trait loci (QTLs) for quality traits related to protein and starch in wheat. Progress in Natural Science, *18*, 825–831. https://doi.org/10.1016/j.pnsc. 2007.12. 013

Suprayogi, Y., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Knox, R. E., and Singh, A. K. (2009). Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadian durum wheat populations. Theoretical and Applied Genetics, *119*, 437–448.

Udall, J. A., Souza, E., Anderson, J., Sorrells, M. E., and Zemetra, R. S. (1999). Quantitative trait loci for flour viscosity in winter wheat. Crop Science, 39, 238–242. https://doi.org/10.2135/cropsci1999.0011183X 003900010036x

Wang, S., Wong, D., Forrest, K., Allen, A. et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90000 single nucleotide polymorphism array. Plant Biotechnol J., *12*, 787–796.

Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A., and Tingey, S. V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., *18*, 6531-6535.

Zhao, L., Zhang, K. P., Liu, B., Deng, Z. Y., Qu, H. L., and Tian, J. C. (2010). A comparison of grain protein content QTLs and flour protein content QTLs across environments in cultivated wheat. Euphytica, *174*, 325–335. https://doi.org/10.1007/s10681-009-0109-z

CHAPTER III

ON ACTUARIAL PREMIUMS FOR JOINT LAST SURVIVOR LIFE INSURANCE BASED ON ASYMMETRIC DEPENDENT LIFETIMES

Emel Kizilok Kara
(Asst. Prof. Dr.), Kirikkale University, Faculty of Arts and Sciences, Department of Actuarial Sciences, e-mail: emel.kizilok@kku.edu.tr
ORCID: 0000-0001-7580-5709

1. Introduction

It is important to model joint-life distributions in premium calculations for multiple lifetimes. In practice, these calculations are made assuming that the future lifetimes of the insured are independent. However, this assumption is not a very realistic approach. Due to factors such as common lifestyle, exposure to the same danger, death, or illness of one of them (broken heart syndrome), the lifetimes of the spouses can be dependent, (Jagger and Sutton 1991).

For this purpose, there are studies on modeling dependent lifetimes in the literature. In most of these studies, the copula approach, which is frequently used to model the dependent random variables, was used. Frees et al. (1996) use both a copula model and a common shock model to a real data-set of married couples from a large Canadian insurance company. Shemyakin and Youn (2006) study the Hougaard copula function to model the same data-set. Denuit and Cornet (1999) use the Frechet-Hoeffding bounds and Norberg's Markov model to determine the effect of dependence in lifetimes on the actuarial quantities. Hsieh et al. (2020) apply the copula approach to insurance policies. In the context of this topic, there are other studies performed by Frees and Valdez (1998), Carriere (2000), Lee et al. (2014), Chen et.al (2015), Zhu et. al (2017), and Lee and Cha (2018).

On the other hand, besides dependence, also determining the symmetric or asymmetric structures are important to model the lifetime as they could cause different actuarial calculations. For example, the death of one of the married couples, which have a dependent future lifetime, will affect the survivor's lifetime. However, this effect may differ depending on whether the first death is a female or a male. The fact that this effect is not the same means that the dependence between the lifetime of the spouses is

asymmetrical. For this purpose, asymmetric copula models should be used to model the dependence structure of unsymmetrical lifetimes.

In this context, recent studies have shown that the dependency between the lifetime of the insured spouses has a significant effect on the premium coefficients (Shemyakin and Youn 2006, Luciano et al. 2008, Bakar 2018). However, there are a few studies in the literature related to the pricing premium including asymmetric calculations. Zhu et. al (2017) compute the swap premiums under asymmetric Hierarchical Archimedean copula (HAC). Finally, Lu (2017) and Dufresne et al. (2018) use symmetrical Archimedean copulas to model the lifetimes of married couples. However, at the end of their study, they say that lifetime data are asymmetrically dependent.

Therefore, this study aims to examine the effect of asymmetric dependence on actuarial premiums. For this purpose, premiums calculated assuming the independent lifetimes are compared with premiums calculated under symmetrical and asymmetrical cases. It is assumed that the symmetrical and asymmetric lifetimes are modeled by the FGM and GFGM copula families, respectively. In the study, firstly, the joint survival and mortality functions are derived for the proposed models according to Gompertz's law of marginal mortality. Then, net single premiums are computed according to the selected ages and the parameter values for the whole life insurance policies under the joint last survivor status. The results show that asymmetry has a significant effect on premium calculations. Also, as explained in the examples given in the last section, it is concluded that the premiums calculated with the asymmetric model are higher than the others.

This paper is organized as follows. Section 2 includes the preliminaries for symmetric and asymmetric dependence, the joint survival and mortality functions, and actuarial net single premium calculations. Section 3 is the original part of this study, where the joint survival and mortality functions are obtained for the assumed lifetime models under the last survivor status. In Section 4, the net single premiums for the proposed copula models are calculated and exemplified for the whole life insurance under the joint last survivor status. Also, the premium coefficients are compared at the selected ages for the proposed models and evaluated according to Spearman's correlations. Finally, the implications and conclusions are given in the last section.

2. Material and Methods

2.1. Asymmetric Dependence

The copula approach is often used to model the dependent random variables. It enables the modeling of dependency independently of marginal distributions without the assumption of normality (Nelsen 2007).

The definition of copula is given by Sklar (1959) with the Sklar theorem. According to this theorem, the joint distribution function is uniquely described as $H(x,y) = C(F(x), G(y))$ with the continuously marginal functions, $F(x)$ and $G(y)$. For uniform marginals, it can be also defined as $C(u,v) = H(F^{(-1)}(u),\ G^{(-1)}(v)$.

Here, a copula C is asymmetric if $C(u,v) \neq C(v,u)$ for all $(u,v) \in [0,1]^2$ whereas C is symmetric otherwise (Nelsen, 2007). For example, Farlie-Gumbel-Morgenstern (FGM) is a symmetric copula, while the generalized FGM (GFGM)- Type II is an asymmetric copula (Jung et al. 2008). In this paper, these copulas are studied and given as follows:

The asymmetric type of the bivariate GFGM introduced by Rodríguez-Lallena and Úbeda-Flores (2004) is defined for $0 \leq u, v \leq 1$ as follows:

$$C(u,v) = uv + \theta u^b v^b (1-u)^\alpha (1-v)^\beta \qquad (1)$$

where $\alpha \geq 1$ and $\beta \geq 1$ are asymmetry parameters; b is any a given value. $\theta \in [-1,1]$ is the dependency parameter for all $b, \alpha, \beta \geq 1$. Here FGM copula is defined by substituting $b = \alpha = \beta = 1$ in Eq (1). The Spearman's rho correlation and consisted intervals are given by Rodríguez-Lallena and Úbeda-Flores (2004).

2.2. The Survival and Mortality Functions for the Joint Lifetimes

The joint-life insurance policies are issued according to the first death and the last survivor cases. Besides, these policy prices vary according to the independence and dependence of future lifetimes. The definitions of the joint survival functions required in determining these prices, introduced by Menge (1938) and Dickson et al. (2013), are given below.

In the first death status, ${}_tp_{xy}$ shows the probability that both lives aged x and y survive t years. *In the last survivor status,* ${}_tp_{\overline{xy}}$ shows the probability that a least one of the lives aged x and y survives t years. The mortality functions are denoted by ${}_t\mu_{xy}$and ${}_t\mu_{\overline{xy}}$, respectively.

$${}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - {}_tp_{xy} \qquad (2)$$

$${}_t\mu_{\overline{xy}} = -\frac{\partial}{\partial t} ln\,[\,{}_tp_{\overline{xy}}] \qquad (3)$$

where ${}_tp_x = \exp(-\int_0^t {}_s\mu_x \, ds)$ with the marginal mortality law, ${}_t\mu_x = -\frac{\partial}{\partial t}\ln[{}_tp_x]$. Also, ${}_tp_{xy} = \exp(-\int_0^t {}_s\mu_{xy} \, ds)$ and ${}_t\mu_{xy} = -\frac{\partial}{\partial t}\ln[{}_tp_{xy}]$. It is noticed under the independence lifetimes, that ${}_tp_{xy} = {}_tp_x \, {}_tp_y$ and ${}_t\mu_{xy} = {}_t\mu_x + {}_t\mu_y$.

On the other hand, the bivariate survival function related to copula under the dependence lifetimes can be given as follows:

$$S_{xy}(s,t) = S_x(s) + S_y(t) - 1 + C\big(1 - S_x(s), 1 - S_y(t)\big).$$

By writing $S_{xy}(t,t) = {}_tp_{xy}$, ${}_sp_x = S_x(s)$, ${}_tp_y = S_y(t)$, ${}_sq_x = 1 - {}_sp_x$ and ${}_tq_y = 1 - {}_tp_y$ for $s = t$

$${}_tp_{xy} = {}_tp_x + {}_tp_y - 1 + C\big({}_tq_x, {}_tq_y\big) \tag{4}$$

If the ${}_tp_{xy}$ given in Eq. (4) is substituting in Eq. (2), then the joint survival probability related to the copula function can be re-determined for the last survivor status as follows:

$${}_tp_{\overline{xy}} = 1 - C\big({}_tq_x, {}_tq_y\big) \tag{5}$$

2.3. Actuarial Net Single Premium

Life insurance policies are issued according to a single life status consisting of one person and a joint life situation consisting of two or more people. The policies in joint life status can be issued according to the first death and the last survivor situations. In the first death life policies, the death benefit is paid to the survivor at the end of the death year when the first death occurs. In the policies last survivor status, the death benefit is paid when the last survivor is dead. In this paper, the last survivor status is evaluated for the premium calculations based on the whole life insurance policies consisting of individuals ages x and y. The net single premium $\bar{A}_{\overline{xy}}$ is defined by (Dickson et al. 2013, Menge 1938) as follows:

$$\bar{A}_{\overline{xy}} = \int_0^\infty e^{-\delta t} \, {}_tp_{\overline{xy}} \, {}_t\mu_{\overline{xy}} dt \tag{6}$$

where the constant δ is the continuously payable interest rate.

In some life insurance, the studies on premium calculations can be seen in Kara (2021), Riaman et al. (2019), and Terzioglu and Sucu (2015).

3. The Deriving of the Survival and Mortality Functions under the Joint Last Survivor Status.

In this section, the inferences of the joint survival and mortality functions for dependent life models are given under the joint last survivor status. For these inferences, Eqs. (5) and (3) are used, respectively, assuming the marginal mortality of Gompertz (1825).

The implications for the independence model are presented in Appendix A. On the other hand, the inferences for the $FGM(\theta)$ and $GFGM(\theta, b, \alpha, \beta)$ constitute the original part of this paper. For this purpose, with the following propositions, survival and mortality functions are given by actuarial notations. The proofs are given in Appendix B (for FGM) and Appendix C (for GFGM) under the last survivor status. Here, firstly, the explicit forms are obtained by Mathematica 10.2, then simplified by substituting the actuarial notations.

Proposition 1. T*he survival and mortality functions for the symmetric dependence (FGM copula) are given by*

$$ {}_t p_{\overline{xy}} = e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}} - Ke^{-\frac{2B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} $$

$$ {}_t\mu_{\overline{xy}} = -\frac{1}{{}_t p_{\overline{xy}}}(((Bc^t(-c^x e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} - c^y e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}} $$

where $K = \theta - e^{\frac{Bc^x(-1+c^t)}{Log[c]}}\theta - e^{\frac{Bc^y(-1+c^t)}{Log[c]}}\theta + e^{\frac{B(-1+c^t)(c^x+c^y)}{Log[c]}}(1+\theta)$

By substituting Eqs. (A1-A4)

(i) $$ {}_t p_{\overline{xy}} = {}_t p_x + {}_t p_y - {}_t p_{xy} - \theta\, {}_t p_{xy} - \theta\left({}_t p_{xy}\right)^2\left(1 - {}_t p_x - {}_t p_y\right) $$

(ii) $$ {}_t\mu_{\overline{xy}} = \frac{1}{{}_t p_{\overline{xy}}}\left({}_t\mu_x\, {}_t p_x + {}_t\mu_y\, {}_t p_y - 2K\, {}_t\mu_{xy}\left({}_t p_{xy}\right)^2 - \left({}_t p_{xy}\right)^2\left(\frac{{}_t\mu_x}{{}_t p_x}\theta + \frac{{}_t\mu_y}{{}_t p_y}\theta - \frac{{}_t\mu_{xy}}{{}_t p_{xy}}(1+\theta)\right)\right) $$

where $K = \frac{1}{{}_t p_x\, {}_t p_y}(\theta\left({}_t p_x\, {}_t p_y - {}_t p_y - {}_t p_x + 1\right) + 1)$.

Proof. The proof is given in Appendix B.

Proposition 2. *The survival and mortality functions for the asymmetric dependence (GFGM copula) are given by Mathematica solution*

$$ {}_t p_{\overline{xy}} = S - \mathrm{G} $$

$$ {}_t\mu_{\overline{xy}} = -\left(\frac{1}{\mathrm{S}-\mathrm{G}}\right)(Bc^t(c^x(-e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} + \frac{G}{-1+e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}}}(-b+(-1+e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}})\alpha)) + c^y(e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} + G\beta + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}(-1 - b(1-e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}})^{-1}G)))) $$

where

$$ S = e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}} - e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} $$

$$ G = (e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}})^{\alpha}(e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}})^{\beta}(1-e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}})^{b}(1 - e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}})^{b}\theta $$

By substituting Eqs. (A1-A4)

(i) $ {}_t p_{\overline{xy}} = {}_t p_x + {}_t p_y - {}_t p_{xy} - \theta({}_t p_x)^{\alpha}({}_t p_y)^{\beta}({}_t q_x)^{b}({}_t q_y)^{b} $

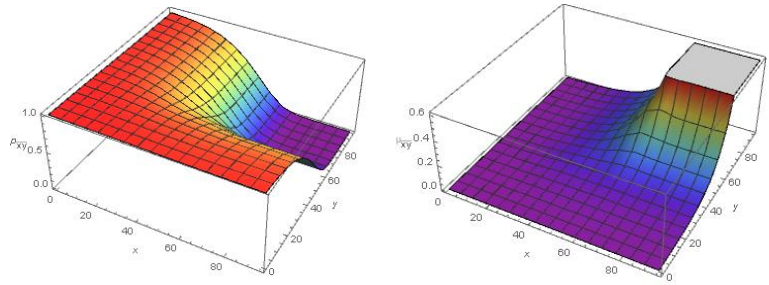
(ii) $ {}_t\mu_{\overline{xy}} = \left(\frac{1}{{}_t p_{\overline{xy}}}\right)({}_t\mu_x({}_t p_x\,{}_t q_y + G(b\frac{{}_t p_x}{{}_t q_x} - \alpha) + {}_t\mu_y({}_t p_y\,{}_t q_x + G(b\frac{{}_t p_y}{{}_t q_y} - \beta))) $

where $S = {}_t p_x + {}_t p_y - {}_t p_{xy}$ and $G = \theta({}_t p_x)^{\alpha}({}_t p_y)^{\beta}({}_t q_x)^{b}({}_t q_y)^{b}$

Proof. The proof is given in Appendix C.

3.1. The Net Single Premiums for the Joint Last Survivor Status

In this section, it is examined the effect of asymmetry on actuarial premiums of the joint last survivor insurance policies. The following tables demonstrate the premium values $\bar{A}_{\overline{xy}}$ for ages 40–60. Computations have been performed according to $FGM(\theta_1)$ and $GFGM\ (\theta_2, b = 1.5, \alpha = 2.5, \beta = 1.8)$ copula models with Gompertz ($B = 0.001,\ c = 1.0887$) marginals. The force of interest is 0.06. Here, the dependence parameters are defined as $\theta_1 \in \{0.3, 0.6, 0.9\}$ and $\theta_2 \in \{3.86744, 7.73489, 11.6023\}$ for the Spearman's correlation values, $\rho \in \{0.1, 0.2, 0.3\}$. These parameters come from using the intervals in Example 4.1 given by Rodríguez-Lallena and Úbeda-Flores (2004). The graphs ${}_t p_{\overline{xy}}$ and ${}_t\mu_{\overline{xy}}$ for GFGM are presented in Figure 1 assuming by $\rho = 0.3$.



(a) (b)

Figure 1. The graphs ${}_tp_{\overline{xy}}$ (a) and ${}_t\mu_{\overline{xy}}$ (b) for GFGM ($\rho = 0.3$ and t=10)

In general, it is seen from these results that dependency has a significant effect on premiums. Remarkably, the premiums increase in policies issued for the individuals with the high correlated lifetimes. For example, in Table 1, $\bar{A}_{\overline{40:50}} = 0.448399$, and $\bar{A}_{\overline{40:50}} = 0.463047$ by assuming $\rho = 0.1$ and $\rho = 0.3$ for the GFGM model, respectively. The other noticeable assessment is that the effect of age differences causes a significant increase in premiums. For example, $\bar{A}_{40:40} = 0.43874$, $\bar{A}_{40:50}$ =0.45774 and $\bar{A}_{40:60} = 0.48233$ by assuming $\rho = 0.3$ and GFGM model.

These inferences are explained in detail below with the examples given according to Table 1 and Table 2 for lives with the same age average and various correlation values, and the results are expressed as corollaries.

3.2. The Illustrative Examples

Even if the insurance policies are issued for those with the same age average, there may be significant differences in premiums due to the asymmetrical effect. In addition, premiums are also affected by the size of the correlation between lifetimes. To explain these situations, net single premiums, $A_{\overline{xy}}$ are computed for various values of ρ and average age 45. The results are given in Table 1 and exemplified below for the whole life insurance policies under the joint last survivor status. Here, it is assumed that the benefits paid at the end of the year of death are 100 000 units. Here, total premiums are denoted by P^i $(= 100\,000 \times A_{\overline{xy}})$ for $i = \{ind, FGM, GFGM\}$.

Table 1. The premiums $A_{\overline{xy}}$ for various values of ρ and ages

		Ages		
Models	ρ	(x=40,y=50)	(x=45,y=45)	(x=50,y=40)
Independence	$\rho = 0$	0.44107	0.45643	0.44107
FGM	$\rho = 0.1$	0.44686	0.46332	0.44686
	$\rho = 0.2$	0.45264	0.47022	0.45264
	$\rho = 0.3$	0.45843	0.47712	0.45843
GFGM	$\rho = 0.1$	0.44840	0.46458	0.44638
	$\rho = 0.2$	0.45572	0.47274	0.45168
	$\rho = 0.3$	0.46305	0.48089	0.45699

Example 1. Assuming by $\rho = 0.3$, net single premiums $(P^{ind}, P^{FGM}, P^{GFGM})$ are (44107,45843,46305) and (44107,45843,45699**)** for the whole life insurance policies issued for pairs of ages (40, 50) and (50, 40), respectively. Here, although the premiums of these age pairs are equal in independent and symmetric dependent models, they change by 606 units in asymmetric dependent models.

Example 2. For a whole life insurance policy issued for ages 40 and 50 with a correlation of $\rho = 0.3$, the net single premiums $(P^{ind}, P^{FGM}, P^{GFGM})$ are (44107,45843,46305), respectively. It is seen here that the premiums increased by 1736 units with the effect of symmetric dependency and 462 units more with the effect of asymmetric dependency.

Example 3. For a whole life insurance policy issued for ages 40 and 50, if the joint lifetimes are independent, the net single premium (P^{ind}) is 44107, if it is asymmetrically dependent, the net single premiums (P^{GFGM}) are 44840 and 46305 for $\rho = 0.1$ and $\rho = 0.3$, respectively. The premium is high for the policies which lifetimes have high correlations.

With these examples, it is shown that there is an increase in premiums due to the effect of asymmetry and that these increases are different according to the correlations. The results can also be explained with the premium change rates calculated by considering the age differences. For this purpose, Table 2 gives the premium change rates computed by $P_{ij} = \frac{|P^i - P^j|}{P^i}$ for $(i,j) = \{$independent (1), FGM (2), GFGM (3)$\}$

Table 2. The change ratios $A_{\overline{xy}}$ for various values of ρ and age differences.

Age	40-40		40-50		40-60	
ρ	P_{12}	P_{13}	P_{12}	P_{13}	P_{12}	P_{13}
$\rho = 0.1$	0.01857	0.02266	0.01312	0.01661	0.00661	0.00718
$\rho = 0.2$	0.03715	0.04531	0.02623	0.03321	0.01323	0.01435
$\rho = 0.3$	0.05572	0.06797	0.03934	0.04982	0.01984	0.02153

In summary, the following conclusions are drawn for the joint last survivor-life insurance policies:

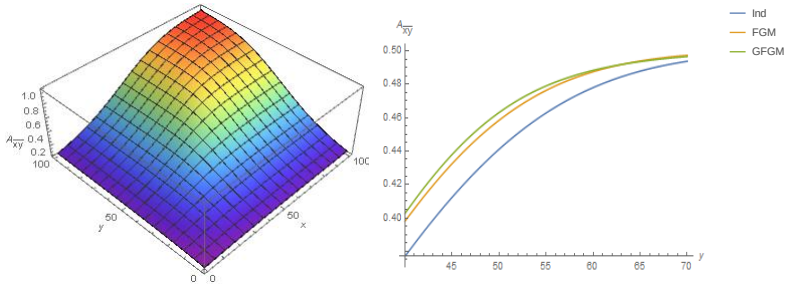
Corollary 1. In asymmetric lifetime models, premiums are affected by asymmetry (Example 1)

Corollary 2. Premiums determined for independent lifetimes increase in insurance policies with symmetrically dependent lifetimes. Premium increases even more in asymmetric policies. That is, the ranked premiums are $P^{GFGM} > P^{FGM} > P^{ind}$. It can also be seen in Figure 2b (Example 2).

Corollary 3. As the correlations between lifetimes increase, the effect of symmetry and asymmetry on premiums also increases. Indeed, the effect of asymmetry increases even more (Example 3).

Corollary 4. As the age difference increases, the symmetrical effect decreases. The asymmetrical effect is further reduced. According to premiums calculated according to independent lifetimes, the highest decrease in FGM and GFGM is observed in the insurance policies issued for lifetimes with correlation ρ = 0.3 (Table 2).

(i) The insurance policies that are most affected by asymmetry are the policies issued for equal ages with 0.3 correlated lifetimes (see Table 2, $P^{13} = 6.797\%$ for ρ = 0,3 and $(x = 40, y = 40)$).
(ii) The insurance policies least affected by asymmetry are the policies issued for the largest age difference with 0.1 correlated lifetimes (see Table 2, $P^{13} = 0.718\%$ for $\rho = 0{,}1$ and $(x = 40, y = 60)$).



(a) (b)

Figure 2. (a) The 3D plots for the GFGM and (b) the plots of $A_{\overline{xy}}$ for the proposed models ($\rho = 0.3$)

4. Conclusion

In this study, the pricing problem has been investigated for the life insurance policies under the joint last survivor status, considering by the independent and dependent lifetimes. It was assumed that the dependence lifetimes were modeled with symmetric FGM and asymmetric GFGM Type II copula. According to these situations, firstly, the joint survival and mortality functions were derived to evaluate policy prices. Then, net single premiums were numerically computed for various correlation values and ages. The results showed that premiums are influenced by dependency, even higher premiums are determined in asymmetric models. It has also been observed that premiums are sensitive to correlation size and age differences.

The evaluation of these results by the insurer will enable other actuarial decisions such as investment and reserve amounts to be made more accurately. In other words, in actuarial modeling, besides being dependent, it is also important whether lifetimes data are symmetrical or not. Therefore, firstly, it should be determined whether the lifetime data are symmetrical or not. Then, the most suitable candidate model should be selected and actuarial quantities should be calculated by one of the models.

In future studies, it will be expanded to other types of life insurance policies with first death and last survivor status. Also, results will be obtained for other asymmetric copula models that give a larger Spearman's correlation range.

Appendicies

Appendix A. As known in the literatüre, the survival functions and the forces of mortality for the Gompertz's law are denoted by

$$ {}_tp_x = e^{-\left(\frac{Bc^x(-1+c^t)}{Log[c]}\right)} \tag{A1} $$

$$ {}_t\mu_x = Bc^{x+t} \tag{A2} $$

The joint survival and mortality functions for the independence model are easily derived as follows:

First death case:

$$ {}_tp_{xy} = {}_tp_x\,{}_tp_y = e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} \tag{A3} $$

$$ {}_t\mu_{xy} = {}_t\mu_x + {}_t\mu_y = Bc^t(c^x + c^y) \tag{A4} $$

For the last survivor status, they can be computed by Eqs. (2) and (3).

Appendix B. Here, the proof of proposition 1 is given for the symmetric FGM copula under Gompertz's laws. That is, the survival (i) and mortality (ii) functions are derived by Eqs. (2 or 5) and (3), respectively.

(i) ${}_tp_{\overline{xy}} = 1 - FGM\left[1 - {}_tp_x, 1 - {}_tp_y, \theta\right]$ has the explicit Mathematica solution as follows:

$$ {}_tp_{\overline{xy}} = e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}} - Ke^{-\frac{2B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} \tag{B1} $$

where $K = \theta - e^{\frac{Bc^x(-1+c^t)}{Log[c]}}\theta - e^{\frac{Bc^y(-1+c^t)}{Log[c]}}\theta + e^{\frac{B(-1+c^t)(c^x+c^y)}{Log[c]}}(1+\theta)$ and by the Eqs. (A1) and (A3), K can be rewritten as follows

$$ K = \theta - \frac{1}{{}_tp_x}\theta - \frac{1}{{}_tp_y}\theta + \frac{1}{{}_tp_x\,{}_tp_y}(1+\theta) $$

$$ K = \theta\left(1 - \frac{1}{{}_tp_x} - \frac{1}{{}_tp_y} + \frac{1}{{}_tp_x\,{}_tp_y}\right) + \frac{1}{{}_tp_x\,{}_tp_y} $$

$$ K = \frac{1}{{}_tp_{xy}}\left(\theta\left({}_tp_{xy} - {}_tp_y - {}_tp_x + 1\right) + 1\right) $$

Therefore Eq. (B1) is formed as below

$$ {}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - K\left({}_tp_{xy}\right)^2 $$

$$ {}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - {}_tp_{xy}\left(\theta\left({}_tp_{xy} - {}_tp_y - {}_tp_x + 1\right) + 1\right) $$

$$ {}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - {}_tp_{xy}\theta\left({}_tp_{xy} - {}_tp_y - {}_tp_x\right)\big) - (\theta+1)\,{}_tp_{xy} $$

$$
{}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - (\theta+1){}_tp_{xy} - \theta\left(\left({}_tp_{xy}\right)^2 - {}_tp_x\left({}_tp_y\right)^2 - \left({}_tp_x\right)^2{}_tp_y\right)
$$

$$
{}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - (\theta+1){}_tp_{xy} - \theta\left({}_tp_{xy}\right)^2\left(1 - {}_tp_x - {}_tp_y\right)
$$

$$
{}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - {}_tp_{xy} - \theta\,{}_tp_{xy} - \theta\left({}_tp_{xy}\right)^2\left(1 - {}_tp_x - {}_tp_y\right).
$$

(ii) The joint mortality function is derived by Mathematica solution using Eq. (3) as follows:

$$
{}_t\mu_{\overline{xy}} = -\frac{1}{{}_tp_{\overline{xy}}}\left(\left(\left(Bc^t\left(-c^x e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} - c^y e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}\right.\right.\right.\right.
$$

$$
+2K(c^x+c^y)e^{-\frac{2B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} + e^{-\frac{2B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}}\left(c^x e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}}\theta\right.
$$

$$
\left.\left.\left.\left.+c^y e^{\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}\theta - (c^x+c^y)e^{\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}}(1+\theta)\right)\right)\right)
$$

$$
= \frac{1}{{}_tp_{\overline{xy}}}\left(\left(\left(Bc^t c^x{}_tp_x Bc^t c^y{}_tp_y - 2KBc^t(c^x+c^y)\left({}_tp_{xy}\right)^2 -\right.\right.\right.
$$

$$
\left.\left.\left.\left({}_tp_{xy}\right)^2\left(c^x Bc^t\frac{1}{{}_tp_x}\theta + c^y Bc^t\frac{1}{{}_tp_y}\theta - Bc^t(c^x+c^y)\frac{1}{{}_tp_{xy}}(1+\theta)\right)\right)\right)\right)
$$

by substituting Eqs. (A2) -(A3) and K

$$
= \frac{1}{{}_tp_{\overline{xy}}}\left({}_t\mu_x\,{}_tp_x + {}_t\mu_y\,{}_tp_y - 2K\,{}_t\mu_{xy}\left({}_tp_{xy}\right)^2 - \left({}_tp_{xy}\right)^2\left(\frac{{}_t\mu_x}{{}_tp_x}\theta + \frac{{}_t\mu_y}{{}_tp_y}\theta - \frac{{}_t\mu_{xy}}{{}_tp_{xy}}(1+\theta)\right)\right).
$$

Appendix C. Here, the proof of proposition 2 is given for the asymmetric GFGM copula under Gompertz's laws. That is, the survival (i) and mortality (ii) functions are derived by Eqs. (2 or 5) and (3), respectively.

(i) ${}_tp_{\overline{xy}} = 1 - GFGM\left[1 - {}_tp_x, 1 - {}_tp_y, \theta, b, \alpha, \beta\right]$ has the explicit Mathematica solution as follows:

$$
{}_tp_{\overline{xy}} = S - \mathrm{G}
$$

where $S = e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}} - e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}}$ and

$$
G = \left(e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}}\right)^{\alpha}\left(e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}\right)^{\beta}\left(1 - e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}}\right)^{b}\left(1 - e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}\right)^{b}\theta
$$

By using the Eq. $(A1)$, S and G can be rewritten as follows:

$$S = {}_tp_x + {}_tp_y - {}_tp_{xy}$$

$$G = \theta({}_tp_x)^{\alpha}({}_tp_y)^{\beta}(1 - {}_tp_x)^{b}(1 - {}_tp_y)^{b}$$

Thus ${}_tp_{\overline{xy}} = S - \mathrm{G}$

$${}_tp_{\overline{xy}} = {}_tp_x + {}_tp_y - {}_tp_{xy} - \theta({}_tp_x)^{\alpha}({}_tp_y)^{\beta}(1 - {}_tp_x)^{b}(1 - {}_tp_y)^{b}$$

(ii) The joint mortality function is derived by Mathematica solution using Eq. (3) as follows

$$\begin{aligned}{}_t\mu_{\overline{xy}} = -\left(\frac{1}{\mathrm{S}-\mathrm{G}}\right)(Bc^t(c^x(-e^{-\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} + e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} \\ + \frac{G}{-1+e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}}}(-b + (-1+e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}})\alpha)) \\ + c^y(e^{-\frac{B(-1+c^t)(c^x+c^y)}{\mathrm{Log}[c]}} + G\beta + e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}}(-1 \\ - b(1-e^{-\frac{Bc^y(-1+c^t)}{\mathrm{Log}[c]}})^{-1}G))))\end{aligned}$$

Let $k_x = -1 + e^{\frac{Bc^x(-1+c^t)}{\mathrm{Log}[c]}} = -1 + \frac{1}{{}_tp_x} = \frac{{}_tq_x}{{}_tp_x}$.

$$\begin{aligned}&= -\left(\frac{1}{{}_tp_{\overline{xy}}}\right)(Bc^t(c^x(-{}_tp_x + {}_tp_{xy} - b\mathrm{G}\frac{{}_tp_x}{{}_tq_x} + G\alpha) + c^y({}_tp_{xy} + G\beta \\ &\qquad - {}_tp_y - b\frac{{}_tp_y}{{}_tq_y}G)))\end{aligned}$$

$$\begin{aligned}&= -\left(\frac{1}{{}_tp_{\overline{xy}}}\right)(Bc^t(c^x(-{}_tp_x\,{}_tq_y - G(b\frac{{}_tp_x}{{}_tq_x} - \alpha) + c^y(-{}_tp_y\,{}_tq_x \\ &\qquad - G(b\frac{{}_tp_y}{{}_tq_y} - \beta)))\end{aligned}$$

$$= \left(\frac{1}{{}_tp_{\overline{xy}}}\right)(({}_t\mu_x({}_tp_x\,{}_tq_y + G(b\frac{{}_tp_x}{{}_tq_x} - \alpha) + {}_t\mu_y({}_tp_y\,{}_tq_x + G(b\frac{{}_tp_y}{{}_tq_y} - \beta))).$$

References

Bakar, Ö. (2018). *Stochastic Analysis of Longevity Risk in Dependent Multiple Life Annuities,* (Master's thesis, Graduate School of Science and Engineering).

Carriere, J. F. (2000). Bivariate survival models for coupled lives. *Scandinavian Actuarial Journal,* 2000, 17–32.

Chen, H., MacMinn, R. D., and Sun, T. (2015) 'Multi-Population Mortality Models: A Factor Copula Approach', *Insurance: Mathematics and Economics* 63:135–146.

Denuit, M., & Cornet, A. (1999). Multilife premium calculation with dependent future lifetimes. *Journal of Actuarial Practice,* 1999 (7), 147-180.

Dickson, D. C., Hardy, M., Hardy, M. R. and Waters, H. R. (2013). *Actuarial mathematics for life contingent risks*, Cambridge University Press.

Dufresne, F., Hashorva, E., Ratovomirija, G. and Toukourou, Y. (2018). On age difference in joint lifetime modelling with life insurance annuity applications, *Annals of Actuarial Science,* 12(2), 350-371.

Frees, E. W., Carriere, J., & Valdez, E. (1996). Annuity valuation with dependent mortality *Journal of Risk and Insurance,* 229-261.

Frees, E. W., & Valdez, E. A. (1998). Understanding relationships using copulas *North American actuarial journal,* 2(1), 1-25.

Gompertz, B. (1825). XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. *Philosophical transactions of the Royal Society of London,* (115), 513-583.

Hsieh, M. H., Tsai, C. J., & Wang, J. L. (2020). Mortality Risk Management Under the Factor Copula Framework-With Applications to Insurance Policy Pools. *North American Actuarial Journal,* 1-13.

Jagger, C., & Sutton, C. J. (1991). Death after marital bereavement-is the risk increased? *Statistics in Medicine,* 10(3), 395-404.

Jung, Y. S., Kim, J. M., & Kim, J. (2008). New approach of directional dependence in exchange markets using generalized fgm copula function. *Communications in Statistics—Simulation and Computation,* 37(4), 772-788.

Kara, E. K. (2021) A Study on Modeling of Lifetime with Right-Truncated Composite Lognormal-Pareto Distribution: Actuarial Premium Calculations. *Gazi University Journal of Science,* 34(1), 272-288 (https://doi.org/10.35378/gujs.646899).

Lee, H., & Cha, J. H. (2018). A dynamic bivariate common shock model with cumulative effect and its actuarial application. *Scandinavian Actuarial Journal*, 2018(10), 890-906.

Lee, I., Lee, H., & Kim, H. T. (2014). Analysis of reserves in multiple life insurance using copula. *Communications for Statistical Applications and Methods*, 21(1), 23-43.

Lu, Y. (2017). Broken-heart, common life, heterogeneity: Analyzing the spousal mortality dependence. *ASTIN Bulletin: The Journal of the IAA*, 47(3), 837-874.

Luciano, E., Spreeuw, J., & Vigna, E. (2008). Modelling stochastic mortality for dependent lives. *Insurance: Mathematics and Economics*, 43(2), 234-244.

Menge, W. O., & Glover, J. W. (1938). *An introduction to the mathematics of life insurance.* Macmillan.

Nelsen, R. B. (2007). *An introduction to copulas.* Springer Science & Business Media.

Riaman, S., Supian, S. & Bon, A.T. (2019). Analysis of Determination of Adjusted Premium Reserves for Last Survivor Endowment Life Insurance Using the Gompertz Assumption. *IEOM Society International*. 430-438.

Rodriguez-Lallena, J. A., & Úbeda-Flores, M. (2004). A new class of bivariate copulas. *Statistics & probability letters*, 66(3), 315-325.

Shemyakin, A. E., & Youn, H. (2006). Copula models of joint last survivor analysis. *Applied Stochastic Models in Business and Industry*, 22(2), 211-224.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ*. Paris, 8, 229-231.

Terzioglu, M. K., & Sucu, M. (2015). Gompertz–Makeham parameter estimations and valuation approaches: Turkish life insurance sector. *European Actuarial Journal*, 5(2), 447-468.

Zhu, W., Tan, K. S., & Wang, C. W. (2017). Modeling multicountry longevity risk with mortality dependence: a Lévy subordinated hierarchical Archimedean copulas approach. *Journal of Risk and Insurance*, 84(S1), 477-493.

CHAPTER IV

ON MODIFIED *m*-SINGULAR GAUSS-WEIERSTRASS OPERATORS

Gumrah Uysal[1] & Basar Yilmaz[2]

[1](Dr), Karabuk University, e-mail: fgumrahuysal@gmail.com
ORCID: 0000-0001-7747-1706

[2](Dr), Kirikkale University, e-mail: basaryilmaz77@yahoo.com
ORCID: 0000-0003-3937-992X

1. Introduction

The approximation by singular-type integral operators is a topic that has been studied for many years and has grown its popularity over the years. We recommend the monographs (Altomare and Campiti, 1994; Anastassiou and Gal, 2000; Butzer and Nessel, 1971; Stein, 1970) for the interested reader who wants to get more detailed information on this topic. In particular, Gauss-Weierstrass operators have immeasurable significance in this sense. In fact, Gauss-Weierstrass operators have been investigated in many different directions with unpredictable speed. Some examples of these studies can be found in the works (Rempulska, and Walczak, 2005; Anastassiou and Aral, 2010; Yilmaz, 2019; Agratini et al., 2017; Yilmaz and Uysal, 2020; Yilmaz, 2020).

The approximation in exponential weighted spaces has gained great importance after Becker et al. (1978) presented their work containing the concept of the indicated spaces. Some studies centered on this topic can be given as (Deeba et al., 1988; Lèsniewicz et al., 1996; Rempulska and Tomczak, 2009; Uysal, 2019; Aral et al., 2020; Uysal, 2020; Yilmaz, 2020).

Throughout this chapter, let $\mathcal{R}$ denote the set of all real numbers ($\mathcal{R} = (-\infty, +\infty)$) and N denote the set of all positive integers. In view of finite

differences, Mamedov (1963) designated the following $m-$singular integral operators:

$$\mathrm{T}_{\mu}^{[m]}(h;u)$$

$$= (-1)^{m+1}\int_{-\infty}^{+\infty}\left[\sum_{i=1}^{m}(-1)^{m-i}\binom{m}{i}h(u+it)\right]K_{\mu}(t)dt, \qquad (1.1)$$

where $u,t\in\mathcal{R}$. In equation (1.1), m is a certain positive integer and the generalized kernel function K_{μ} with $K_{\mu}:\mathcal{R}\to\mathcal{R}$, where μ stands for a positive parameter which belongs to the index set P, is enriched with some determinative properties analogous to the well-known approximate identity-type kernels. These operators generalize the already existing convolutional integrals for $m=1$ through allowing the number m to make impact on the operators based on the construction of the sum. Such generalizations provide high utility, for example, in the approximation of some classes of smooth and (weighted) integrable functions. This can be easily seen by making comparison between increasing values of the number m. Later, Mamedov (1991) defined also Mellin analogue of the operators of type (1.1) having beneficial convergence aspects. Some works which are related to m-singularity concept can be given as (Rydzewska, 1978; Bardaro et al. 2013; Karsli, 2014; Uysal, 2018; Uysal, 2019). Jackson type and/or Jackson-Stechkin type generalizations of the operators are related to this type generalization, for example, Yilmaz (2014) obtained some results concerning convergence in variation for Jackson type generalization of nonlinear convolution operators considered by Angeloni and Vinti (2006). In addition, Rydzewska (1978) gave some pointwise convergence results for Jackson-Stechkin type one-dimensional singular integrals.

Agratini et al. (2017) introduced the following modification of Gauss-Weierstrass operators:

$$W_n^*(h;x)=\frac{\sqrt{n}}{\sqrt{\pi}}\int_{-\infty}^{+\infty}h(\beta_n(x)+t)\,e^{-nt^2}dt, \qquad (1.2)$$

where $x,t\in\mathcal{R}$, $n\in\mathrm{N}$ and $\beta_n(x)=x-\frac{b}{2n}$ with $b>0$. Here, the functions h were chosen to ensure that the operators are well-defined. The

modification of Picard operators were also defined in the same work. The operators in equation (1.2) fix constant functions and e^{2bx} with $b > 0$ and $x \in \mathcal{R}$. Also, if we consider to take $b = 0$ for a second, usual Gauss-Weierstrass operators W_n are recovered (see, for example, (Agratini et al., 2017; Yilmaz, 2019); see also (Butzer and Nessel, 1971)). This study covers numerous fundamental and important results. Also, the mentioned study provides the basis for many studies as it involves new operators. Some of the studies based on this study can be given as (Aral et al. 2020; Yilmaz and Ari, 2020; Yilmaz, 2020). Also, based on the works by Mamedov (1963), Agratini et al. (2017) and Aral et al. (2020), Uysal (2019) studied m −singular generalizations of modified Picard operators.

Aral (2018) constructed a modification of Picard operators preserving some exponetial functions. Uysal (2020) studied the operators defined by Aral (2018) in exponential weighted spaces.

By mixing the operators (1.1) and (1.2), the following modified m −singular Gauss-Weierstrass operators are set:

$$W_n^{*[m]}(h;x) = \frac{\sqrt{n}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \left[\sum_{i=1}^{m} (-1)^{i-1} \binom{m}{i} h(\beta_n(x) + it) \right] e^{-nt^2} dt, \qquad (1.3)$$

where m is a certain positive integer, $x, t \in \mathcal{R}$, $n \in$ N and $\beta_n(x) = x - \frac{b}{2n}$ with $b > 0$. By taking $m = 1$, the operators in equation (1.3) turn out to be the operators in equation (1.2).

In the end of the article by Rempulska and Tomczak (2009), the authors state that the results which are similar to that of they proved can be obtained for Gauss-Weierstrass operators by taking the weight function w_b as $w_b(x) = e^{-bx^2}$ with $w_b: \mathcal{R} \to \mathcal{R}^+$ and $b > 0$ for the space $L_{p,b}(\mathcal{R})$. In view of this idea and referring to (Becker et al., 1978) (see also (Rempulska and Tomczak, 2009)), the exponential weighted Lebesgue spaces $L_{p,b}(\mathcal{R})$ with $b > 0$ and $1 \leq p < \infty$ are formed of all functions $h: \mathcal{R} \to \mathcal{R}$ which are measurable on $\mathcal{R}$ in the sense of Lebesgue such that $\int_{-\infty}^{+\infty} |w_b(x)h(x)|^p dx < \infty$. The norm of the function in the space $L_{p,b}(\mathcal{R})$ is given by the following formula:

$$\|h\|_{p,b} = \left(\int_{-\infty}^{+\infty} |w_b(x)h(x)|^p dx \right)^{1/p} \quad (1 \leq p < \infty).$$

As in (Yilmaz, 2020), in the current work, we consider the spaces $L_{p,b}(\mathcal{R})$ with respect to the definition given above (for more details, see (Becker et al., 1978; Rempulska and Tomczak, 2009)). We replace $w_b(x)$ by e^{-bx^2} with $b > 0$ for the space $L_{p,b}(\mathcal{R})$ and by e^{-2bx^2} with $b > 0$ for the space $L_{p,2b}(\mathcal{R})$.

The current work is an extention of (Yilmaz, 2020) and contains some useful generalizations of the results of indicated work.

2. Existence of the Operators

Now, we prove a generalization of existence lemma stated in (Yilmaz, 2020) for *m*-singular modified Gauss-Weierstrass operators.

Lemma 2.1. Let $n > 2bm^2$. If $h \in L_{p,b}(\mathcal{R})$ with $b > 0$ and $1 \leq p < \infty$, then $W_n^{*[m]}h \in L_{p,2b}(\mathcal{R})$ and the inequality:

$$\|W_n^{*[m]}h\|_{p,2b} \leq \sum_{i=1}^{m} \binom{m}{i} e^{\frac{b^3}{2n^2-4bi^2n}} \frac{\sqrt{n}}{\sqrt{n-2bi^2}} \|h\|_{p,b}$$

holds.

Proof. Using the norm definition in the space $L_{p,2b}(\mathcal{R})$ with $1 < p < \infty$, we have

$$\|W_n^{*[m]}h\|_{p,2b} = \left(\int_{-\infty}^{+\infty} \left| e^{-2bx^2} \left(W_n^{*[m]}h \right)(x) \right|^p dx \right)^{1/p}.$$

In view of generalized Minkowski inequality (see, for example, (Stein, 1970)) for $1 < p < \infty$ and using change of variables, we have

$$\|W_n^{*[m]}h\|_{p,2b}$$

$$\leq \frac{\sqrt{n}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-nt^2} \left(\int_{-\infty}^{+\infty} \left| \sum_{i=1}^{m} \binom{m}{i} \left| \frac{h(\beta_n(x)+it)}{e^{b(x-\frac{b}{2n}+it)^2}} \right| \frac{e^{b(x-\frac{b}{2n}+it)^2}}{e^{2bx^2}} \right|^p dx \right)^{1/p} dt$$

$$\leq \frac{\sqrt{n}}{\sqrt{\pi}} \sum_{i=1}^{m} \binom{m}{i} \left(\int_{-\infty}^{+\infty} \left| \frac{h(u)}{e^{bu^2}} \right|^p du \right)^{1/p} \int_{-\infty}^{+\infty} e^{2b\left(\frac{b}{2n}-it\right)^2} e^{-nt^2} dt$$

$$\leq \sum_{i=1}^{m} \binom{m}{i} e^{\frac{b^3}{2n^2-4bi^2n}} \frac{\sqrt{n}}{\sqrt{n-2bi^2}} \text{ll}h\text{ll}_{p,b}, \text{ for all } n > 2bm^2.$$

Using the norm definition in the space $L_{1,2b}(\mathcal{R})$, we have

$$\text{ll}W_n^{*[m]}h\text{ll}_{1,2b} = \int_{-\infty}^{+\infty} \left| e^{-2bx^2} \left(W_n^{*[m]}h \right)(x) \right| dx.$$

In view of Fubini's theorem (see, for example, (Butzer and Nessel, 1971)) and using change of variables, we have

$$\text{ll}W_n^{*[m]}h\text{ll}_{1,2b} \leq \sum_{i=1}^{m} \binom{m}{i} e^{\frac{b^3}{2n^2-4bi^2n}} \frac{\sqrt{n}}{\sqrt{n-2bi^2}} \text{ll}h\text{ll}_{1,b},$$

where $n > 2bm^2$.

■

3. Main Result

In this section, we will state and prove a pointwise convergence theorem similar to the theorem proved in (Aral et al., 2020; Yilmaz, 2020). The following definition is expressed by blending the definition of $m-p$ −Lebesgue point given in (Mamedov, 1963) and a weighted Lebesgue point definition given in (Yilmaz, 2020).

Definition 3.1. A point $x \in \mathcal{R}$ at which the limit relation:

$$\lim_{k\to 0^+} \frac{1}{k} \int_0^k \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it) + h(\beta_n(x)-it) - 2h(x)}{e^{2b(it)^2}} \right] \right|^p dt$$

$$= 0 \qquad (3.1)$$

holds, is called weighted $m-p$ −Lebesgue point of $h \in L_{p,b}(\mathcal{R})$ with $b > 0$ and $1 \leq p < \infty$.

Generalized version of pointwise convergence theorem given in (Yilmaz, 2020) for modified m −singular Gauss-Weierstrass operators is as follows:

Theorem 3.1. Let $n > 2bm^2$, $1 \le p < \infty$ and $b > 0$. If $h \in L_{p,b}(\mathcal{R})$, then

$$\lim_{n\to+\infty} \left| W_n^{*[m]}(h;x) - h(x) \right| = 0$$

holds at each weighted $m-p$ −Lebesgue point $x \in \mathcal{R}$ of $h \in L_{p,b}(\mathcal{R})$ for which (3.1) holds.

Proof. Let $p = 1$. In view of the well-known fact written as

$$2\int_0^{+\infty} e^{-nu^2} du = \frac{\sqrt{\pi}}{\sqrt{n}}, \qquad n \in \mathrm{N},$$

we may write

$$W_n^{*[m]}(h;x) - h(x)$$
$$= \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{+\infty} \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it) + h(\beta_n(x)-it) - 2h(x)}{e^{2b(it)^2}} \right] \times e^{-t^2(n-2bi^2)} dt.$$

In view of (3.1), for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\int_0^k \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right| dt \le \varepsilon k, \qquad (3.2)$$

where $0 < k \le \delta$, holds.

Next, the following inequality can be written with respect to $\delta > 0$ found above:

$$\left| W_n^{*[m]}(h;x) - h(x) \right|$$
$$\le \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it) + h(\beta_n(x)-it) - 2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bm^2)} dt$$
$$+ \frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it) + h(\beta_n(x)-it) - 2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bi^2)} dt$$

$$=: J_{[1]} + J_{[2]},$$

where

$$J_{[1]} = \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x) + it) + h(\beta_n(x) - it) - 2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bm^2)} dt$$

and

$$J_{[2]} = \frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x) + it) + h(\beta_n(x) - it) - 2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bi^2)} dt.$$

We continue with $J_{[1]}$ in the following fashion:

$$J_{[1]} = \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bm^2)} dt$$

$$= \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} e^{-t^2(n-2bm^2)} d\mathcal{F}(t),$$

where

$$\mathcal{F}(t) := \int_0^t \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+iw)+h(\beta_n(x)-iw)-2h(x)}{e^{2b(iw)^2}} \right] \right| dw.$$

Applying method of integration by parts to $J_{[1]}$ and using relation (3.2), we have

$$|J_{[1]}| = \left| \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} e^{-t^2(n-2bm^2)} d\mathcal{F}(t) \right|$$

$$= \left| \frac{\sqrt{n}}{\sqrt{\pi}} e^{-\delta^2(n-2bm^2)} \mathcal{F}(\delta) - \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} \mathcal{F}(t) d[e^{-t^2(n-2bm^2)}] \right|$$

$$\leq \frac{\sqrt{n}}{\sqrt{\pi}} e^{-\delta^2(n-2bm^2)} |\mathcal{F}(\delta)| + \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} |\mathcal{F}(t)| d[-e^{-t^2(n-2bm^2)}]$$

$$\leq \frac{\sqrt{n}}{\sqrt{\pi}} e^{-\delta^2(n-2bm^2)} \varepsilon\delta + \varepsilon \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} t d\left[-e^{-t^2(n-2bm^2)}\right].$$

One more application of integration by parts leads to

$$\left|J_{[1]}\right|$$

$$\leq \varepsilon \frac{\sqrt{n}}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-t^2(n-2bm^2)} dt$$

$$= \varepsilon \frac{\sqrt{n}}{\sqrt{\pi}} \frac{\sqrt{\pi}}{\sqrt{n-2bm^2}}$$

$$= \varepsilon \left(\frac{\sqrt{n}}{\sqrt{n-2bm^2}} \right), \quad n > 2bm^2.$$

Therefore, $\left|J_{[1]}\right| \to 0$ as $n \to +\infty$ with $n > 2bm^2$.

Now, it remains to show that $\left|J_{[2]}\right| \to 0$ as $n \to +\infty$. Observe that

$$J_{[2]}$$

$$= \frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right| \times e^{-t^2(n-2bi^2)} dt$$

$$\leq \sum_{i=1}^{m} \binom{m}{i} \frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \left| \frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right| e^{-t^2(n-2bi^2)} dt$$

$$\leq \sum_{i=1}^{m} \binom{m}{i} \frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \frac{|h(\beta_n(x)+it)|+|h(\beta_n(x)-it)|}{e^{2b(it)^2}} e^{-t^2(n-2bm^2)} dt$$

$$+ \frac{2\sqrt{n}}{\sqrt{\pi}} 2^m |h(x)| \int_{\delta}^{+\infty} e^{-nt^2} dt$$

$$\leq \sum_{i=1}^{m} \binom{m}{i} \frac{\sqrt{n}}{\sqrt{\pi}} e^{-\delta^2(n-2bm^2)} \int_{\delta}^{+\infty} \frac{|h(\beta_n(x)+it)|}{e^{b(\beta_n(x)+it)^2}} \frac{e^{2b\left((\beta_n(x))^2+(it)^2\right)}}{e^{2b(it)^2}} dt$$

$$+ \sum_{i=1}^{m} \binom{m}{i} \frac{\sqrt{n}}{\sqrt{\pi}} e^{-\delta^2(n-2bm^2)} \int_{\delta}^{+\infty} \frac{|h(\beta_n(x)-it)|}{e^{b(\beta_n(x)-it)^2}} \frac{e^{2b\left((\beta_n(x))^2+(it)^2\right)}}{e^{2b(it)^2}} dt$$

$$+ \frac{2\sqrt{n}}{\sqrt{\pi}} 2^m |h(x)| \int_{\delta}^{+\infty} e^{-nt^2} dt$$

$$\leq \frac{2\sqrt{n}}{\sqrt{\pi}} 2^m \left\{ e^{-\delta^2(n-2bm^2)} e^{2b(\beta_n(x))^2} \text{ll}h\text{ll}_{1,b} + |h(x)| \int_\delta^{+\infty} e^{-nt^2}\, dt \right\}.$$

Since $\sqrt{n} e^{-\delta^2(n-2bm^2)} \to 0$ with $n > 2bm^2$ for every $\delta > 0$ as $n \to +\infty$, and $\sqrt{n} \int_\delta^{+\infty} e^{-nt^2}\, dt \to 0$ as $n \to +\infty$ for every $\delta > 0$, the claim follows for the case $p = 1$.

Now, let $p > 1$. In view of (3.1), for every $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\int_0^k \left| \sum_{i=1}^m \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right|^p dt \leq \varepsilon k, \quad (3.3)$$

where $0 < k \leq \delta$, holds.

Next, the following inequality can be obtained with respect to $\delta > 0$ found above:

$$\left| W_n^{*[m]}(h;x) - h(x) \right|^p$$

$$\leq 2^p \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_0^\delta \left| \sum_{i=1}^m \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right| \times \right.$$
$$\left. \times e^{-t^2(n-2bm^2)} dt \right)^p$$

$$+ 2^p \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_\delta^{+\infty} \left| \sum_{i=1}^m \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right| \times \right.$$
$$\left. \times e^{-t^2(n-2bi^2)} dt \right)^p$$

$$=: 2^p \left(I_{[1]} + I_{[2]} \right).$$

Via using Hölder's inequality (see, for example, (Rudin, 1987)) in $I_{[1]}$ with $\frac{1}{p} + \frac{1}{q} = 1$, where $1 < p, q < \infty$, we proceed as follows:

$$I_{[1]} \leq \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_0^\delta \left| \sum_{i=1}^m \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right|^p \right.$$
$$\left. \times e^{-t^2(n-2bm^2)} dt \right)$$

$$\times \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_0^\delta e^{-t^2(n-2bm^2)} dt \right)^{\frac{p}{q}}$$

$$\leq \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_0^\delta \left| \sum_{i=1}^m \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right|^p \right.$$
$$\left. \times e^{-t^2(n-2bm^2)} dt \right)$$

$$\times \left(\frac{\sqrt{n}}{\sqrt{n-2bm^2}} \right)^{\frac{p}{q}}$$

$=: I_{[1,1]} \times \left(\frac{\sqrt{n}}{\sqrt{n-2bm^2}}\right)^{\frac{p}{q}}, \quad n > 2bm^2,$

where

$$I_{[1,1]} = \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+it)+h(\beta_n(x)-it)-2h(x)}{e^{2b(it)^2}} \right] \right|^p \times e^{-t^2(n-2bm^2)} dt.$$

With the help of the function G defined by

$$G(t) := \int_0^t \left| \sum_{i=1}^{m} \binom{m}{i} (-1)^{i-1} \left[\frac{h(\beta_n(x)+is)+h(\beta_n(x)-is)-2h(x)}{e^{2b(is)^2}} \right] \right|^p ds,$$

we may write

$$\left|I_{[1,1]}\right| = \left| \frac{\sqrt{n}}{\sqrt{\pi}} \int_0^{\delta} e^{-t^2(n-2bm^2)} dG(t) \right|.$$

In view of (3.3) and processing as in the case of $p = 1$, we easily obtain that

$$\left|I_{[1,1]}\right| \leq \varepsilon \left(\frac{\sqrt{n}}{\sqrt{n-2bm^2}}\right), \quad n > 2bm^2.$$

Thus, $\left|I_{[1]}\right| \to 0$ as $n \to +\infty$ with $n > 2bm^2$.

Recalling $I_{[2]}$, we see that the following inequality holds:

$$I_{[2]}$$

$$\leq 2^p \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \sum_{i=1}^{m} \binom{m}{i} \left(|h(\beta_n(x)+it) + h(\beta_n(x)-it)|\right) e^{-nt^2} dt \right)^p$$

$$+ 2^{2p} |h(x)|^p 2^{mp} \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} e^{-nt^2} dt \right)^p$$

$$=: 2^p I_{[2,1]} + 2^{(m+2)p} |h(x)|^p I_{[2,2]}.$$

Now, considering $I_{[2,1]}$, we get

$$I_{[2,1]}$$

$$\leq \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \sum_{i=1}^{m} \binom{m}{i} \frac{\left(|h(\beta_n(x)+it)+h(\beta_n(x)-it)|\right)}{e^{2b(it)^2}} e^{-t^2(n-2bm^2)} dt \right)^p$$

$$\leq 2^p \left(\frac{\sqrt{n}}{\sqrt{\pi}} \int_{\delta}^{+\infty} \sum_{i=1}^{m} \binom{m}{i} \frac{|h(\beta_n(x)+it)|}{e^{2b(it)^2}} e^{-t^2(n-2bm^2)} dt \right)^p$$

$$+2^p\left(\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)-it)|}{e^{2b(it)^2}}e^{-t^2(n-2bm^2)}dt\right)^p.$$

Using Hölder's inequality with $\frac{1}{p}+\frac{1}{q}=1$, where $1<p,q<\infty$, we may deal with $I_{[2,1]}$ as follows:

$$I_{[2,1]}$$

$$\leq 2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)+it)|}{e^{2b(it)^2}}\right)^p e^{-t^2(n-2bm^2)}dt$$

$$+2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)-it)|}{e^{2b(it)^2}}\right)^p e^{-t^2(n-2bm^2)}dt$$

$$\leq 2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)+it)|}{e^{b(\beta_n(x)+it)^2}}\right)^p \frac{e^{bp(\beta_n(x)+it)^2}}{e^{2bp(it)^2}}e^{-t^2(n-2bm^2)}dt$$

$$+2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)-it)|}{e^{b(\beta_n(x)-it)^2}}\right)^p \frac{e^{bp(\beta_n(x)-it)^2}}{e^{2bp(it)^2}}e^{-t^2(n-2bm^2)}dt$$

$$\leq 2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)+it)|}{e^{b(\beta_n(x)+it)^2}}\right)^p e^{2bp(\beta_n(x))^2}e^{-t^2(n-2bm^2)}dt$$

$$+2^p a[n]\frac{\sqrt{n}}{\sqrt{\pi}}\int_\delta^{+\infty}\left(\sum_{i=1}^m\binom{m}{i}\frac{|h(\beta_n(x)-it)|}{e^{b(\beta_n(x)-it)^2}}\right)^p e^{2bp(\beta_n(x))^2}e^{-t^2(n-2bm^2)}dt$$

$$\leq 2^{p+1}2^{mp}a[n]\frac{\sqrt{n}}{\sqrt{\pi}}e^{-\delta^2(n-2bm^2)}e^{2bp(\beta_n(x))^2}(\mathrm{ll}h\mathrm{ll}_{p,b})^p$$

for all n with $n>2bm^2$, where $a[n]$ stands for the term $\left(\frac{\sqrt{n}}{\sqrt{n-2bm^2}}\right)^{\frac{p}{q}}$.

Since $2^pI_{[2,1]}+2^{(m+2)p}|h(x)|^pI_{[2,2]}\to 0$ as $n\to+\infty$ with $n>2bm^2$, the desired conclusion for $p>1$ is finally obtained.

■

Example 3.1. Let $h:\mathcal{R}\to\mathcal{R}^+$ be defined by $h(t)=e^{-t^2}$. Table 3.1 demonstrates how the approximation by $W_n^{*[m]}$, W_n^* and W_n change with respect to the different values of n, m and b at $x=0$. The results in Table 3.1 are obtained by using Mathematica 12.2.

Table 3.1: Demonstration of Example 3.1

n	$m=1$ $b=1$ $W_n^*(h;0)$	$m=10$ $b=1$ $W_n^{*[10]}(h;0)$	$m=1$ $b=8$ $W_n^*(h;0)$	$m=10$ $b=8$ $W_n^{*[10]}(h;0)$	$m=1$ $b=0$ $W_n(h;0)$
10	0.95129	0.98210	0.82439	0.84970	0.95346
40	0.98757	1.00036	0.97814	0.99054	0.98772
70	0.99288	0.99998	0.98974	0.99677	0.99293
100	0.99501	0.99997	0.99346	0.99839	0.99503
130	0.99616	0.99998	0.99524	0.99904	0.99617

4. Conclusion

In this work, $m-$singular modifications of the operators defined in equation (1.2) are studied and a pointwise approximation result is proved. The operators in equation (1.3) indeed give better results for some classes of functions with respect to some higher values of m. This situation is struggled to be exemplified in Table 3.1.

Singular integral operators, particularly the Gauss-Weierstrass operators, are very common tools in engineering applications. In order to see this, it is enough to look at the abundance of the studies on convolutional neural networks in the literature and how often they are used in image reconstruction and related current life issues. Therefore, the newly defined operators may do a good contribution in this field.

In the future, the operators in equation (1.3) can be examined in many different directions, especially in the sense of Korovkin-type approximation.

References

Agratini, O., Aral, A. and Deniz, E. (2017). On two classes of approximation processes of integral type. Positivity 21(3), 1189–1199.

Altomare, F. and Campiti, M. (1994). Korovkin-type Approximation Theory and its Applications. De Gruyter Studies in Mathematics, 17. Walter de Gruyter & Co., Berlin.

Anastassiou, G.A. and Gal, S.G. (2000). Approximation Theory. Moduli of Continuity and Global Smoothness Preservation. Birkhäuser Boston, Inc., Boston, MA.

Anastassiou, G.A. and Aral, A. (2010). On Gauss-Weierstrass type integral operators. Demonstratio Math. 43(4), 841–849.

Angeloni, L. and Vinti, G. (2006). Convergence in variation and rate of approximation for nonlinear integral operators of convolution type. Results Math. 49(1-2), 1-23. (Erratum: Results Math. 57(3-4) (2010) 387-391)

Aral, A. (2018). On generalized Picard integral operators. In: Advances in Summability and Approximation Theory. Springer, Singapore, 157–168.

Aral, A., Yilmaz, B. and Deniz, E. (2020). Weighted approximation by modified Picard operators. Positivity 24(2), 427-439.

Bardaro, C., Karsli, H. and Vinti, G. (2013). On pointwise convergence of Mellin type nonlinear *m*-singular integral operators. Comm. Appl. Nonl. Anal. 20(2), 25–39.

Becker, M., Kucharski, D. and Nessel, R.J. (1978). Global approximation theorems for the Szász-Mirakjan operators in exponential weight spaces. In: Linear Spaces and Approximation (Proc. Conf., Math. Res. Inst., Oberwolfach, 1977), Internat. Ser. Numer. Math., Vol. 40, Birkhäuser, Basel, 319-333.

Butzer, P.L. and Nessel, R.J. (1971). Fourier Analysis and Approximation. Volume 1: One-dimensional Theory. Pure and Applied Mathematics, Vol. 40. Academic Press, New York, London.

Deeba, E., Mohapatra, R.N. and Rodriguez, R.S. (1988). On the degree of approximation of some singular integrals. Rend. Mat. Appl. (7) 8(3), 345–355.

Karsli, H. (2014). Fatou type convergence of nonlinear *m*-singular integral operators. Appl. Math. Comput. 246, 221-228.

Lèsniewicz, A., Rempulska L. and Wasiak, J. (1996). Approximation properties of the Picard singular integral in exponential weighted spaces. Publ. Mat. 40(2), 233–242.

Mamedov, R.G. (1963). On the order of convergence of *m*-singular integrals at generalized Lebesgue points and in the space $L_p(-\infty,\infty)$. (in Russian) Izv. Akad. Nauk SSSR Ser. Mat. 27, 287–304.

Mamedov, R.G. (1991). The Mellin Transform and Approximation Theory. (in Russian) "Elm", Baku.

Rempulska, L. and Walczak, Z. (2005). On modified Picard and Gauss-Weierstrass singular integrals. Ukrain. Mat. Zh. 57(11), 1577–1584.

Rempulska, L. and Tomczak, K. (2009). On some properties of the Picard operators. Arch. Math. (Brno) 45, 125–135.

Rudin, W. (1987). Real and Complex Analysis. (Third Ed.) McGraw-Hill Book Company, New York.

Rydzewska, B. (1978). Point-approximation des fonctions par des certaines intégrales singulières. (in French) Fasc. Math. 10, 13–24.

Stein, E.M. (1970). Singular Integrals and Differentiability Properties of Functions. (PMS-30) Princeton University Press, New Jersey.

Uysal, G. (2018). Nonlinear *m*-singular integral operators in the framework of Fatou type weighted convergence. Commun. Fac. Sci. Univ. Ank. Sér. A1 Math. Stat. 67(1), 262–276.

Uysal, G. (2019). Approximation by *m*-singular modified Picard operators. In: Proceeding Book of Applied Sciences of International Symposium on Recent Developments in Science, Technology and Social Sciences, 20-22 December. IKSAD International Publishing House, Ankara, 255-261.

Uysal, G. (2020). Some remarks on modified Picard operators preserving some exponential functions. Advances in Mathematics: Scientific Journal 9(9), 7679–7688.

Yilmaz, B. (2014). Jackson type generalization of nonlinear integral operators. J. Comput. Anal. Appl. 17(1), 77–83.

Yilmaz, B. (2019). On some approximation properties of the Gauss-Weierstrass operators. Commun. Fac. Sci. Univ. Ank. Ser. A1. Math. Stat. 68(2), 2154-2160.

Yilmaz, B. and Ari, D.A. (2020). A note on the modified Picard integral operators. Math. Meth. Appl. Sci. 1-6.

Yilmaz, B. and Uysal, G. (2020). On Kirov-type generalization of Gauss-Weierstrass operators. In: Academic Studies in Science and Mathematics Sciences. Livre De Lyon, Lyon, 59-69.

Yilmaz, B. (2020). Approximation properties of modified Gauss-Weierstrass integral operators in exponentially weighted spaces, accepted in: Facta Univ. Ser. Math. Inform. (to appear)

CHAPTER 5

NUCLEAR REACTORS

Mehtap Düz

(Assoc. Prof. Dr.), İnönü Üniversity, e-mail: mehtap.gunay@inonu.edu.tr
ORCID: 0000-0002-0445-1648

1. Introduction

With the increase in the world population, the acceleration of industrialization and technology increases the need for energy exponentially. Increasing income level, ease of city life and increased access to electricity, energy; It ceases to be a necessity and becomes an indispensable factor affecting the socioeconomic status of countries and even their prestige in the global world.

Energy resources are examined in two groups as renewable and non-renewable resources. Renewable energy sources are hydraulic energy, solar energy, hydrogen energy, energy obtained from biomass, geothermal energy, tidal energy and wave energy. Non-renewable energy sources are fossil fuels (oil, natural gas, coal) and nuclear energy sources (Koç and Şenel, 2013).

Fossil resources continued to be used increasingly until the 20th century and became the most preferred fuels. Reasons such as greenhouse gas emissions that cause climate changes together with fossil fuels, decrease in fossil fuels and their different distribution on the earth; it caused developed countries to decide to reduce fossil fuel use, but it did not fully materialize. The total amount of energy produced worldwide is 25.721 TWh. Fossil energy sources such as coal and gas, which increase CO_2 emissions, took part in the production of most of this energy. Renewable energy sources, on the other hand, cover only 12.2% of this cake, since they are energy sources that are higher in cost efficiency, can be negatively affected by natural conditions but are claimed to pollute the environment less.

Access to electricity and clean air is vital today. As the demand for electricity continues to increase, cleaner resources must be introduced to reduce greenhouse gas emissions in order to prevent climate change. This will result in a large increase in the use of all low-carbon energy sources, of which nuclear energy is an important part.

Thermal power plants where nuclear energy is produced under high security are called nuclear power plants. Its difference from other thermal power plants is that fission energy is used instead of chemical energy. Nuclear power plants are preferred as an important energy source due to their many positive features. One of the positive features of nuclear energy is that it has low CO_2 emission, is a cheap, sustainable and easily accessible source.

Uranium is the basis of nuclear energy. Uranium is mined in different geographies of the world. The amount of nuclear fuel used for electricity generation is much cheaper compared to other fossil fuels. The energy obtained from 2 grams of uranium, which is a fission energy source, is equivalent to the energy produced by 400 grams of coal, 350 m^3 of natural gas and 450 m^3 of oil from other fossil sources. However, since nuclear power plants do not release greenhouse gases during operation, they do not cause global warming.

Nuclear energy is obtained as a result of fusion, fission and half-life reactions. Half-life reaction, also known as natural fission; It is the breakdown of the nucleus and stabilizing.

Devices designed to convert the energy generated as a result of chain fission reaction into heat and energy in a controlled, safe and permanent manner are called nuclear reactors. The first nuclear reactor was an assembly produced in the USA by Enrico Fermi in 1942 and destroyed due to the high amount of radiation.

All nuclear reactors are configured according to almost the same operating principle. As we mentioned in the previous section, controlled energy production in a nuclear reactor takes place by keeping the chain fission reaction in balance. This can only be achieved by bringing the effective reproduction factor to the equilibrium state called critical. The working mechanism of the reactor consists of three parts:

- Primary cycle
- Secondary cycle
- Refrigeration cycle

In the primary cycle, nuclear energy converts nuclear fuel into heat energy. In the secondary cycle, heat energy is converted into kinetic energy in the turbine and into electrical energy in the generator. In the cooling cycle, the water vapor coming from the turbine is condensed and converted into water.

2. Basic Elements of Nuclear Reactors

Nuclear reactors are designed with great precision from the fuel pellet in the center to the concrete protection container at the outermost. The

insulation of the reactor against radiation leakage is provided at the highest level. Although there are different types of nuclear reactors, the basic components that make up a typical reactor are generally the same.

2.1. Fuels

Fuel materials, which form the heart of a reactor, are the most important part of the reactor's design and operation. Uranium is used as the main fuel in fission reactors. Uranium ore is passed through various stages from its natural state to being transformed into a fuel to be used in a nuclear reactor. These stages undergoing uranium are also called fuel cycle.

Two different methods are used in the fuel cycle: closed loop and open loop. While the fuel is cooled and sent to the temporary storage area in the open cycle, the U and Pu elements obtained from the spent fuel in the closed cycle are re-used as nuclear fuel. Since closed cycle is more expensive than open cycle in today's conditions, many countries prefer open fuel cycle (NRC, 2021; Sever, 2019).

If the fuel cycle is briefly summarized:

- Uranium ore is crushed in treatment plants and melted with the help of acid and carbonate. With the help of chemicals, it turns into a granular substance called U_3O_8 (yellow paste) containing 75% Uranium. Powdered U_3O_8 is transformed into UF_6 (hexafluoride) by subjecting to various chemical processes for enrichment.

- The enriched UF_6 brought to the recycling facility is converted into UO_2.

- UO_2 fuel pellets, which take the shape of a cylinder by pressing and heating in the fuel manufacturing facility, are armored and turned into a standard fuel form.

- Fuel bundles are sent to storage facilities after being used for 3 to 4 years to generate energy in nuclear power plants.

- If a closed fuel cycle is used, the fuels are recycled and made ready for use (World Nuclear Association, 2020; Öngü, 2014)

Natural uranium contains 99.282% ^{238}U, 0.712% ^{235}U and 0.006% ^{234}U isotopes. Although there are reactors using natural uranium as fuel, fission reactors generally use ^{235}U enriched at 3-5%. The enrichment process is applied to increase the fissile (fissionable) feature of the fuel.

Enriched uranium, which is formed into 1 inch (2.5 cm) long pellet (fuel rod) bundles, is placed in a pressurized boiler as fuel groups. In a

1000 MWe PWR grade, more than 18 million pellets and 51.000 fuel rods can be found.

In addition to these nuclear fuel types, we can also list the minor actinides that are being studied and are highly likely to be used in nuclear power reactors in the future. The biggest factor in the distance of some parts of the society to nuclear reactors is the nuclear waste generated as a result of the reaction and the storage of these wastes. These wastes can be seen in Table 1; uranium, plutonium, minor actinides (^{273}Np, ^{241}Am, ^{245}Cm) and fission products.

Table 1: Radioactive elements waste percentages and half-lives (Bardakçı, 2019).

Radioactive Elements	**Waste Percentage**	**Half-life (years)**
Uranium	**94.6%**	
^{236}U	0.8%	$7.04.10^{8}$
^{237}U	0.6%	$2.34.10^{7}$
^{238}U	98.6%	$4.47.10^{9}$
Neptunium	**0.06%**	
^{237}Np	100%	$2.14.10^{6}$
Plutonium	**1.1%**	
^{238}Pu	2.5%	87.7
^{239}Pu	54.2%	$2.41.10^{4}$
^{240}Pu	23.8%	$6.56.10^{3}$
^{241}Pu	12.6%	14.4
^{242}Pu	6.8%	$3.75.10^{5}$
Americium	**0.05%**	
^{241}Am	63.8%	432
^{243}Am	36%	$7.37.10^{3}$
Curium	**0.01%**	
^{243}Cm	1%	29.1
^{244}Cm	92.2%	18.1
^{245}Cm	5.7%	$8.5.10^{3}$
^{246}Cm	1.1%	$4.76.10^{3}$
Fission Products	**4.2%**	

Since minor actinides have a radiooxide effect and an undeniable power, if they are reused as fuel, the amount of nuclear waste will decrease as well as a large gain in nuclear energy production. For the transformation of minor actinides, light water reactors and fast reactors can be

considered. Many countries carry out theoretical and experimental studies on this.

In a fuel newly installed reactor, neutron sources such as a mixture of Po, Ra, Cf and Am-Be are needed to initiate the reaction. Alpha particles from the decay cause neutrons to be released from Be as they turn to ^{12}C. The core of the reactor covers all nuclear fuel and generates all heat. The core can contain hundreds of thousands of individual fuel pins (Yıldız & Köse, 2020).

Fuel rods are shielded to prevent the possibility of corrosion of the coolant and the spread of radiation. Fuel armor materials are expected to have low neutron absorption cross section, resistant to any corrosion, high resistance to corrosion with coolant and low cost.

2.2. Moderator

It is one of the main components of the reactor. Moderators are used to slow down the fast neutrons formed as a result of the fission reaction. Moderators;

• Being economical and easy to find,

• Low thermal neutron capture cross section,

• Small mass number,

• Chemical stability under radiation,

• Basic qualities such as high density are expected.

In natural uranium reactors, since graphite, beryllium and heavy water neutron capture cross sections are low; light water is used as a moderator in enriched uranium reactors due to its high neutron capture cross section.

2.3. Cooler

Cooler; It is a material that cools the reactor vessel heated by the fission reaction and transfers the heat from the fuel to the turbine. Examples of water (H_2O), heavy water (D_2O), Na, He, KCl, LiCl, NaOH, Sn, Hg, Pb, Bi, K, CO_2, H_2 and air coolants can be given (Baltacıoğlu, 1995). The most widely used of these is water. Basic qualities expected from reactor coolers; Low melting point, high boiling point, not causing corrosion on the structure, low neutron absorption cross section, radiation balance, heat balance, not interacting with turbine fluids, high heat transfer coefficient. The refrigerants used in nuclear reactors cannot meet all of these qualities. For this reason, nuclear reactors are classified according to the type of refrigerant they use. Pressurized water, boiling water, heavy water, supercritical water, organic (diphenyl, diphenyloxide), liquid metal (sodium, potassium, lead, bismuth and their mixtures), gaseous (air,

nitrogen, carbon dioxide, helium) and dissolved salt. reactor type has been developed.

2.4. Reflector

It is an important material for the continuation of the fission reaction in reactors. Its task is to reflect the neutrons leaking from the core and make them return to the chora. A good moderator for thermal reactors is also a good reflector. In order for the material used to be a good reflector, it must have low neutron absorption, high reflection coefficient, radiation stability (high radiation resistant) and resistant to oxidation. Materials such as water, heavy water, beryllium, graphite, steel, tungsten, carbide and zirconium silicide (Zr_3Si_2) are used as reflectors (Thermopedia, 2020; Nucleer Power, 2021; Nupex, 2021; General Atomics, 2021).

2.5. Control Sticks

It is made of neutron-absorbing materials such as B, Ag, In, Cd and Hf and is used to control the rate of reaction and stop it when necessary. Nuclear Operators increase the neutron density in the environment and the temperature depending on the density by lifting the control rods to initiate the nuclear reaction. If the system is to be kept in balance, the temperature is reduced by lowering the control rods. Besides the control rods, water pumps are also used to stabilize the temperature in the reactor (Elektrik port, 2021). Control rods are manufactured by companies such as General Electric (GE), Toshiba Westinghouse, Hitachi and MHI. A typical reactor has about 200 control rods and about 1000 fuel elements. Control rods are placed from the upper part of the reactor in the PWR, while in the BWR they are placed from the lower part of the reactor. This situation ensures that the fuel change in BWR is performed more easily than in PWR.

2.6. Reactor Pressure Vessel (RPV)

It is the structure that separates the reactor from the environment. These are usually in the form of a dome made of high density steel reinforced concrete. The outer casing of the nuclear power system is made of steel casing and concrete lining to prevent radioactive leakage. The concrete liner acts as a radiation shield, while the steel casing prevents leakage of liquid or gaseous materials (Elektrik port, 2021).

RPVs are steel containers made of alloy resistant to high temperature and radiation with a thickness ranging from 23 to 25 cm, including fuel rods and control rods.

This reinforced concrete and steel structure is designed to be highly resistant to natural disasters and any accident that may arise from outside (World Nüclear Association, 2020; Elektrik port, 2021).

Cooling towers are needed by some power plants to shed excess heat that cannot be converted into energy due to the laws of thermodynamics. These are symbols of nuclear energy. They just emit clean water vapor.

In order to prevent radioactivity from adversely affecting the environment, nuclear power plants are built within the framework of the "Safety at Depth" principle. This principle is within the reactor:

• Coating the fuel with ceramic,

• Placing the fuel in closed cylindrical fuel tubes resistant to radioactive leakage,

• The fuel tubes should be placed in a reactor vessel made of stainless steel designed in such a way that coolant passes between them,

• The primary protection container, also called steel jacket, covers the reactor,

• It includes all these systems to be located in a 1 meter thick domed concrete shelter building, also called secondary protection container.

All nuclear reactors are devices designed to sustain a chain fission reaction. They are grouped according to their purpose and design features (Yakut, 2018; Akyüzlü, 2008).

a. According to Usage Purposes;

• Research Reactors

• Conversion Reactors

• Power Reactors (Commercial reactors)

b. According to Fuel;

• Natural Uranium Fired Reactors (0.72% ^{235}U)

• Enriched Uranium Fired Reactors (0.72% ^{235}U), ^{239}Pu, ^{233}U)

• Plutonium Fired Reactors

• Thorium Fired Reactors

c. According to their coolers;

• Light Water Cooled Reactors

• Heavy Water Cooled Reactors

• Gas Cooled Reactors

d. According to Neutron Retardants;

• Light Water Reactors

• Heavy Water Reactors

• Graphite Reactors

e. According to Neutron Energies;

• Slow Energy Neutrons Operated Reactors

• Medium Energy Neutrons Powered Reactors

• Fast-Energy Neutron-Powered Reactors

Nuclear power reactors are generally classified according to refrigerant types. Light water reactors are the most preferred reactors. Pressurized water reactor (PWR) comes first among the reactors that are actively used for electricity generation and Boiling water reactor (BWR) comes second.

Research reactors are used in universities and research centers in many countries, including where power reactors are not operated. Research reactors are not used to generate electricity. These reactors produce neutrons for many purposes, including manufacturing radiopharmaceuticals for medical diagnosis and treatment, testing materials, and conducting basic research. Compared to power reactors, they have a simpler and smaller structure and operate with lower energy.

Conversion reactors are nuclear reactors that produce more than the amount of fissile material it consumes while generating energy. It is used in nuclear fuel production. Although at first these reactors were attractive to people because they were more economical than light water reactors, they lost all their attraction in the 1960s with the discovery of uranium reserves and the invention of new uranium enrichment methods. The production of ^{239}Pu from ^{238}U and production of ^{233}U from ^{232}Th can be given as examples of the conversions realized in conversion reactors. In both cases, one neutron is absorbed and two β particles decay.

Power reactors are nuclear power systems that are widely used today for commercial purposes. In addition to electricity generation, these reactors are used in drinking water production; It can also be used to generate energy for power ships and spacecraft. In addition, the heat released as a result of reaction in nuclear power reactors can meet the heat needs of many residential centers and industrial areas (Tokgöz, 2016).

3. Result

Many reasons such as population increase, increase in welfare level, rapid industrialization and technological developments have increased the need for energy in today's world. Human beings who started their energy adventure with fossil fuels; Realizing that these fuels have begun to run out and more importantly, the air necessary for their survival is

contaminated by the toxic gases emitted from these fuels, they have turned to alternative energy sources, which they consider cleaner.

The most preferred energy sources as alternative energy sources are wind and solar energy. These energy resources; Due to some negative features such as installation and maintenance cost, continuity, climate and exposure to natural disasters, it must be supported with a complementary energy source. This is possible with nuclear energy produced in nuclear power plant in accordance with the intended use of alternative energy sources.

Contrary to popular belief, nuclear energy provides clean, reliable and continuous energy production under high-level security measures. Today, energy is produced in 442 nuclear power plants in 33 countries.

Another situation that occupies the public opinion on the international platform is the increased spent fuels and their toxic effects with the developing nuclear technology. Spent fuels consist of uranium, plutonium, minor actinides and fission products. In many developed countries, uranium and plutonium, which make up the majority of waste fuel, are separated by chemical methods and used as fuel again. The fission products in the remaining waste fuel do not pose a problem in terms of their amount and degradation rate. Minor actinides cause concern due to their high nuclear power and toxic effects (Düz, 2020; Özgener, 2009). Since minor actinides have radiooxide effect and undeniable power, if they are reused as fuel, they will contribute greatly to both the reduction of nuclear waste and nuclear energy production. For the transformation of minor actinides, light water reactors and fast reactors can be considered. Many countries carry out theoretical and experimental studies on this.

References

Akyüzlü, Ö. F., (2008), "Nükleer Reaktör Yakıt İmalatı Öncesi Uranyumun Saflaştırılması ve Zenginleştirme Prosesleri", *Marmara Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

Baltacıoğlu, N., (1995), Fisyon Reaktörlerinde Kullanılmış Yakıt Çubuklarının Füzyon-Fisyon Hibrid Reaktörlerinde Yeniden Kullanılabilir Hale Dönüştürülmesi, *Erciyes Üniversitesi Fen Bilimleri Enstitüsü*, Doktora Tezi.

Bardakçı, H., (2019), "Bir Füzyon-Fisyon Hibrit Reaktöründe Monte Carlo Tekniği Kullanılarak Bazı Kütüphaneler İçin Üç Boyutlu Nötronik Hesaplamalar", *İnönü Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

Düz, M., (2020), "Bir Kaynar Su Reaktör Modellemesinde Np ve Am Yakıt Çubukları İçin Bazı Nötronik Değerlerin İncelenmesi", *Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi Fen Dergisi*, Vol. 15, pp. 140-147.

Elektrik Port, (2021), https://www.elektrikport.com.

General Atomics, (2021), "Sıga™ Sıc Composite", https://www.ga.com/nuclear-fission/siga-sic-composite.

Koç, E. and Şenel, M. C., (2013), "Dünyada ve Türkiye'de Enerji Durumu- Genel Değerlendirme", *Mühendis ve Makina*, Vol. 54, pp. 32-44.

Nrc, (2021), "Stages of the Nuclear Fuel Cycle", https://www.nrc.gov/materials/fuel-cycle-fac/stages-fuel-cycle.htm.

Nucleer Power, (2021), "Neutron Nuclear Reactions", https://www.nuclear-power.net/nuclear-power/reactor-physics/nuclear-engineering-fundamentals/neutron-nuclear-reactions.

Nupex, (2021), "Nuclear Reactors", http://nupex.eu/index.php?g=textcontent/nuclearenergy.

Öngü, S., (2014), "Nükleer Reaktörler, Yakıt Tipleri ve Mersin Akkuyu Nükleer Santrali", *Niğde Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

Özgener, H., (2009), "Nükleer Atık Sorunu ve Uranyum Ötesi Elementlerin Dönüşümleri", *X. Ulusal Nükleer Bilimler ve Teknolojileri Kongresi*, (pp. 27-46), İstanbul Teknik Üniversitesi, Enerji Enstitüsü, 6-9 Ekim.

Sever, O., (2019), "Çevre ve Stratejik Bakış Açısıyla Türkiye'de Nükleer Santral Çalışmaları*", Aksaray Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

Thermopedia, (2020), http://www.thermopedia.com.

Tokgöz, S. R., (2016), "Araştırma ve Güç Reaktörlerinde Kontrol ve Yakıt Malzemelerinin İncelenmesi", *Sakarya Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

World Nuclear Association, (2020), https://www.world-nuclear.org.

Yakut, F., (2018), "Nükleer Reaktörlerde Kullanılan Kontrol Çubuklarının Malzeme Açısından İncelenmesi", *Osmaniye Korkut Ata Üniversitesi Fen Bilimleri Enstitüsü*, Yüksek Lisans Tezi.

Yıldız, A. and Köse, E., (2020), Nükleer Santrallerde İş Kazaları, Ejons İnternational Journal On Mathematic, *Engineering And Natural Sciences*, Vol. 13, pp. 93-111.

CHAPTER 6

IDENTIFYING AND RANKING IMPORTANT PERSONALITY TRAITS OF PH. D. DEGREE STUDENTS: APPLICATION OF LINEAR CHEBYSHEV BEST WORST METHOD

Nazlı Ersoy
(RA. Dr.), Kilis 7 Aralık University, e-mail: ersoynazli3@gmail.com
ORCID: 0000-0003-0011-2216

1. Introduction

Education is important factor in building the individual and social identity of people and shaping their lives. A good educational background increases the number of qualified people, qualified people increase productivity, and this increases the level of development. Countries with well-educated manpower have a more effective position than those with inadequately educated and populated countries.

Ph.D. education is the center of academic practice (Pyhältö et al., 2012: 1) and one of the most critical points of the academic profession, which requires personal development as well as the course load. In this context, in Ph.D. degree programs, candidates are expected to develop scientific innovations, to develop a method, or to use the existing scientific methods in the relevant field.

In such an important process, student selection process needs to be addressed seriously. Admission requirements for doctoral programs are not very easy but vary from university to university. However, it is known that the candidates go through stages that can be described as difficult in general. In the process of selection of students for Ph.D. degree programs, candidates are required to have a degree certificate, proficiency in a second language that may differ from one country to another, expected to meet certain requirements such as high scores from certain assessment tests and scientific tests given by universities aimed to measure candidates' competence, and in cases of interviews, candidates' personality traits can have effects on the process.

In the presence of multiple criteria and alternatives, MCDM methods are suitable for decision-making process. To see a problem as an MCDM

problem, there must be at least two alternatives and more than one criteria. In such a case, the MCDM provides the decision-maker with an appropriate solution framework to reach the final decision.

This study was aimed to identify and rank important personality traits of Ph.D. degree students through the eyes of academics. In this direction, as a result of a comprehensive literature review, eight personality traits were included in the analysis. Seven decision-makers from different universities were determined and analysis was performed with the new MCDM method which is called Linear Chbyshev Best Worst Method (BWM) method. The basic contribution of this study is two-way;

i) The subject of this study was examined for the first time with the Linear Chbyshev BWM method
ii) It is thought that this study will constitute a reference model for researchers and academicians in the future.

The rest of the paper is organised as follows: The literature review and notional framework are included in section 2. The mathematical model used in the application is described in Section 3. The findings of the study are included in Section 4. The concluding remarks, discussion and suggestions are involved in Section 5.

2. Conceptual Framework and Literature Review

The studies in the literature are mostly focused on the content analysis of doctoral theses (see Horton & Hawkings, 2010), identification of problems encountered during the dissertation process and coping strategies (see Devonport & Lane, 2006), the contribution of Ph.D. students (see Larivière, 2012), doctoral student socialization (see Russell et al. 2016), the knowledge and abilities of doctoral students (see Melin & Janson, 2006), dispute between doctoral students and supervisors (see Gunnarsson et al. 2013), educational environment and working conditions of doctoral students (see Kolmos et al. 2008), mental health problems and work for organization in Ph.D. students (see Levecque et al. 2017). The qualitative method was mostly preferred in the mentioned studies. The following examples can be given to the studies that deal with similar research topics with MCDM methods.

The following examples can be given to the studies conducted in the education sector using MCDM methods. Salimi and Rezaei (2016) measured the efficiency of projects in the universities using the BWM. They determined the weights using the BWM. Kabak and Dağdeviren (2014) assesssed the sustainability of university selection choices of the

students. The ANP was used to identfy the criteria weights and the PROMETHEE method was used to determine performance ranking. Genç and Masca (2018) evaluated the preferences of the university students using fuzzy VIKOR method. Aghdaie and Behzadian (2010) applied the TOPSIS and AHP approaches to select the thesis subject. Musani and Jemain (2013) applied fuzzy VIKOR to evaluate the school performance. Pekkaya (2015) used TOPSIS, VIKOR and PROMETHEE to determine the Career Preference of University Students. Fadlina et al. (2017) used Extended Promethee II method for determining the best college student. Deliktas and Ustun (2017) applied fuzzy MULTIMOORA and goal programming for Erasmus student selection in Dumlupinar University. Baki et al. (2017) applied AHP to select the best student for the academic semester. Aires et al. (2017) developed hybrid ELECTRE–TOPSIS method for selecting the best students in Brazilian university. Wu et al. (2012) used AHP and VIKOR methods for Ranking 12 private universities based on performance evaluation. Tuan et al. (2020) applied fuzzy AHP and TOPSIS method for evaluating and ranking the lecturers' productivity.

As can be seen from the examples above, many different studies have been made in the education sector using MCDM methods. BWM, a new method, has been handled by researchers in different subjects and in different fields. The following examples can be given to the studies using BWM:

Rezaei (2015a) handled the mobile phone selection problem using the BWM. In this direction, four different phone models have been chosen under various criteria and made the selection with BWM. Also, comparisons between the two methods have been made using the AHP method and it has been shown that BWM performs better. Rezaei et al. (2015b) proposed a supplier development model for developing the suppliers and they used BWM to segment them. Firstly, suppliers were evaluated considering their capacity and willingness to cooperate. The BWM model was then applied to determine the criteria weights. Rezaei et al. (2016) benefited from the BWM approach in supplier selection. In the study, which has a three-stage selection process including pre-selection, selection and aggregation, seven suppliers were evaluated considering 8 selection criteria. Ahmadi et al. (2017) evaluated the sustainability performance of the manufacturing companies. In the study where 38 experts evaluated the social sustainability criteria, the most important criteria was determined as contractual stakeholders' influence. Gupta (2018) used the hybrid BWM and VIKOR methods to measure the service quality of the airlines. BWM was used for obtaining the criteria weights while VIKOR methodology was used to rank the airlines. Salimi and Rezaei (2016) evaluated the effectiveness of Ph.D. projects using the

BWM. Guo and Zhao (2017) performed three case studies using the fuzzy BWM. Van de Kaa et al. (2017) used BWM to select the thermochemical conversion of biomass technology in the Netherlands. Rezaei et al. (2018) benefited from the BWM to assess logistic performance. Nawaz et al. (2018) proposed the BWM and Markov Chain approach for selection the cloud service. The Markov Chain approach was used to determine the pattern, BWM was used to rank the services. Shojaei et al. (2018) measured the airports performance using the integrated Taguchi Loss Function, VIKOR and BWM techniques. Wan Ahmad et al. (2017) measured the outside factors affecting the sustainability of the gas and oil supply chain with the BWM. Kocak et al. (2018) suggested the Euclidean BWM method to analyze the car selection and transportation mode selection problems. At the end of the study, the proposed model was found to be a more effective method for the solution than linear and nonlinear Chebyshev BWM models.

3. Method

3.1. Research Design

In this study, multi-stage design was used and the importance levels of the features that should be found in doctoral students were listed by BWM method in line with the opinions of decision makers. In the study, which included seven decision makers and eight criteria, the criteria were specified after the literature review and the criterion weights were determined using the evaluations of the decision makers working at Muğla Sıtkı Koçman and Akdeniz Universities.

3.2. Research sample

The research was carried out with academicians working at Akdeniz and Muğla Sıtkı Koçman universities. Three research assistants and five lecturers conducted the evaluations to designate the criteria weights. The working group consists of four female and four male volunteer academicians. The more detailed information is presented in Table 1.

Table 1**.** Information for Decision Makers

University	Muğla Sıtkı Koçman University	4
	Akdeniz University	4
Position	Associate professor	5
	Research assistant	3
Gender	Female	4
	Male	4

The criteria were determined by comprehensive literature review and presented in Table 2.

Table 2. Decision Criteria

Criteria	**Symbol**
Problem solving ability	C_1
Enthusiasm	C_2
Critical thinking	C_3
Reliability	C_4
Be open to new ideas	C_5
Cheeriness	C_6
Being a people person	C_7
Researcher personality	C_8

3.3. Research Instrument and Procedures

In the last decade, MCDM techniques have been used in many fields such as business and energy management, health and safety management. All these techniques require binary comparisons between different criteria. But, according to Rezai (2015a) most of these techniques have consistency issues in the process of pairwise comparison. To come through this problem, Rezai (2015a)

developed a new MCDM technique called BWM. This method solves the problem of inconsistency during pairwise comparison using a new procedure in the comparison of alternatives and allows fewer comparisons compared to other MCDM techniques (Gupta and Barua, 2016: 5-6).

Table 3. Scale for Best Worst Methodology

Equally significant	Equal to moderately more significant	Partially more significant	Partially to powerfully more significant	Powerfully more significant	Powerfully to verypowerfully more significant	Very powerfully more significant	Very powerfully to extremely more significant	Extremely more significant
1	2	3	4	5	6	7	8	9

Source: Gupta, 2018: 40.

In the BWM, the scale 1/9 in Table 3 was used to make a pairwise comparison of criteria (n criteria). The matrix obtained in this direction is as follows;

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

where a_{ij} indicates the relative criterion of i according to j.
$a_{ij}=1$ indicates that i and j are of the same significance,
$a_{ij}>1$ indicates that i is more significant than j,
$a_{ij}= 9$ indicates the extreme significance of i to j.

a_{ji} indicates the significance of j to i. For matrix A to be reciprocal, it is required that $a_{ij}=1/a_{ji}$ and $a_{ii}=1$for all i and j. When executing a pairwise comparison, a_{ij} refers to both the aspect and force of the decision maker i relative to j.

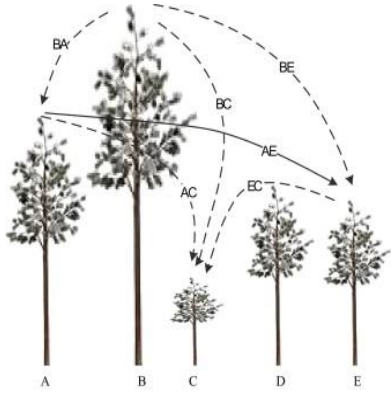


Figure 1. The Preference of A over E.

Source: Rezaei, 2015a: 50.

According to Figure 1, comparing A to other trees, it is easy to say that A is taller than other trees and shorter than B but it is difficult to designate a number to state the relative length level. When a person wishes to give a number to show his judgment for comparing A and B, he takes into account the connection between the two and others. Suppose you want to designate a number to specify the choice of tree A on tree E. Then, it is easy to say that A is longer than E, and A to E will be assigned a number bigger than 1 to indicate A's preference. This number cannot be 9 because

C is quite shorter than E and B is longer than A. In this case, a number such as four or five should be assigned. Anyone who wants to compare B and C assign the number 8. If A is preferred to E, the best and worst alternatives are also considered by the decision-maker. In the example above, the best (highest) tree is B, while the worst tree is C. D has no role in this example. If more trees that are longer than C and shorter than B are added, D still has no role in this comparison. The same problem emerges for A and E when comparing A and D (Rezaei, 2015a: 50-51).

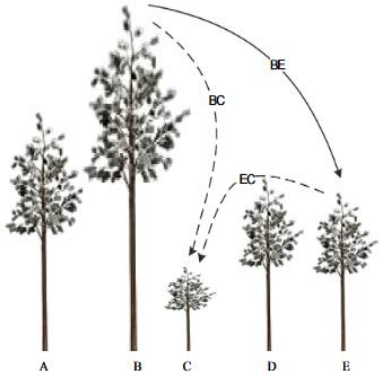


Figure 2. The Preference of B over E.

Source: Rezaei, 2015a: 50.

On the other hand, it is easier to compare when a member of a couple is the best or worst one (e.g, B to E). In this case, following the same logic line, only three comparisons (Figure 2) are required with regard to the six comparisons in the former example (Figure 1). In this case, evaluation is easier (Rezaei, 2015a: 50-51).

The steps of the BWM whose general logic is given above are as follows (Rezaei, 2015a: 51-52):

Step 1. Determining the decision criteria.

A set of criteria $\{c_1, c_2, ..., c_n\}$ is obtained to make a decision. For instance, in the case of purchasing a car, quality, style, price, safety can be the decision criteria.

Step 2. Determining the best and worst decision criteria**.**

The best and worst criteria are defined by decision-makers.

Step 3. Determining the best criterion over all other criteria using a number between 1-9.

A score of 1 indicates the equal choice between the best and another criteria. A score of 9 indicates the over preference of the best criterion over the other criteria. Accordingly, the BO (Best-to-Others) vector is obtained as follows:

$$A_B = (a_{B1}, a_{B2}, ..., a_{Bn})$$

a_{Bj} shows the choice of the best criterion B over criterion j, and $a_{BB} = 1$.

Step 4. Determining a number between 1-9 to choose all criteria over the worst criterion.

The OW (Others-to-Worst) vector is obtained as follows:

$$A_W = (a_{1W}, a_{2W}, ..., a_{nW})^T$$

where a_{jW} shows the choice of the criterion j instead of worst criterion W, $a_{WW} = 1$.

Step 5. Find the optimum weights ($w_1^*, w_2^*, w_3^*, \ldots, w_n^*$) .

The optimum weights are calculated in this step. The optimum criteria weights will perform the following necessaries: The ideal case for each pair of w_j/w_W and w_B/w_j is $w_j/w_W = a_{jW}$ where and $w_B/w_j = a_{Bj}$. For this reason, to reach the ideal solution, we must minimise the maximum value between sequences $\left|\frac{w_B}{w_J} - a_{Bj}\right|$ and $\left|\frac{w_j}{w_w} - a_{jw}\right|$, and it can be expressed as:

$$\min \max_j \left\{ \left| \frac{w_B}{w_j} - a_{Bj} \right|, \left| \frac{w_j}{w_w} - a_{jw} \right| \right\}$$

Subject to

$$\sum_j w_j = 1$$

$$w_j \geq 0, \text{ for all j.} \qquad \textbf{(1)}$$

Problem (1) is transformed into a linear model as follows:

$$\min \xi$$

s.t.

$$\left|\frac{w_B}{w_j} - a_{Bj}\right| \leq \xi, \text{ for all j}$$

$$\left|\frac{w_j}{w_W} - a_{jW}\right| \leq \xi, \text{ for all j.}$$

$$\sum_j w_j = 1$$

$$w_j \geq 0, \text{ for all j.} \qquad \textbf{(2)}$$

Model **(2)** is solved to gain optimized weights (w_1*, w_2*, w_3*, … ,w_n*) and optimum value ξ^*. As the ξ^* value approaches zero, the consistency increases and comparisons become more reliable.

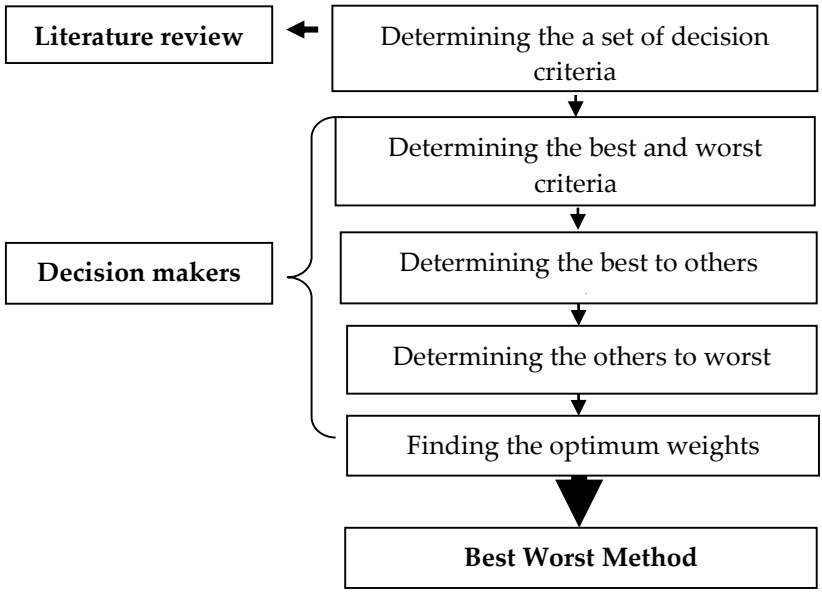


Figure 3**.** Application Steps of BWM

3.4. Data Analysis

To analyze the data, decision-makers were enquired to obtain the best and worst criteria using Table 2. BWM steps were applied in order to list the personality traits that are considered important in doctoral students. Analyzes were made using Microsoft Excel 2016 office program.

4. Results

In this study, a MCDM model is proposed to rank the personality traits seen in doctoral students according to their importance. For this purpose, the BWM steps given in Figure 3 were followed sequentially.

4.1. Identifying the Best and the Worst Criteria

The first step of the BWM model is the establishment of decision criteria. Accordingly, decision-makers were enquired to designate the best and worst criteria using Table 3. They were determined through a questionnaire and presented in Table 4.

Table 4. The Best and Worst Criteria Described by Respondents

Criteria	Best	Worst
C_1		
C_2	2/3/4	1
C_3	6	
C_4		
C_5		
C_6		2/3/4/5/6/7
C_7	1	
C_8	5/7	

4.2. Determining the Best Criterion Preference Among All Criteria

At this stage, participants were enquired to designate the preference of the best criterion using the 1-9 scale (Table 3). The results obtained according to the preferences of each participant are presented in Table 5.

Table 5. BO Vectors for Seven Respondents

Respondents	Best Criterion	Criteria							
		C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8
1	C_7	6	9	4	3	5	2	1	6
2	C_2	2	1	3	3	4	9	8	4
3	C_2	2	1	3	8	3	9	7	2
4	C_2	6	1	3	4	5	8	7	2
5	C_8	3	2	4	5	4	9	3	1
6	C_3	3	4	1	6	5	9	7	2
7	C_8	2	5	3	6	4	8	7	1

4.3. Identifying the Other Criteria Preference Over the Worst Criterion

Participants were asked to specify the choice rate of all criteria over the least significant criterion through a questionnaire. The results are presented in Table 6.

Table 6. OW Vectors for Seven Respondents

Respondents	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Worst criterion:	C_2	C_6	C_6	C_6	C_6	C_6	C_6
C1	3	9	8	3	6	7	7
C2	1	8	9	8	7	6	4
C3	4	7	7	6	5	9	6
C4	5	7	2	4	4	3	5
C5	6	6	6	5	5	5	3
C6	7	1	1	1	1	1	1
C7	9	8	3	2	6	2	2
C8	3	6	8	7	8	8	8

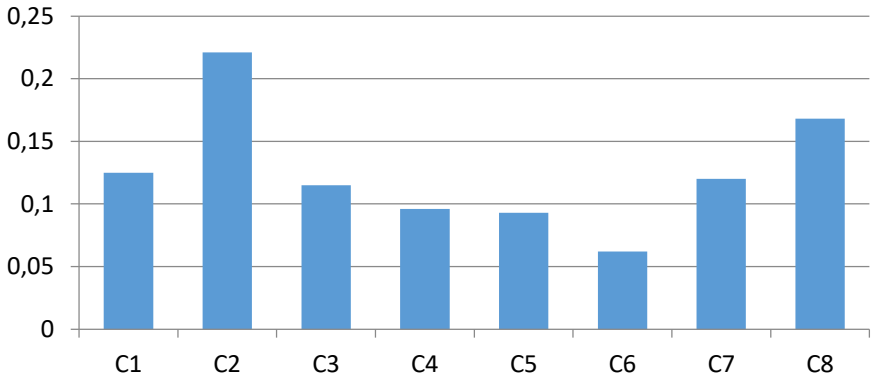
4.4. Finding the Optimum Criteria Weights

In the light of the results obtained from the opinions of seven participants, the optimum criteria weights were calculated using the Linear Chbyshev BWM. A simple weighted average was calculated for each criterion to obtain a one weight vector, and results are presented in Table 7.

Table 7. The Criteria Weights Determined by the Seven Respondents

Criteria	**Weight**
Problem solving ability **(C1)**	0,138
Enthusiasm **(C2)**	0,183
Critical thinking **(C3)**	0,149
Reliability **(C4)**	0,088
Be open to new ideas **(C5)**	0,093
Cheeriness **(C6)**	0,052
Being a people person **(C7)**	0,102
Researcher personality **(C8)**	0,195
ξ^{L*}	0,0781054

According to Table 7, average consistency rate (ξ^{L*}) is close to zero. The fact that the consistency indicator is about to zero indicates that the comparisons made have a high consistency.



Graph 1. Criteria Weights

According to the results, it was found that C_8 had the highest importance level among the personality traits of doctoral students. Also, it was concluded that C_8 followed by C_2,C_3,C_1criteria. On the other hand, the C6 criterion was determined to be the lowest criterion. The C_6 were followed by C_4, $C_5$and C_7, respectively.

5. Discussion, Conclusion and Recommendations

Education, which is one of the most important parameters of the modern world and whose importance is increasing day by day, is the most important way to connect to life. Education, which serves many purposes such as equality and justice, independence and freedom, self-reliance, national development, has an important place in our self-discovery and understanding of the world. In this context, with the doctorate process, which is one of the most important steps of education, people gain knowledge, skills and experience in various fields. The students who take part in such an important process that contributes to science have quite different characteristics from each other. This may be a point affecting production processes and thus the contribution to society.

In this direction, the personality traits of doctoral students were determined and the importance of these traits was listed with the help of Linear Chbyshev BWM. Accordingly, eight criteria and seven decision-makers were included in the analysis. At the end of the study, the most important criterion was determined as C_8 (researcher personality), while the criterion with the lowest significance level was determined as C_6 (cheeriness). Besides, it was determined that the obtained consistency ratio was close to 0.This shows that the comparisons between the criteria have a high consistency.

In future studies, the AHP method based on paired comparison of criteria can be used to designate the significance of the criteria and the results can be compared. On the other hand, methods such as Euclidean BWM, Nonlinear Chebyshev BWM andFuzzy BWM can be used by gaining different perspectives to the study.

References

Aghdaie, M. H.,& Behzadian, M., 2010, A hybrid Fuzzy MCDM Aproach to Thesis Subject Selection. *J. Mat. Comp. Sci*, *1*(4), 355-365.

Ahmadi, H.B., Kusi-Sarpong, S., and Rezaei, J., 2017, Assessing the social sustainability of supply chains using best worst method. *Resources, conservation and recycling*, 126, 99–106.

Aires, R. F., Ferreira, L., de Araujo, A. G., & Borenstein, D., 2017, Student selection in a Brazilian university: using a multi-criteria method. *Journal of the operational research society*, 1-14.

Baki, N. A., Norddin, N. I., & Azaman, W. A. W., 2017, Application of Analytic Hierarchy Process for Selecting Best Student. *Journal of Applied Environmental and Biological Sciences*, *7*(1), 69-73.

Deliktas, D. and Ustun, O., 2017, Student selection and assignment methodology based on fuzzy MULTIMOORA and multichoice goal programming. *International Transactions in Operational Research*, *24*(5), 1173-1195.

Devonport, T. J.,& Lane, A. M., 2006, Cognitive appraisal of dissertation stress among undergraduate students. *The psychological record*, *56*(2), 259-266.

Fadlina, Sianturi, L. T., Karim, A. Mesran, Siahaan, A.P.T., 2017, Best Student Selection Using Extended Promethee II Method,International Journal of Recent Trends in Engineering & Research (IJRTER),03:08, 21-29.

Genç, T.,& Masca, M., 2018, An application of fuzzy VIKOR method on evaluation of the university students' preferences on selected Turkish banks. *International Journal of Business and Systems Research*, *12*(1), 13-24.

Gunnarsson, R., Jonasson, G., and Billhult, A., 2013, The experience of disagreement between students and supervisors in PhD education: a qualitative study. *BMC Medical Education*, *13*(1), 1-8.

Guo, S., and Zhao, H., 2017, Fuzzy best-worst multi-criteria decision-making method and its applications. *Knowledge-Based Systems*, 121, 23–31.

Gupta, H., 2018, Evaluating service quality of airline ındustry using hybrid best worst method and vikor. *Journal of Air Transport Management*, 68, 35–47.

Gupta, H. and Barua, M. K., 2016, Identifying Enablers of technological ınnovation for ındian msmes using best–worst multi criteria decision making method. *Technological Forecasting and Social Change*, 107, 69–79.

Horton, E. G.,& Hawkins, M., 2010, A content analysis of intervention research in social work doctoral dissertations. *Journal of Evidence-Based Social Work*, *7*(5), 377-386.

Kabak, M.,& Dağdeviren, M., 2014, A hybrid MCDM approach to assess the sustainability of students' preferences for university selection. *Technological and Economic Development of Economy*, *20*(3), 391-418.

Kocak, H., Caglar, A., and Oztas, G. Z., 2018, Euclidean best–worst method and ıts application. *International Journal of Information Technology & Decision Making*, 17(05), 1587-1605.

Kolmos, A., Kofoed, L. B., and Du, X. Y., 2008, PhD students' work conditions and study environment in university-and industry-based PhD programmes. *European Journal of Engineering Education*, *33*(5-6), 539-550.

Larivière, V., 2012, On the shoulders of students? The contribution of PhD students to the advancement of knowledge. *Scientometrics*, *90*(2), 463-481.

Levecque, K., Anseel, F., De Beuckelaer, A., Van der Heyden, J., and Gisle, L., 2017, Work organization and mental health problems in PHD students. *Research Policy*, *46*(4), 868-879.

Melin, G. and Janson, K., 2006, What skills and knowledge should aphd have? Changing preconditions for phd education and post doc work. *Wenner Gren International Series*, *83*, 105-118.

Musani, S., and Jemain, A. A., 2013, November, A fuzzy MCDM approach for evaluating school performance based on linguistic information. In *AIP Conference Proceedings* (Vol. 1571, No. 1, pp. 1006-1012).

Nawaz, F., Asadabadi, M. R., Janjua, N. K., Hussain, O. K., Chang, E., and Saberi, M., 2018, An mcdm method for cloud service

selection using a markov chain and the best-worst method. *Knowledge-Based Systems,* 159, 120-131.

Pekkaya, M., 2015, Career preference of university students: An application of MCDM methods. *Procedia Economics and Finance*, *23*, 249-255.

Pyhältö, K., Toom, A., Stubb, J., and Lonka, K., 2012, Challenges of becoming a scholar: a study of doctoral students' problems and well-being. *ISRN Education*, 2012, 1–12.

Rezaei, J., 2015a, Best-worst multi-criteria decision-making method. *Omega*, 53, 49–57.

Rezaei, J., Nispeling, T., Sarkis, J., and Tavasszy, L., 2016, A Supplier selection life cycle approach ıntegrating traditional and environmental criteria using the best worst method. *Journal of Cleaner Production*, 135, 577–588.

Rezaei, J., van Roekel, W. S., and Tavasszy, L., 2018, Measuring the relative importance of the logistics performance ındex indicators using best worst method. *Transport Policy*, 68, 158–169.

Rezaei, J., Wang, J., and Tavasszy, L., 2015b, Linking supplier development to supplier segmentation using best worst method. *Expert Systems with Applications*, 42(23), 9152–9164.

Russell, J., Gaudreault, K. L., & Richards, K. A. R., 2016, Doctoral student socialization: Educating stewards of the physical education profession. *Quest, 68*(4), 439-456.

Salimi, N. and Rezaei, J., 2016, Measuring efficiency of university-ındustry phd projects using best worst method. *Scientometrics*, 109(3), 1911-1938.

Shojaei, P., Seyed Haeri, S. A., and Mohammadi, S., 2018, Airports evaluation and ranking model using taguchi loss function, best-worst method and vıkor technique. *Journal of Air Transport Management*, 68, 4–13.

Tuan, N., Hue, T., Lien, L., Thao, T., Quyet, N., Van, L., & Anh, L., 2020, A new integrated MCDM approach for lecturers' research productivity evaluation. *Decision Science Letters*, *9*(3), 355-364.

Van de Kaa, G., Kamp, L., and Rezaei, J., 2017, Selection of biomass thermochemical conversion technology in the netherlands: a best worst method approach. *Journal of Cleaner Production*, 166, 32–39.

Wan Ahmad, W. N. K., Rezaei, J., Sadaghiani, S., and Tavasszy, L. A., 2017, Evaluation of the external forces affecting the sustainability

of oil and gas supply chain using best worst method. *Journal of Cleaner Production*, 153, 242–252.

Wu, H. Y., Chen, J. K., Chen, I. S., & Zhuo, H. H., 2012, Ranking universities based on performance evaluation by a hybrid MCDM model. *Measurement*, *45*(5), 856-880.

CHAPTER 7

THE EVOLUTION AND GENETICS OF TURKISH EMMER WHEAT

Dogan Ilhan

(Asst. Prof. Dr.), Kafkas University, e-mail:doganilhan@kafkas.edu.tr
ORCID: 0000-0003-2805-1638

1. Introduction

The Anatolian geography, which has been the cradle of many civilizations throughout the historical process, has played an important role in the cultivation of the plant species used in the nutrition of human and other mammals and in their distribution throughout the world. Turkey, due to its location at the crossing point for plant species spread throughout the world is a very important gene center. In addition, it is the gene center of other important field crops as well as wheat, since it is located at the intersection of two different gene centers (Mediterranean and Fertile Crescent Gene Centers) (Baloch et al. 2017). Wild relatives are important in expanding the genetic pool of agricultural products. Therefore, Turkey is a very important origin that hosts 23 wild relatives (*Triticum*, *Aegilops*, *Amblyopyrum*, *Dasypyrum* etc.) in the primary and secondary gene pool of wheat (Feldman and Levy, 2015). Wheat has gone through some important stages along with the historical process. Wheat and barley were in the wild form that shed their grains and germinated irregularly when they first appeared. There were two in the wild emmer ear. If one of these grains germinated in fall, the other could germinate in spring the following year. Because of this, germination was a big problem in the cultivation of wheat. In the following years, wheats that underwent natural mutations turned into regular germinating forms that do not shed their grains. The best example that can be given to this situation is emmer wheat. Wild emmer (*Triticum dicoccoides*) evolved by mutation or natural selection into a primitive form called emmer (*Triticum dicoccum*) (Zohary, 1969). In the Anatolian lands, grains were first planted by human beings and they started to be fed with the products obtained. Wheat, which is considered to be one of the first products of the settled order, was first consumed by boiling and turning it into a mushy consistency. Archaeobotanists, who examined the wheat grains they found in the soil remains, came across einkorn and emmer wheat. These wheats, which belong to the same family, have shaped themselves according to the climatic conditions. It is named as einkorn wheat in milder regions, and its relative that grows in cold climate is called

emmer wheat (Haldersen et al. 2011). In ancient times, primitive species were directly collected from nature and consumed as food during drought and famine years. Primitive wheat varieties such as *Triticum monococcum* L. (Einkorn) and *Triticum dicoccum* S. (Emmer) were used for this purpose for many years and then were cultivated by the producers (Zaharieva and Monneveux, 2014). Information such as the genetic diversity, gene content and genome structure of the bread wheat used today can be reached in detail thanks to the genome information of wild emmer wheat (*T. turgidum* ssp. *dicoccoides*) (Haudry et al. 2007). These ancestral species in the wheat taxon also contain many important features. Emmer wheat (*T. dicoccum*), one of the most important of these, is a type of wheat that has potential for the sustainable development of agriculture. It is stated that it has very favorable advantages for farmers such as resistance to wheat diseases and resistance to drought. Studies show that emmer wheat fields are very resistant to pathogen attack and rust disease. It is also a potential product that should be used for the production of many new products due to its high protein content (Konvalina et al. 2012b). Again, emmer wheat is a highly competitive variety against weeds and many stress factors (fertilization with low levels of nitrogen). In studies comparing emmer wheat and other wheat; emmer wheat is seen to be superior in terms of adaptation to soil, resistance to climatic environmental conditions, grain quality and crude protein content (Konvalina et al. 2012a). When the adaptability of this wheat type was investigated, it was found that special and general adaptation abilities for all quantitative characters were extremely high (Mahantashivayogayya et al. 2004). Considering all these features, it is preferred in wheat breeding not only because of its taste and its usefulness in terms of content, but also because of its resistance to many negative factors (Zaharieva et al. 2010).

Based on the studies conducted in recent years, the cultivation areas of local wheat varieties are gradually narrowing. Although there is no clear statistical information about how much of the local wheat sown covers, there are studies to determine the most grown local wheat. Among the local wheat grown in Turkey, the ones with the most cultivation area can be listed as follows; Zerun wheat, White wheat, Red wheat, Yellow wheat, Charcoal wheat, Kirik wheat, Einkorn, Koca Wheat, Toptas wheat (Kan et al. 2016). Emmer, a close relative of einkorn, is an ancient wheat that has been growing in Anatolia for centuries and is endangered. It is also called Kavilca, Kabluca, wild wheat, especially in Kars province in Turkey (Ertugay et al. 2020). Being one of the ancestors of wheat, this variety has many features different from modern wheat types. It is high in fiber, high protein but low gluten. One of the most important features of local wheat varieties is that the amount of gluten differs from modern varieties (Heun et al. 1997). Due to their natural population, local varieties are more resistant to environmental conditions, diseases and pests than modern

wheat. In Turkey, on the slopes, rocky areas; It is grown in lands where other crops / modern wheat varieties cannot be grown (van Ginkel and Ogbonnaya, 2007). Wild and local forms of plant genetic resources that can develop by adapting to different geographical conditions are very valuable for the continuity of genetic heritage. Therefore, these genetic resources should be used with correct methods (Malik and Singh, 2006).

Due to the general structure of the province of Kars, agricultural lands are at high altitudes. In addition, seasonal and day-night temperature differences are high. The dry farming system is dominant in the province, as there is little variety in agriculture due to the short plant growing period and the unsuitability of the ecological structure. Crop production is mostly concentrated on cereals. In Kars, only barley and wheat are produced as cereals, the others are almost nonexistent (Karagöz, 1996). Emmer wheat, which is a local genetic resource specific to Kars, is considered within the emmer group wheats. In order to adapt to cold climate, emmer increased the number of husks surrounding the seed and thickened the forks in the ear (Feldman and Kislev, 2007). Emmer (*T. dicoccum*) and einkorn (*T. monococcum*) varieties of wheat were also found in Çayönü, one of the oldest settlements in Anatolia. Along with the famous einkorn wheat of Kastamonu, emmer wheat of Kars is also included in this ancient wheat group (Ertugay et al. 2020). In Kars region, local seeds and healing village products have tended to disappear as a result of modern agricultural practices. With the support of the Turkish-American Association Anatolia Foundation, which was established to keep Anatolia's products alive, since 2006, the endangered local cereal and seed varieties have been brought back and propagated by organic farming methods in the project villages. An additional one-year support was received from the United Nations (GEF Small Support Program-SGP) through the Yer Gök Anadolu Association (2007-2010), which was established locally in 2007. Nearly 450 farmer families participated in the project during this period. Today, the project continues under the leadership of village associations established for sustainability and experts, and the infrastructure for the cultivation, harvesting, packaging and marketing of the products is developed. Emmer, which is cared for by the villagers but planted in a small amount, is one of the most important products taken under protection in the project (www.karsdogal.org/kavılca-ant304k, 15 March 2019).

In the light of all this information, this review study aims to enlighten the evolutionary development process and genetic structure of Turkish emmer wheat, which is an important ancestral gene source for wheat.

2. Evolutionary Process

It is accepted that wheat was first grown as food between 8000-10000 BC. wheat cultivation is done in Turkey about 8000 years, once Turkey has

become one of the centers of the wheat. In addition, Turkey is one of the world's largest country in terms of land planted in wheat (Kan et al. 2016). Emmer group wheats have survived since the Neolithic period around 10,000 BC and it was cultivated approximately ten thousand years before its wild ancestor, *Triticum dicoccoides* (Willcox, 1998) (Fig. 1). In this process, bare wheats, especially tetraploid species, gradually changed the absorbent. Shell wheats have changed with more productive varieties of durum wheat. However, in isolated areas such as southern Russia, emmer continued to be popular until the early 1900s. Currently, emmer is an important product in Ethiopia. Emmer wheat is referred to as hulled wheat, also called 'farro' in Italy. Einkorn wheat cultivation declined in the bronze age and was generally used as bulgur or mixed with other grains as animal feed. In recent years, production of einkorn in France, India, Turkey, is limited to a small area in Yugoslavia and Italy (Harlan, 1981; Perrino and Hammer, 1982). During the Middle Bronze Age (1900-1700 BC) in Central Anatolia, it was determined that bread wheat was seen more than einkorn. The description of common wheat is based on straw residues. Einkorn wheat samples are quite common in the middle bronze age, but samples of emmer wheat are encountered in Ottoman plants during the bronze age together with rye (Nesbitt, 1993).

During evolutionary development, plants have undergone natural genomic changes caused by environmental conditions. Some DNA modifications may occur as a result of mutations resulting from environmental conditions or pollination with foreign plants (Stokstad, 2002; Conner, 2002). The formation of diploids, tetraploids and hexaploids in wheat and forage crops can be given as an example of naturally occurring genetic changes. Comparative genetic studies can explain how these changes occur. It also provides information about the relationships between genus and species (Karaca, 2002).

All over the world, and Turkey is also an important role in nutrition found in wheat, between the Euphrates and Tigris river "Fertile Crescent," it is stated that the area known as geographic spread across the world. The wild relatives of wheat in Turkey, although there is a maximum in the Southeastern Anatolia Region is widespread. Although the exact place where emmer wheat was cultivated is not known, the Northern part of the Fertile Crescent has been proposed (Özkan et al. 2002). It is accepted that it was cultivated in Karacadağ in Southeastern Anatolia for the first time in the world and then spread to the whole world from here (Özkan et al. 2005). Luo et al. (2007) stated that the domesticated emmer gene center was enriched by the gene flow from Lebanon, Southwest Syria and Israel. Genus *Triticum* L. and *Aegilops* L. have many species of 3 ploidy level, diploid (2n = 14), tetraploid (4n = 28) and hexaploid (6n = 42) (Fig. 1). 27 species of wild wheat grains present in the world there are 20 of them in

Turkey. Wild wheat species have survived until today despite the harsh environmental conditions and heavy grazing (Özkan et al. 2010).

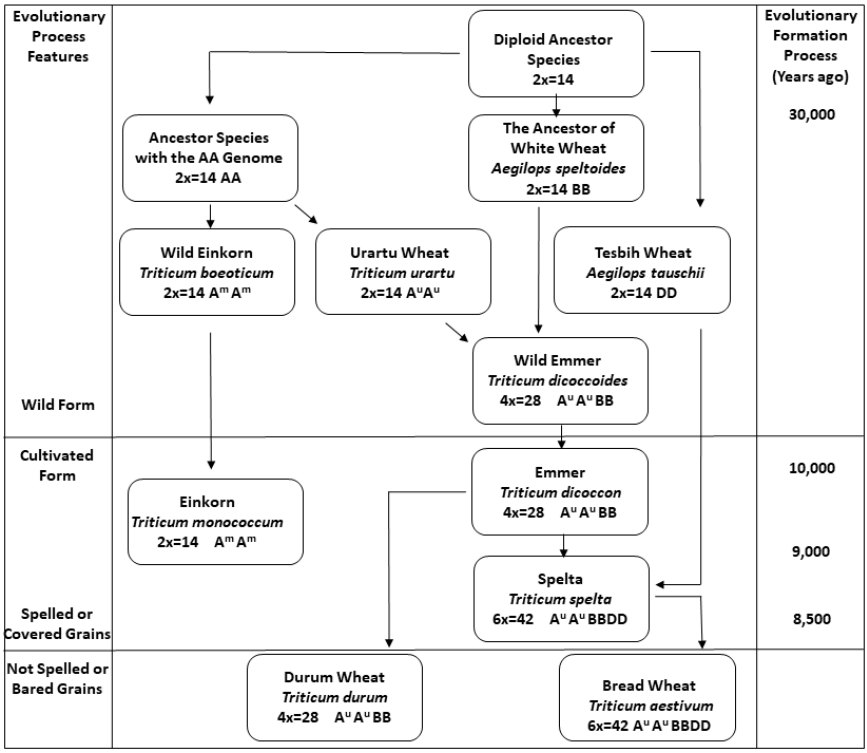


Figure 1. The Evolutionary Development Process of Cultivated Wheat Emmer (Peng et al. 2011b)

3. Scientific Classification

Emmer wheat is included in the Wheat (Gramineae = Poaceae) family. There are approximately 300 species within the *Triticum* genus. Cultured wheat and other wheat types are also included in the *Triticum* genus and the Poaceae family (Matsuoka, 2011). The first classification related to the genus *Triticum* was made by Linnaeus (1753). In his work "Species Plantarum", he included the genus *Triticum* in the Triticeae group and stated that the genus *Triticum* included the culture forms. This classification has been supported by scientists for many years.

In the other classification, Schulz (1913) grouped the wheat primarily as einkorn (spa), emmer (gernik) and bread wheat (dinkel). Then the groups; bare-grained culture forms were rearranged as spelled culture forms and wild forms (Table 1).

Table 1. Classification of *Triticum* Species by Schulz (1913)

		Cultured Wheat	
Classified	**Wild Ancestor**	**Spelled**	**Not Spelled**
Einkorn	*T. aegilopoides*	*T. monococcum*	----------
Emmer	*T. dicoccoides*	*T. dicoccon*	*T. durum*
			T. turgidum
			T. polonicum
Dinkel	Unknown	*T. spelta*	*T. compactum*
			T. vulgare

In 1929, *Aegilops* and *Triticum* species were examined by Eig (1929) and it was suggested to use the name *Aegilops* for those who could not be cultured. Eig focused more on morphological characters. He gave detailed information about the morphological characters in determining the species and varieties. He divided the genus *Aegilops* into 22 species and many varieties. In addition, 13 of the 21 species belonging to the Asian continent were found to be in Anatolia.

In the classification made by Stebbins in 1956, it was emphasized that *Aegilops* and *Triticum* genera should be classified under the genus *Triticum*.

In 1962, Zohary and Feldman announced that the *Triticum* and *Aegilops* types were in the wheat group and that today's polyploids are the forms that adapt very well to the environment. Also different polyploid species and that there are intermediate forms in Turkey stated that Turkey *Aegilops* the evolution of the field.

In 1970, Zohary pointed out the importance of wild wheat in terms of vegetative gene resources. He suggested that the *Aegilops* (22 wild species) and *Triticum* types are in the wheat group. He also stated that three of the four wild gene sources are diploid and one is tetraploid.

In 1994, Van Slageren made studies on the geographical distribution and morphological characters of Van Slageren *Triticum* and *Aegilops* genera. He evaluated the *Triticum* and *Aegilops* genera as two different genera. In order to be able to identify 22 *Aegilops* and 4 *Triticum* species, he studied their geographical distribution and morphological features.

Emmer wheat has a very complex classification within the species. Dorofeev et al. (1979) and Vavilov (1964a), four subspecies were identified on an eco-morphological and geographical basis *abyssinicum* Vav. (Abyssinian emmer), *maroccanum* Flaksb (Moroccan emmer),

asiaticum Vav. (Eastern emmer) and *dicoccum* (European emmer). Both *asiaticum* subspecies and *dicoccum* subspecies are also divided into 2 among themselves (*serbicum* (A. Schulz) Flaksb. (Volga-Balkanian emmer) - *transcaucasicum* Flaksb. (Asian emmer) and *dicoccum* (West European emmer) - *euscaldunense* Flaksb. (Pirene emmer). *Asiaticum* was divided by Vavilov (1964a) into 3 regions (*armenoanatolicum* (Ermenisatan, Anatolia), *iranicum* (North-West Iran), *carabachicum* (Karabakh)) and *maroccanum* into 2 regions (*montanum* Vav. (Mountain regions), *oasilicum* (plain regions)). *Abyssinicum* subspecies are also divided into different regions (Dorofeev et al. 1979). Afterwards, emmer wheats were divided into 3 different ecogeographic regions (Ethiopia, Yemen and Indian emmers) (Dedkova et al. 2007). Dorofeev et al. (1979) identified 8, 13, 40 and 3 types of *abyssinicum*, *asiaticum*, *dicoccum* and *marocanum* subspecies, respectively. It has been understood that these varieties are important in terms of evolutionary process and breeding (Szabo and Hammer, 1996).

Cultivated emmer wheat was described as *Triticum dicoccum* (Schrank ex Schübl) or *T. dicoccon* [syn. *T. turgidum* L. subsp. *dicoccon* (Schrank) Thell.]. There are many synonyms for this nomenclature (Zaharieva et al. 2010) (Table 2).

Table 2. Scientific Classification of Emmer Wheat

Scientific Category	**Classification**
Kingdom	Plantae
Clade	Angiosperms
Clade	Monocots
Order	Poales
Family	Poaceae
Genus	*Triticum*
Species	*T. dicoccum*
Binomial Name	*Triticum dicoccum* (Schrank ex Schübl)

4. Genetic Structure

The genome size of emmer wheat is approximately 12 Gb and it has a repeat content of 80% due to its polyploid nature and large size (Avni et al. 2017). This value is 30 times the size of rice and about 6 times the size of the corn genome. Duplications and polyploidy mechanisms in chromosome segments play a critical role in the formation of genome size in grains. Wheat included in cereals is the youngest polyploid plant

(Arumuganathan and Earle, 1991). While grain types were an ancestral species approximately 30 thousand years ago, they started to be separated from each other due to natural mutations and environmental interactions in the historical process. *Triticum monococcum* and *Triticum urartu* (diploid) diverged from each other about 10,000 years ago and found their way into two different species on earth (Fig. 1). It has been reported that *Triticum turgidum* and *Triticum timopheevii* wheat species, which are tetraploid and have AABB and AAGG genome structure, emerged after the separation of *T. monococcum* and *T.urartu* species as a result of some morphological and genetic studies (Matsuoka, 2011; Peng et al. 2011a). *T. turgidum dicoccoides* (wild ancestor of emmer) was discovered by Kornicke in southern Syria in 1873. It is stated that it was rediscovered by Aaronsohn in Lebanon, Jordan, Israel and Syria in 1910. All these discoveries proved the correctness of Candolle's suggestion that "wheat cultivation started in the Euphrates River basin" stated in 1886 (Gill et al. 2004). *Triticum urartu* (wild diploid wheat, 2n = 14, AuAu) and *Aegilops speltoides* (wild grass plant, 2n = 14, BB) hybridized naturally approximately 30-50 thousand years ago and their chromosomes were folded to *T. dicoccoides* (wild emmer, 4n = 28, AuAuBB) has occurred. As a result of the investigation of the genetic relationships of einkorn and emmer wheats, it has been determined that the gene center of these wheat species is the Southeastern Anatolia Region in Turkey (Diyarbakır-Karacadağ region). These species grow naturally in the Fertile Crescent region. As a result of natural-artificial selection made from *T. dicoccoides* species, *T. dicoccum* (cultured einkorn wheat 4n = 28 and AuAuBB) was cultivated (Heun et al. 1997; Gill et al. 2004). *T. dicoccum* Schrank (emmer wheat) is known as two grain peeled wheat. Two grain wheats represent tetraploid and hexaploid types. Somatic chromosome number in tetraploid wheats is 2n = 4x = 28. Emmer wheat, also known as einkorn, is known as emmer in Europe. It is stated that it emerged by crossing a grass of the genus *Aegilops* (goat) with *Triticum urartu* (wild red einkorn), which is close to blacky (Peng et al. 2011b). One of the wild forms in the tetraploid group is *T. dicoccoides* and carries the AABB genomes. *T. dicoccum* is the husked culture form of the group. *T. dicoccum* is regarded as the ancestor of modern durum wheat (Tahir and Pashayani, 1990).

T. dicoccoides, which is a wild tetraploid wheat species, plays an important role in the production and cultivation of durum and bread wheats, which are widely cultivated in the world today. Because this type played a main role in the domestication and cultivation of both wheat cultivars. As a result, bread and durum wheats, which are widely grown in the world today, received the gene/genome (BB genome) from *T. dicoccoides*. The *T. dicoccoides* diversification resulted in other tetraploid wheat subspecies, and *T. dicoccoides* provided the genome (AABB) to the hexaploid species.

In summary, *T. dicoccoides* has a very important place in the cultivation process of wheat (Peng et al. 2011a).

There are basically 6 species in the genus *Triticum*. There are 17 subspecies of these species and 3 of these subspecies are wild and 14 are cultural forms. While *Triticum urartu* (2n = 2x = 14, AuAu) is found only as wild form, *Triticum aestivum* (6n = 6x = 42, AuAuBBDD) and *Triticum zhukovskyi* L. (6n = 6x = 42, AmAmAuAuGG) species are found only as culture forms (Matsuoka, 2011) (Table 3).

Table 3. Species and Subspecies, Ploidy Levels, Chromosome Numbers, Genome Formulas and Vernacular Names of Cultivated and Wild Wheats (Matsuoka, 2011)

Ploidy and Chromosome Number	**Species and Subsp.**	**Genome**	**Vernacular Names**
Diploid wheat 2n=14	***Triticum monococcum L.***	**AA**	**Kaplica, Siyez**
	-subsp. *aegilopoides (boeticum*)		Wild Einkorn
	-subsp. *monococcum*		Cultivated Einkorn (Siyez wheat)
	Triticum urartu	**AA**	Wild Urartu wheat
Tetraploid wheat 4n=28 (Durum)	***Triticum turgidum L.***	**AABB**	**Gernik, Catal Siyez**
	-subsp. *dicoccoides*		Wild Emmer (Gernik wheat)
	-subsp. *dicoccon*		Cultivated Emmer (Gernik wheat)
	-subsp. *durum*		Durum wheat
	-subsp. *polonicum*		Polanya wheat (Crane beak)
	-subsp. *turanicum*		Horasan wheat
	-subsp. *turgidum*		Coarse grain wheat
	-subsp. *carthlicum (persicum)*		Persian wheat
	-subsp. *paleocolchicum*		Georgian wheat
	Triticum timopheevii	**AAGG**	Russian wheat
	-subsp. *armenicum (araraticum)*		Wild timophevi
	-subsp. *timopheevii*		Cultivated timophevi
Hekzaploid wheat 6n=42 (Bread)	***Triticum aestivum L.***	**AABBDD**	**Bread wheat**
	-subsp. aestivum		Bread wheat
	-subsp. compactum		Topbaş wheat
	-subsp. sphaerococcum		Dwarf wheat
	-subsp. macha		Maha wheat
	-subsp. spelta		Spelt wheat
	Triticum zhukovskyi L.	**AAAAGG**	Zhukovski wheat

Turkey is the center of genetic diversity for wild wheat species (*Aegilops* sp.). Wild wheat varieties, which we can find in most regions in Turkey, are used in researches on the improvement and evolution of wheat. In addition, it is of great importance in genetic improvement studies conducted to increase the quality of today's durum and common wheat (Heun et al. 1997). The wild and relative forms of the wheat in Turkey first formed the husked culture forms as a result of selection studies and natural hybridization among themselves. Later, bare-grained culture forms emerged. It is reported that there is no human intervention in the hybridization of these species and the change of chromosome numbers, that is, on the basis of the culture varieties found today. Chromosome differences and crossbreeding and folding occurred naturally as a result of natural interactions that lasted for thousands of years. Of course, in recent wheat breeding studies, new varieties have been developed by crossing the varieties within the same species with each other. When needed, new varieties can be developed by using natural and artificial hybridization methods with related species. However, today, the cultivation of transgenic (gene transferred from other organisms) commercial wheat varieties is not carried out in Turkey as well as in the world (Avsar et al. 2020).

5. CONCLUSION

Turkey is an important gene center hosting many wheat species due to its geo-strategic structure. Numerous species and subspecies of Poaceae and their wild relatives, dating back to about 10,000 years, are increasing day by day in parallel with the development of new modern varieties. In addition to meeting the basic needs of people such as nutrition, wheat varieties, which are an important cultural element in Turkish society, are still intensely cultivated today. Especially in the geography of Turkey, ancestral wheat species (emmer and einkorn wheats) grown in limited quantities in certain regions are very valuable genetic resources. One of their most important features is that they are very efficient even under harsh geographic and climatic conditions. These wheat varieties are highly resistant to various biotic and abiotic environmental conditions, they are very rich in nutritional content and most importantly they have a high level of allelic capacity for the development of new varieties. In the light of this basic information, in this chapter, information was given about the evolutionary development processes and parallel scientific classifications of emmer wheats grown in certain regions in Turkey, and their genetic structure was examined. In this sense, it is thought that this review can be a reference by contributing to the utilization of more genetic resources and the development of new varieties by introducing emmer wheat to the agriculture in Turkey as well as the scientific recognition of emmer wheat.

References

Arumuganathan, K., Earle, E. D. (1991). Nuclear DNA content of some important plant species. Plant Molecular. Biology. Rep., *9*, 208-218.

Avsar, B., Sadeghi, S., Turkec, A., and Lucas, S. J. (2020). Identification and quantitation of genetically modified (GM) ingredients in maize, rice, soybean and wheat-containing retail foods and feeds in Turkey. Journal of Food Science and Technology, 57, 787–793. https://doi.org/10.1007/s13197-019-04080-2

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O. and Gundlach, H. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. Science, *357*, 93-97.

Baloch, F. S., Alsaleh, A., Shahid, M. Q., et al. (2017). A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from Central Fertile Crescent. PloS One., *12* (1), e0167821.

Conner J. K. (2002). Genetic Mechanisms of Floral Trait Correlations in A Natural Population, Nature *420*, 407-410.

Dedkova, O. S., Badaeva, E. D., Mitrofanova, O. P., Bilinskaya, E. N., Pukhalskiy, V. A. (2007). Analysis of intraspecific diversity of cultivated emmer *Triticum dicoccum* (Schrank.) Schu¨bl. using C-banding technique. Genetika, *43*,1517– 1533.

Dorofeev, V. F., Filatenko, A. A., Migushova, E. F., Udachin, R. A., Jakubziner, M. M. (1979). Wheat. In: Dorofeev VF, Korovina ON (eds) Flora of cultivated plants, vol 1. Kolos, Leningrad, p 346.

Eig, A. (1929). Momograpisch-kritische übersiht der gattung Aegilops. Repertorium specierum novarum regni vegetabilis, Berlin.

Ertugay, M. F.,Yangılar, F. And Çebi, K. (2020). Ice cream with organic Kavilca (Buckwheat) fibre: Microstructure, thermal, physicochemical and sensory properties. Carpathian Journal of Food Science and Technology, *12* (3), 1–16. https://doi.org/10.34302//CRPJFST/2020.12.3.3

Feldman, M. and Kislev, M. E. (2007). Domestication of emmer wheat and evolution of free-threshing tetraploid wheat. Israel J. Plant Sci., *55*, 207–221.

Feldman, M. and Levy, A. A. (2015). Origin and evolution of wheat and related Triticeae species. In Alien Introgression in Wheat (Molnar-Lang, M., Ceoloni, C. and Dolezel, J., eds), pp. 21–30. Switzerland: Springer International Publishing.

Gill, B. S., Apells, R., Botha-Oberholster, A. M., Buell, C. R., Bennetzen, J. L., Chaloub, B., Chumley, F., Dvorak, J., Iwanaga, M., Keller, B., Li, W., McCombie, W. R., Ogihara, Y., Qutier, F., Sasaki, T. (2004). A workshop reporting on wheat genome sequencing: international genome research on wheat consortium, Geneticks, *168*, 1087-1096.

Haldersen, S., Akan, H., Çelik, B, and Heun, M. (2011). The climate of the Younger Dryas as a boundary for Einkorn domestication. Vegetation History and Archaeobotany, *20* (4), 305–18.

Harlan, J. R. (1981). The early history of wheat: earliest traces to the sack of Rome. In: Evans L.T. and Peacock W.J. (eds.), Wheat Science – Today and Tomorrow. Cambridge Press, 1–19.

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glemin, S., David, J. (2007). Grinding up wheat: A massive loss of nucleotide diversity since domestication. Mol Biol Evol., *24*, 1506–1517.

Heun, M., Schäfer-Pregl, R., Klawan, D., Castagna, R., Accerbi, M., Borghi, B., Salamini, F. (1997). Site of Einkorn Wheat Domestication Identified by DNA Science, *278* (5341), 1312, 1314.doi: 10.1126/science.278.5341.132.

Kan, M., Küçükçongar, M., Keser, M., Morgounov, A., Muminjanov, H., Özdemir, F., Qualset, C. (2016). Wheat Landraces in Farmers Fiels in Turkey: National Survey, Collection and Conservation, 2009-2014. FAO publication, pp. 178.

Karaca, M. (2002). Simple Sequence Repeat (SSR) Markırs Linked to the Lign Lintless (Li(1)) Mutant in Cotton, Journal Hered., *93*, 221-224.

Karagöz, A. (1996). Agronomic practices and socioeconomic aspects of emmer and einkorn cultivation in Turkey. In: Padulosi S, Hammer K, Heller J (eds) Hulled wheats, promoting the conservation and used of underutilized and neglected crops. IPGRI, Rome, pp 172–177.

Konvalina, P., Capouchova, I., Stehno, Z. and Moudry, J. (2012a). Differences in yield parameters of emmer in comparison with old and new varieties of bread wheat. African Journal of Agricultural Research, Vol. *7* (6), pp. 986-992.

Konvalina, P., Capouchova, I., Stehno, Z. (2012b). Agronomically important traits of emmer wheat. Plant Soil Environment, *58* (8), 341-346.

Linneaus, C. (1753). Species plantarum, Laurentii Salvii, Holmiae.

Luo, M. C., Yang, Z. L., You, F. M., Kawahara, T., Waines, J. G. and Dvorak, J. (2007). The structure of wild and domesticated emmer wheat populations, gene flow between them, and the site of emmer domestication. Theor. Appl. Genet., *144*, 947–959.

Mahantashıvayogayya, K., Hanchınal, R. R., and Salımath, P. M. (2004). Combining Ability in *Dicoccum* Wheat. Journal Agri. Science, *17* (3), (451-454).

Malik, S., Singh, S., (2006). Role of plant genetic resources in sustainable agriculture. Indian. J. Crop Sci., *1* (1–2), 21–28.

Matsuoka, Y. (2011). Evaluation of polyploid *Triticum* wheats under cultivation: The role of domestication, natural hybridization and allopolyploid specification in their diversification. Plant and Cell Physiology, *52* (5), 750-764.

Nesbitt, M. (1993). Ancient crop husbandry at Kaman-Kalehöyük: 1991 archaeobotanical report. In: Mikasa T (ed) Essays on Anatolian archaeology. Bull of Middle Eastern Culture Center in Japan no 7. Harrassowitz, Wiesbaden, pp 75-97.

Özkan, H., Brandolini, A., Schafer-Pregl, R. and Salamini F. (2002). AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. Mol. Biol. Evol., *19*, 1797–1801.

Özkan H, Brandolini A, Pozzi C, Effgen S, Wunder J, Salamini F (2005) A reconsideration of the domestication geography of tetraploid wheats. Theor Appl. Genet., *110*, 1052–1060.

Özkan, H., Willcox, G., Graner, A., Salamini, F., Kilian, B. (2010). Geographic distribution and domestication of wild emmer wheat (*Triticum dicoccoides*). Genetic Resources and Crop Evoluation, *58*, 11–53.

Peng, J. H., Sun, D., Nevo, E. (2011a). Will emmer wheat, *Triticum dicoccoides*, occupies a pivotal position in wheat domestication process. Australian Journal of Crop Science, *5* (9): 1127-1143.

Peng, J. H., Sun, D., Nevo, E. (2011b). Domestication, evaluation, genetics and genomics in wheat. Molecular Breeding, *28*, 281-301.

Perrino, P. and Hammer, K. (1982). *Triticum monococcum* L. and *T. dicoccum* S. are stil cultivated in İtaly. Genetica agraria, 36, 303-311.

Schulz, A. (1913). Die Geschiche der kultivierten Getreide. Halle:Louis Neberts.

Stebbins, G. L. (1956). Taxonomy and the evolution of genera, with special reference to the family Gramineae. Evolution, *10*, 235-245.

Stokstad, E. L. (2002). A Little Pollen Goes A Long Way, Science, *296*, 2314.

Szabo, A. T., Hammer, K. (1996). Notes on the taxonomy of farro: *Triticum monococcum*, *T. dicoccon* and *T. spelta*. In: Padulosi S, Hammer K, Heller J (eds) Hulled wheats, promoting the conservation and used of underutilized and neglected crops. IPGRI, Rome, pp 2–40.

Tahir, M., Pashayani, H. (1990). Transfer of agronomic traits from wild *Triticum* species to *T. turgidum* L. var. *durum* in Wheat Genetic Resources: Meeting Diverse Needs. New York, 317.

Van Slageren, M. W. (1994). Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. & Spach) Eig (Poaceae). ICARDA/Wageningen Agricultural University Paper, 94 (7), 1-512.

van Ginkel, M., Ogbonnaya, F. (2007). Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. Field Crops Res., *104*, 86–94.

Vavilov, N. I. (1964a). Mirovye resursy sortov khlebnykh zlakov, zernovykh bobovykh, l'na i ikh ispolzovanie v selektzii. (World resources of cereals, leguminous seed crops and flax, and their utilization in breeding). Nauka press, Moskow and Leningrad, 122 p (in Russian).

Willcox, G. (1998). Archaeobotanical evidence for the beginnings of agriculture in Southwest Asia. In: Damania AB, Valkoun, J, Willcox G, Qualset CO (eds) The origins of agriculture and crop domestication, ICARDA, Aleppo, Syria, ICARDA, IPGRI, FAO and UC/GRCP.

Zaharieva, M., Geleta, A. A., Al, H., Satish, C. M., Monneveux, P. (2010). Cultivated emmer wheat (*Triticum dicoccon* Schrank), an old crop with promising future. Genet. Resour. Crop Evol., *57*, 937-962.

Zaharieva, M. and Monneveux, P. (2014). Cultivated einkorn wheat (*Triticum monococcum* L. subsp. *monococcum*): the long life of a founder crop of agriculture. Genet. Resour. Crop Evol., *61*, 677–706.

Zohary, D., Feldman, M. (1962). Hybridization between amphidiploids and the evolution of polyploids in the wheat (*Aegilops-Triticum*) group. Evolution, *16*, 44-61.

Zohary, D. (1969). The progenitors of wheat and barley in relation to domestication and agricultural dispersal in the Old World.

Zohary, D. (1970). Wild wheats. Frankel, O. H., Bennett, E. (eds). Genetic resources in plant-their exploration and conservation. IBP Handbook No:11 Blackwell Scientific Publ, 239-248, Oxford.

www.karsdogal.org/kavılca-ant304k, (15 Mart 2019).