CURRENT RESEARCHES IN ARCHITECTURE AND ENGINEERING SCIENCES

EDITED BY
Assist. Prof. Dr. Çiğdem BOGENÇ

AUTHORS
Prof. Dr. Anil K. BERA
Prof. Dr. Bülent GÜLOĞLU
Assoc. Prof. Dr. Banu BEKCİ
Assist. Prof. Dr. Cüneyt YÜCELBAŞ
Assist. Prof. Dr. Çiğdem BOGENÇ
Assist. Prof. Dr. Muhammed Ernur AKINER
Assist. Prof. Dr. Şule YÜCELBAŞ
Dr. Osman DOĞAN
Lecturer Gökalp ÇINARER

İKSAD
Publishing House

CURRENT RESEARCHES IN ARCHITECTURE AND ENGINEERING SCIENCES

EDITED BY

Assist. Prof. Dr. Çiğdem BOGENÇ

AUTHORS

Prof. Dr. Anil K. BERA
Prof. Dr. Bülent GÜLOĞLU
Assoc. Prof. Dr. Banu BEKCİ
Assist. Prof. Dr. Cüneyt YÜCELBAŞ
Assist. Prof. Dr. Çiğdem BOGENÇ
Assist. Prof. Dr. Muhammed Ernur AKINER
Assist. Prof. Dr. Şule YÜCELBAŞ
Dr. Osman DOĞAN
Lecturer Gökalp ÇINARER

İKSAD
Publishing House

(The Licence Number of Publicator: 2014/31220)
TURKEY TR: +90 342 606 06 75
USA: +1 631 685 0 853
E mail: iksadyayinevi@gmail.com
www.iksadyayinevi.com

It is responsibility of the author to abide by the publishing ethics rules.
Iksad Publications – 2020©

ISBN: 978-625-7687-10-2
Cover Design: İbrahim KAYA
December / 2020
Ankara / Turkey
Size = 16 x 24 cm

CONTENTS

PREFACE

Science which reaching a set laws interms of the truth and reality is the basic element for living and being in the process of understanding life. In addition to our existing scientific knowledge, whether it is about art or engineering, the chapters in this book, which will open the doors to various fields of science, have been carefully prepared by valuable scientists. This work, which consists of researches covering the fields of design, mathematics, engineering and health sciences, prepared under the name of "Current Researches in Architecture and Engineering Sciences", will constitute an academic data source for academicians and researchers. I believe that this study, which consists of seven parts, will make significant contributions to the world of science and I would like to thank the researchers who have contributed.

Sincerely Yours

Assist. Prof. Dr. Çiğdem BOGENÇ

CHAPTER 1

SUSTAINABILITY IN LANDSCAPE PLANNING

Assoc. Prof. Dr. Banu BEKCİ[1]

[1] Recep Tayyip Erdoğan University, Faculty of Engineering and Architecture, Department of Landscape Architecture, Rize, Turkey, banu.bekci@erdogan.edu.tr. 0000-0003-1789-535X

The most important factors affecting the morphological structure of cities are the sociological structure of the society, environmental factors, and the unique values of the region, which are addressed in the city's planning strategies. The competence of planning decisions taken in shaping the morphological structure of cities during the urbanization process is high. Also, it is important on an individual basis to reflect the process and practices in the field in every planning decision. Considering the developments in technology and industry, it is also seen that planning and design approaches cause different transformations in spaces (Scott, 2008). Along with industrialization, the first planning approaches in cities were handled and the living conditions of individuals were tried to be improved. First of all, applications were made on the infrastructure and rehabilitation works of the cities, and then system studies were carried out by considering the use of natural and cultural resources of the cities. Since the planning strategies developed for the city will create a regular city in a holistic spatial structure, it will also increase individuals' living standards by preventing health problems. Individuals will be able to easily create structure and human dynamics as they lead a life within planning. As the dynamics set up in planning change with the involvement of the individual in planning, the macro form and structure of the city are also affected (Latouche, 1993). For this reason, in order for the city to form a healthy mechanism, it is necessary to create integrity in the human-space-city texture.

INTRODUCTION

The transformation of cultural data is as important as natural data in shaping planning. Cultural heritage areas, shaped by space and time, have increasingly become an expression of culture and identity (Bogenç and Bekci, 2020). These cultural heritage areas, which have always been together with individuals, develop with the influence of contemporary approaches and experiences (Harvey, 2010). The cultural diversity of a field is evaluated as an evolutionary process (Yücel Kap and Salt, 2018), which is related to tangible or intangible, spatial or temporal, technology or nature-based features that represent the continuity in the codes hidden in its identity and consists of multi-scale (local to global) approaches. Cultural landscapes can be evaluated in two groups as rural and urban landscapes. While rural landscapes reflect the vital activities, traditions, beliefs, and socio-cultural values of the local people living in a particular region, urban landscapes consist of a mosaic of ecosystems shaped by natural and social processes. In this respect, urban landscapes create an environment shaped by social structure, culture, and human behavior (Yıldız Meydan, 2018).

There is a rapid urbanization process on an unprecedented scale in today's world. This process causes severe consequences on the natural environment as well as on societies. There has been attempted to develop urban growth management policies with different qualities in order to stop or prevent external interventions in cities caused by this rapid and uncontrolled development (Perveen et. al., 2017). While

developing policies for the development of cities, local governments have started to make serious efforts to achieve sustainable results. Urban areas are the places where people use most and which are most functional (Clark and Dickson, 2003). Due to the diversity of these functionalities, sustainable management in urban growth is a very important issue for city planners and managers (Watson, 2009). A perfect urban growth policy could not be achieved due to the complexity, socialization, and economic concerns in urban ecology (Shkaruba et. al., 2017). The most common understanding of the concept of urban planning involves taking construction measures for the planning approach in urban and rural areas that aim to create preconditions for people to live together in an environment that suits them in pursuit of sociopolitical goals. In short, urban design is expressed as the collaboration of functional elements for the configuration of spatial structure (Almusaed and Almssad, 2019). In disciplines such as landscape architecture and landscape urbanism, the concept of urban planning is a newly discussed form of discipline. The relationship between the built environment and human behavior in recent urban planning approaches is as interesting as urban design and transportation planning. While theoretical, empirical, and practical studies aiming at the physical health of the population rather than the individual health of individuals, they generally aim at improving the quality of life, improving system efficiency, reducing environmental impacts (Handy et. al., 2002), and support urban planning approaches.

Ebenezer Howard, Frank Lloyd Wright, and Le Corbusier, three urban planners between 1890 and 1930, tried to answer their questions "What should be the ideal cities of the 20th century that best reflect the power and beauty of modern technology and the closest ideas of social justice? (Fishman, 2012). Howard was an enthusiastic, ambitious, benevolent socialist and a democrat supporting Wright Jefferson. Le Corbusier was a revolutionary trade unionist, and a supporter who published his designs in magazines he spoke (Barnhisel, 2005). Each planner tried to show that urban designs are rational and beautiful in their own right and social goals they believe in. These three planners shared a common assumption, each evaluating the ideal from the perspective of his social theory, his own national tradition, and personality (Lynch, 1984). This phenomenon is an evolutionary process. The urban form means the embodiment of the urban phenomenon in time and space. This phenomenon has been interpreted as a process that supports the emphasis of the city's functions, growth, shape, and appearance (Yanga et. al., 2015). The process is in continuous change. It is seen that the morphology, functional structure, and spatial order of the city are affected the most by this change.

Sustainable urban planning approaches should include very ambitious policies such as limiting energy consumption, reducing pollution, and increasing arable lands by protecting natural areas. For this purpose, reuse of urban areas and the determination of application areas more effectively come to the agenda (Bekci and Bogenç, 2019). However, the ever-increasing building stock in cities will make it difficult to use

urban development in rich countries within an ecologically sustainable and egalitarian framework (Naess, 2009). Urban environments where contact with nature is provided are valuable in their own right and in meeting other needs specific to more natural environments. The city dwellers expressed how important this situation is, as feeling nature, finding natural environments attractive, being able to use recreational areas where they can play games and feel the sense of privacy, being in the community, and experiencing the desire to communicate (Matsuoka and Kaplan, 2008). Lopes de Souza (2007) interprets this situation as a strong actor in the management approach and implementations of civil society in the sustainability process of urban planning.

None of the planning can be designed without considering environmental protection, economic development, and social inequality. Planners continue their work by considering three main objectives called this "planning triangle". They try to eliminate uncertainties by completing sustainability again so that they can complete the process comfortably. Thus, planners develop suggestions to identify and resolve conflicts in communities by confronting economic and environmental injustices (Campbell, 2007).

RESULTS

When urban planning concepts are examined today, it is seen how remarkable their contents are. The qualities that cities have in their formation constitute the diversity of these conceptual foundations (Ahvenniemi et. al., 2017). However, the differences in the user needs felt in the cities over time caused the diversification of planning theories. The most important difference between these theories is that they are shaped according to the period of need. The established theories are prepared in line with the needs and characteristics of the society and the environment, as in our laws and regulations. While the theories are listed as "Urban Planning Theory, Pluralist and Defensive Planning Theory, Comprehensive Planning Theory, Step-by-Step Planning Theory, Strategic Spatial Planning Approach", the most recently developed strategic spatial planning approach has been addressed in line with the local needs. It advocates the necessity of creating a competitive environment for local administrations to offer their services to society.

Within the scope of the study, the authorized people's opinions who involved in different stages of the planning processes of the cities on urban planning theories and processes were asked. They were asked to examine the process in line with their experiences and knowledge. The evaluations made are discussed in Table 1.

In the evaluations made, it is seen that competent people are responsive and sensitive about urban planning. These sensitive approaches are addressed through joint studies in terms of both the

implementation of the planning criteria used in the arrangement of the cities and the evaluation of the qualities of the region. Although managers are in the position of decision-makers, they are also aware of how important the support of the public and other organizations is. The most important purpose in planning decisions is service to individuals. However, its quality and sustainability are as important as the service to be provided.

Table 1. Participants' evaluation of urban planning theories

Academician (Landscape architect)	Considering ecological data in urban planning approaches, planning and design decisions should be made in accordance with the human scale and meeting the individuals' needs. Planning approaches in which all living things of the ecosystem are taken into consideration in planning decisions should be evaluated in a multidisciplinary manner. Ensuring competition in all areas of the cities may adversely affect the living standards and quality of life in cities. Therefore, international organization approaches are needed in planning as well as regional. The important thing is to develop strategies that will increase the livability of the city by preparing long-term (50-100 years) development plans.
Active NGO Member	Planning decisions compatible with the topography of the area where the city history is highlighted are necessary. In the planning decisions to be taken, light green areas with green texture should be frequently included. In addition, the designs of the building and open green areas should be designed to meet different activities and include multi-purpose uses. In "step-by-step planning concept" decisions, each situation should be evaluated within itself. Each decision regarding the environment can be gathered under a single roof in activity reports. In short, the concept of naturalness should be integrated into the development plans to be prepared, and even long-term planning decisions should be supported. However, evaluation meetings where interim evaluations will be made in planning decisions to be prepared should be included in the work package. Considering the conditions of the period, planning decisions should be discussed in the evaluation meetings, and if necessary, the process should be revised without allowing wrong practices. The process should always be under control.

Landscape architect (Public employee)	While planning the city and region, zoning permission (especially floor height) should be given by considering the topography. First of all, road widths and parking areas should be designed in line with the users' needs. Human availability should be considered as the most important planning strategy. The number of green areas per person should be increased. Planning decisions for everyone should be supported, as is the design for all approaches. Decisions should be made by examining each region and city within the scope of their unique values. "Step-by-step planning concept" decisions should be carried out step by step. Each planning and design approach should be evaluated separately, and the results should be examined. The results obtained should be used by considering the planning decisions in today's technology where the usage values after 30 or 50 years can be determined.
Municipal Manager (Environmental Education)	Cities need qualified open green spaces. The open green areas included in the planning will also positively affect the quality of life, as they will meet the children's playground and individuals' needs for sports activity in cities. "Strategic Spatial Planning Approach" decisions should be taken into account while making planning decisions. Many factors will be considered with this approach. Although each planning approach has its own separate working principles, different units must work in a coordinated manner. The more important the duration is in development plans, the more important is the quality of the planning decisions taken. Planning and design decisions should be made according to the characteristics of the areas (eg, technological designs should be given to people more in urban spaces).

(1: Do you have alternative urban planning for the city you live in?, 2: What do you think about Regional Plans, Environmental Plans and Zoning Plans from the prepared planning decisions ?, 3: Urban planning approaches "Urban Planning Theory, Pluralist and Defensive Planning Theory, What do you think about "Comprehensive Planning Theory, Step-by-Step Planning Theory, Strategic Spatial Planning Approach?" 4: How old do you suggest the development plans to be, what is your opinion?)

To provide a quality service, it is to benefit from the planning theories that have been transformed until today, whose deficiencies have been completed. Each city has a special value in its own ecosystem. Cities blended with these special features should be evaluated within themselves, and planning approaches should be determined. As a

result, every life should be shaped by the place where they live and presented to the users' service. Although the concepts of culture and society are different from each other, they are phenomena that complement and direct each other. Shared common values form people who share the common culture and interact with each other while creating a culture. Therefore, it shows us that everyone who has knowledge and experience in planning can be a determining factor in the cities' planning strategies. Having sufficient knowledge about the socio-cultural structure as well as the natural structure of the place where individuals live will positively affect future planning decisions. The evaluations have shown that the more knowledge of individuals, the higher the planning decisions to be taken for cities.

A sustainable city should be planned in a way that meets all individuals' needs and increases their well-being without harming the natural structure of the world today and in the future (Huovila et. al., 2019). There are many planning approaches utilizing science and technology for sustainability. While researching solutions to the problems encountered with sustainability, it is also focused on the dynamic interaction between nature and society by dealing with how social changes shape the environment. These approaches try to understand the components in nature-society systems one by one and reflect the basic complexity of the interaction between them (Clark and Dikson, 2003).

DISCUSSION

The protection of cultural resources is an issue related to social awareness and cultural level. Individuals' relationship with space depends on the individual's cultural knowledge and production of knowledge (Hall, 2014). It increases the social development level when combined with the freedom to create knowledge that the individual has. Globalization in the world has different effects on cities as it affects the identity in cities (Cengiz Hergül and Kahveci, 2019). The cultural changes affect societies and cause them to experience a process of cultural change by dragging them to new structures. This change causes developing cities to move away from their original values. In fact, this change in cities is a large phenomenon that includes human, society, and management issues. City managers and contributors should be aware of planning disciplines and follow developments by adapting to the rhythm of the rapidly changing world (Yavuz et. al., 2020).

In urban planning, phenomena of "contributing to a better quality of life in integrated urban planning approaches" and "promoting environmentally sensitive human health and increasing social welfare and quality of life" guide strategies (Caragliu et. al., 2011). For this reason, environmental pollution should be prevented primarily by considering environmental factors in the regulations that are planned to be made. The challenge is to understand the interrelation between the built environment and human behavior, and then develop models that can predict environmental conditions in which people will be

more physically active (Bekci et. al., 2013). Such models assist planners in designing and managing the built environment to promote physical activity (Handy et. al., 2002). The most important goal in urban space designs is to create successful open green spaces that can provide a living space. However, by meeting the needs and expectations of people with urban spaces with high standards, positive contributions are made to urban life (Alpak et. al., 2020). It is also known that the urban landscape creates a climate that affects human comfort, air quality, and energy consumption. Although it is not accepted that the effect of urban planning approaches on the creation of the climate of cities is not much, the answers sought are hidden in the climatogues, planners, and their cooperation in the planning process (Eliasson, 2000). In environmental impact assessments, climate change should be considered as a determining parameter in environmental protection and sustainable planning approaches (Sarı and Karaşah, 2020). Cities need to be monitored regularly in line with defined guidelines and laws to measure their goals and systematically monitor progress towards their goals (Munier, 2011; Bogenç and Sabaz, 2019). Cities must keep track of their annual indicators by communicating with cities through different development models to keep track of their current progress (Dameri, 2017). To be able to monitor the improvement of sustainable approaches in cities means to establish consensus among academics, researchers, and city planners who will play a role in the measurement of local administration and government policies. This consensus reflects the sustainable indicators of the environment, social systems, economics, and trends in people's

quality of life. As a result, what is important to humans is a measurable concept (Josza and Brown, 2005).

Planning methods that can combine socio-environmental issues and generate alternative visions that can manage urban growth in sustainable ways are necessary to ensure the healthy growth of cities (Perveen et. al., 2017). There are uncertainties about whether urban open green spaces and forests are sustainable in the context of rural and urban environments. For this reason, most European countries have established sheltered green spaces in surrounding cities to provide healthy recreational services to urban users. However, as open green space ecosystems are under the pressure of suburban systems, various problems have been encountered in the fragmentation of ecosystems and forest loss (Shkaruba et. al., 2017).

CONCLUSION

In the last years, administrative boundaries and urban policies cannot be defined by administrative units in the administration of cities. This process has now begun to be expressed with multi-level governance models. The models that are tried to be created are associated with national, regional, and local policies at the European level. While preparing urban development plans, human-space-environment sensitivities should be prioritized and the compatibility of cities in these systems should be taken into consideration. To increase sustainability in cities, the environmental impact of urban activities should be minimized.

The active involvement of local people in the programs prepared by the administrations is now an inevitable reality. This process forms the basis of sustainable policies that try to harmonize the relationship between the environment, population, and industrial development. Controlled growth in cities can also be defined as smart growth. Individuals who have sufficient knowledge, culture, and experience in planning should transfer the user needs to planning by passing through the mind filter.

It is seen that the planning of cities will be improved with the participation of not only planners but also decision-makers and the public. Since planning is considered as a political activity, especially as scientific knowledge, it has been concluded that some of the restrictions determined can only be resolved with institutional approaches improved in the social context of planning.

REFERENCES

Ahvenniemi, H., Huovila, A., Pinto-Seppa, I., Airaksinen, M. (2017). What are the differences between sustainable and smart cities?, Cities, 60: 234-245.

Almusaed, A., Almssad, A. (2019). City Phenomenon between Urban Structure and Composition, Book Citation Index, 29p.

Alpak, E. M., Düzenli, T., Mumcu, S. (2020). Raising awareness of seating furniture design in landscape architecture education: physical, activity-use and meaning dimensions, International Journal of Technology and Design Education, 30:587-611.

Barnhisel, G. (2015). Cold War Modernists Art, Literature, and American Cultural Diplomacy, Columbia University Press, ISBN: 978-0-231-53862-6, 336p.

Bekci, B., Bogenç, Ç. (2019). Recreatıonal Plannıng Recommendatıons for Campuses, III. Uluslararası Batı Asya Turizm Araştırmaları Kongresi.

Bekci, B., Taşkan, G., Bogenç, Ç. (2013). A habitant-focused approach to the concept of eco-city at industrialized cities with a rural city character (Bartın-Karabük) Turkey sample, Journal of food Agriculture&Enviroment, 2(11):867-872.

Bogenç, Ç., Sabaz, M. (2019). Dünya Miras Alanı Safranbolu'nun Alan Yönetim Planının Geliştirilmesi Sürecinde Tema, Hedef ve Eylemlerin Belirlenmesi, İnsan ve Toplum Bilimleri Araştırmaları Dergisi, 8(3): 1527-1544.

Bogenç, Ç., Bekci, B. (2020). Regeneration of Historical Akçaabat Ortamahalle, Journal International Refereed Journal On Social Sciences, 6(33):810-823.

Campbell, S. (2007). Green Cities, Growing Cities, Just Cities?: Urban Planning and the Contradictions of Sustainable Development, Journal of the American Planning Association, 62(3):296-312.

Caragliu, A., Del Bo, C., Nijkamp, P. (2011). Smart cities in Europe, Journal of Urban Technology, 18(2):65-82.

Clark, W.C., Dickson, N. M. (2003). Sustainability science: The emerging research program, Proceeding of the National Academy of Sciences of the United States of America, 100(14):8059–8061.

Cengiz Hergül, Ö. C., Kahveci, H. (2019). Slow City Movement against Globalization and Loss of Urban Identity, American Journal of Engineering Research, 8(2):78-83.

Dameri, R. P. (2017). Smart City Implementation: Creating Economic and Public Value in Innovative Urban Systems, Springer, ISBN-13: 978-3319833699, 167p.

Eliasson, I. (2000). The use of climate knowledge in urban planning, Landscape and Urban Planning, 48(1-2):31-44.

Fishman, R. (2012). Urban Utopies in the Twentieth Century *Ebenezer Howard, Frank Lloyd Wright, and Le Corbusier,* Blackwell Publishing Ltd., p:27-53.

Hall, P. (2014). Cities of Tomorrow: An Intellectual History of Urban Planning and Design Since 1880, Wiley-Blackwell, ISBN:978-1-118-45647-7, 640p.

Handy, S. L., Boarnet, M. G., Ewing, R., Killingworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning, Journal of Preventive Medicine, 23(2): 64-73.

Harvey, D. C. (2010). Heritage Pasts and Heritage Presents: temporality, meaning and the scope of heritage studies, International Journal of Heritage Studies, 7(4), p:319-338.

Huovila, A., Bosch, P., Airaksinen, M. (2019). Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when?, Cities, (89):141-153.

Josza, A., Brown, D. (2005). Neighborhood Sustainability Indicators Report on a BestPractice Workshop – Report. School of Urban Planning, McGill University and the Urban Ecology Center/SodemC, Montrea, QC, Canada.

Latouche, S. (1993). Dünyanın Batılılaşması, İstanbul: Ayrıntı Yayınları, ISBN: 9789755390383, 135p.

Lopes de Souza, M. (2007). Social movements as 'critical urban planning' agents, City, 10(3):327-342.

Lynch, K. (1984). Good City Form, ISBN: 0-262-62046-4 Massachusetts Institute of Technology, 493p.

Matsuoka, R. H., Kaplan, R. (2008). People needs in the urban landscape: Analysis of *Landscape And Urban Planning* contributions, Ladscape and Urban Planning, 84(1):7-19.

Munier, N. (2011). Methodology to select a set of urban sustainability indicators to measure the state of the city, and performance assessment, Ecological Indicators, 11:1020-1026.

Naess, P. (2009). Urban Planning and Sustainable Development, Eupean Planning Studies, 9(4):503-524.

Perveen, S., Kamruzzaman, M., Yiğitcanlar, T. (2017). Developing Policy Scenarios for Sustainable Urban Growth Management: A Delphi Approach, Sustainability, 9(1787):2-27.

Sarı. D., Karaşah, B. (2020). Future Adaptability of Urban Trees due to theEffects of Climate Change: the Case of Artvin,Turkey, Journal of Environmental Science and Management, 23(1):60-70.

Scott, W. (2008). Strategic Flood Risk Assessment for Greater Manchester:sub-regional assessment. Scott Wilson, Manchester.

Shkaruba, A., Kireyeu, V., Likhacheva, O. (2017). Rural–urban peripheries under socioeconomic transitions: Changing planning contexts, lasting legacies, and growing pressure. Landscape and Urban Planning, 165:244–255.

Watson, V. (2009). Seeing from the South: Refocusing Urban Planning on the Globe's Central Urban Issues, Geography, Urban Studies&Planning, 46(11):2259-2275.

Yanga, T., Lib, M., Shenb, Z. (2015). Between morphology and function: How syntactic centers of the Beijing city are defined, Journal of Urban Management, 4(2):125-134.

Yavuz, A., Canbakkal Ataoğlu, N., Acar, H. (2020). The Identification of The City on The Legibility and Wayfinding Concepts: A Case of Trabzon, Journal of Contemporary Urban Affairs, 4(2):1-12.

Yıldız Meydan, S.G. (2018). Kültürün Mekânsal Değişimler Üzerindeki Etkisi, Gazi İktisat ve İşletme Dergisi, 4(3):173-184.

Yücel Kap, S. D., Salt, E. (2018). Kültürel Peyzajların Tarihsel Süreç İçerisindeki Değişimlerinin Tespiti: Bursa İznik Örneği, Planlama, 28(1):40-55.

CHAPTER 2

A STUDIO EXPERIENCE IN BASIC DESIGN

Assist. Prof. Dr. Çiğdem BOGENÇ[1]

[1] Recep Tayyip Erdoğan University, Faculty of Fine Arts, Design, and Architecture, Department of Landscape Architecture, 53100, Rize / Turkey.
cigdem.bogenc@erdogan.edu.tr. 0000-0003-2790-6614

INTRODUCTION

Basic design education, which aims to develop the mind and hand skills of students of fine arts, architecture and design faculties in thinking, reasoning, comprehending and solving a problem, started in the Bauhaus School established in Weimar in 1919. It has progressed systematically and taken place in the curriculum of the faculties within certain programs since then. When the basic design, which emerged as one of the most basic elements of Modernism and the Bauhaus school, came to an end with Bauhaus Nazi Germany, its effects on education continued to spread (Lloyd Jones, 1969: 156). Bauhaus School has been a very effective school on architecture and related disciplines. In the field of art and design (Itten, 1975), it has been very effective in the development of worldwide design education on the design of various products from teapot to building design in many countries of the world (Cross 1983).

Sketches, prototypes, and models are developed with design methodology, which is about doing and trying something (Lloyd 2019). Design, which expresses the process of thinking, planning and realization, is expressed as thinking and forming in the mind, formulating for a purpose, finding a method, planning systematically, having a purpose, goal, intention, and creating or inventing in an artistic field requiring a high skill (Seylan, 2005, p.16).

It is aimed to develop conceptual-critical thinking skills through the use of different techniques and materials in the basic design education, in which the foundations of design skills are formed within the scope of

the education curriculum. In this process, artistic perception and design skills in which the design language is acquired through design elements and principles are developed. The basic design course, which is thought to encourage creativity, plays an important role in the education of architects (Aşaşoğlu et al.2010).

With the basic design course taught in architecture faculties, students' skills in analytical analysis, divergent thinking, intuitive approach and concentration skills can be developed and students are provided to develop original and creative methods in solving problems and to recognize their own cognitive and affective characteristics. It is aimed that students will be able to overcome design problems through these acquisitions. Thus, students' gaining the experience of generating solutions to the design problem they encounter and developing new approaches is the basic characteristics of the course (Seylan, 2005, p. 23). In art classes where it is emphasized that teaching, learning and actions cannot be considered separate from values and therefore the content is important, the understanding of socialization stands out by engaging in group works (Cengiz, et. al., 2016) and activities and doing project works.

Basic Design Education in the Faculty of Architecture

The Basic Design is a must course in the first year of many Architecture and Planning schools in Turkey and in the world. In this course, in which 2 and 3-dimensional abstract representation methods are taught, Gestalt design principles adopted by the Bauhaus school from perception psychology are generally used. However, some educational

institutions have started to apply different design methodologies today. Computational design is one of them, and it is an algorithm-based design using the computer as a tool. It is a method that separates the design process into stages such as the abstraction and reconstruction of the existing one and ensures (Bekci, et. al., 2013), the control of the process (Erbaş 2015).

The schools of Architecture and Planning in Turkey accept students assuming that creativity and design can be taught. Students are admitted with quantitative scores without an aptitude test. In this understanding, there is an assumption that creativity and design can be taught. In this sense, students' problem-solving capacity and analytical thinking skills gain importance. However, since the intuitive teaching method is used within the course, students have difficulty in the design process and creativity (Çubukçu and Gökçen 2007) because students who are accustomed to studying with written texts and formulas need to work in an abstract world consisting of lines, surfaces, volumes, colors, and textures (Günay 2007). However, the design course, which has an abstract language for expression compared to other lessons, improves students' visual perception and thinking skills in the process of transforming the information collected from the concrete world into an abstract visual expression (Tekel et al. 2015).

The design, which is the form or idea envisioned in mind, can be a shape, an object, a formation, or a form (Bogenç, et. al., 2019).. The idea that comes to the designer's mind can serve for education, entertainment purposes, and problem-solving. It is also possible for a

need to be thought over for resolution. Therefore, a process is necessary for revealing this form or this design. This process might be called the design process (Tokman, 1999: 74.75).

Gestalt Principles and Concepts used in the Course of Basic Design

"Gestalt" is a German word, which means "unified whole." The first Gestalt Principles were devised in the 1920s by German psychologists Max Wertheimer, Kurt Koffka, and Wolfgang Kohler, who aimed to understand how humans typically gain meaningful perceptions of their chaotic stimuli. The introductory course given at the Bauhaus Design School in Germany was based on the Gestalt Theory of Perception. About Gestalt, related to concepts such as integrity and context, Koffka (2000) explained that "the whole is different from the sum of the parts." This explanation, which emphasizes the independence of the whole in the perception system, is one of the most basic statements for the Gestalt theory.

Gestalt theory has had two main contributions to the design process. First, it tries to formulate the rules of visual perception by analyzing object patterns and groupings, and secondly, it formulates the principles of problem-solving and creativity. Gestalt means shape, form, pattern, or configuration in German, and Koffka defines it as the organization of sensory elements in addition to the sensory elements of a perceived object, although it is derived from something else (Koffka, 2000, 1). Since the main contribution of Gestalt theory to design was based on the analysis of the whole-piece relationship, it has also had a significant impact on learning theories. Hufnus (1999), who associates theory with

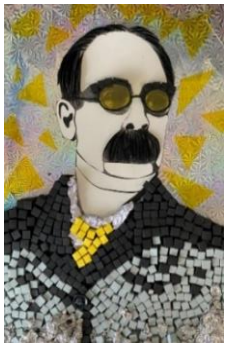
the educational world, argued that the gestalt approach in teaching is based on the relationship of the whole and piece. In the method, trainers provide an overview of the subject to be learned. They then divide the topic into its various parts and ultimately contribute (Bekci, et. al., 2019) to the whole with the parts by presenting the whole subject again (Hufnus 1999).

Basic Design Education and a Problem Solution in the Department of Landscape Architecture

This work has been shaped by the landscape architecture department of studio experience in a similar nature, which provides training courses given at the Bauhaus school of design in Turkey. After the education curriculum on basic design elements, basic design principles, Gestalt theory and design theories given to the students of the department for 14 weeks, the students said, *"You can create such a design with the elements or units you will create by organizing the elements you will choose using the whole of your field according to certain design principles; It has been requested that the images you will create are pictured with non-organic additions with new meanings arising from relationships, and the singular image is turned into reality using the collage method''.* In the 35x50cm working area, we create a portrait of a world-renowned architect in 1/100 scale with the collage method, with the collage method, the basic design elements (point, line, shape, form, color, texture, tone, etc.) and principles (balance, movement, rhythm, contrast, accent, pattern. , integrity, etc.). In the study, the collage work done by the students of the Department of Landscape Architecture in

the "Basic Design Course" and their evaluation were included. The designs for the chosen study topic have begun to concretize the given problem with the collage method, adhering to the basic design principles. In this process, the students were asked to use their talents and imaginations by completing the shortcomings of the students through mutual dialogues between the lecturer and the student. When the finished work of the students was evaluated at the end of the 8-hour study in the studio environment, it was seen that the students were able to interpret different works by using Gestalt theory and design theories and collage technique. It was seen that the collage technique, which is an artistic expression, was interpreted as follows by the students within the scope of basic design education.

Table 1. Problem solving with basic design and collage technique

Architecture/ UrbanPlanner	Photograph	Working on Architect's Desıgn Prıncıples Used ın The Study	Photography with Collage Technique
Jane JACOBS, (1916-2006) was an urbanist and activist whose writings championed a fresh, community-based approach to city building.			***Harmony*** ***Contrast*** ***Repetition***
Ebenezer HOWARD, (1850 –1928) was an English urban planner and founder of the garden city movement,.A Peaceful Path to Real Reform (1898), the description of a utopian city in which people live harmoniously together with nature.			***Harmony*** ***Contrast*** ***Repetition***
Zaha HADİD, (1950-2016) Iraqi-born British architect known for her radical deconstructivist designs. In 2004 she became the first woman to be awarded the Pritzker Architecture Prize.			***Harmony*** ***Contrast*** ***Repetition*** ***Proportion***

Antonio Gaudi CORNET (1852- 1926) was a Catalan architect known as the greatest exponent of Catalan Modernism. **Gaudi's** works have a highly individualized, sui generis style.			***Harmony*** ***Repetition***
Kevin LYNCH, (1918-1984) was an American urban planner and author. He is known for his work on the perceptual form of urban environments and was an early proponent of mental mapping.			***Harmony*** ***Repetition*** ***Balance*** ***Proportion***

CONCLUSION

Landscape architecture trainings, which are a space design discipline, are design-based trainings. Aiming to combine design with technical and functional solutions, these discipline trainings aim to provide students with design concepts and the philosophical perspective behind these concepts. Architects can convey their ideas to others through conceptual schemes, the plans and drawings of the building they designed, the open spaces they designed by landscape architects, and the city they plan. Conceptual schemes are often used in design studios as a way of thinking and communicating it. Instead of the approach in which only form concerns stand out, designs containing data such as meaning, place, history and culture can thus be made.

Based on the fact that the fact of the development of "creative intelligence", which is the main determinant in art education, is seen as a problem, the students were asked to do collage work within the scope of the lesson. When the end-of-term collage assignments of the students are examined, the basic design education given in the Department of Landscape Architecture gives the student the ability to develop himself in research, questioning, judgment, and design ideas in solving a conceptual problem with the education system of the design world in which the student will be in his future education and professional life. It has been seen that they use. The works, which are the reflection of the dialogue between mind and hand, also reflect the development of the student's imagination.

REFERENCES

Asasoğlu, A., Gür, Ş., and Erol Ş., 2010. Basic design dilemmas in architectural education. Scientific research and essays Vol. 5 (22), pp. 3538-3549, 18 November, 2010 Available online at http://www.academicjournals.org/SRE ISSN 1992-2248 ©2010 Academic Journals

Bekci, B., Bogenç, Ç., Üçok, M., 2019. Kandemir R., Kemah Kalesi ve Girlevik Şelalesi Kayaçlarının Mineralojik Yapısının Rekreasyonel Potansiyellerinin Farklı Tasarım Yaklaşımları ile İrdelenmesi, III. International Applied Social Sciences Congress, İzmir

Bogenç, Ç., Bekci, B., Üçok, M., 2019 Temel Tasar Elemanlarını Mekân Tasarımına Yansıtma Deneyimleri, III. International Applied Social Sciences Congress, İzmir

Bekci, B., Taşkan, G., Bogenç, Ç., 2013. "The effect of courtyard designs on young people, which have been made according to different functional preferences: Bartın university (Turkey)", Journal of Food, Agriculture & Environment, Vol. 11 (3&4): 1804-1813

Cengiz, B., Cengiz, C., Bekci, B., 2016. "Assessing The Visual Quality of Rural Highway Landscape Reclamation in The Western Black Sea Region of Turkey", Fresenius Environmental Bulletin, 25(3), 933-952.

Çubukçu, E. ve Gökçen Dündar, Ş. (2007) Can creativity be taught ? An empirical study on benefits of visual analogy in basic design education, ITU AZ 4, No 2 67-80.

Erbaş, S., 2015 The effect of computational design approach on computer-aided design education. Journal of research in education and teaching. Ağustos 2015 Cilt:4 Sayı:3 Makale No: 03 ISSN: 2146-9199

Cross, A., 1983. The educational background to the bauhaus. 4 (1) : Pages 43-52

Günay, B. (2007) Gestalt theory and city planning education. METU Journal of Faculty of Architecture, 24(1), 93-113.

HUFNUS, B. (1999) Gestalt theory, (http://coe.sdsu.edu/eet/Articles/ gestalt/), Department of educational technology, San Diego State University (retrieved: July 2005).

İtten, j., 1975. Design and form. John Wiley& Sons, İnc.. A VNR Book. Germany

KOFFKA, K. (2000) Introduction to: "Perception: An introduction to the Gestalt theorie" (1922), Classics in the history of psychology, an internet resource developed by Christopher D. Green, York University, Toronto, Ontario: Last revised February 2000.

Lloyd, P. J. (1969) The failure of basic design. Leonardo, 2(2), 155-160.

Lloyd P., 2019 You make it and you try it out: Seeds of design discipline futures. Design Studies, 65 (2).

Seylan, A. (2005). Basic Design. Samsun: Dağdelen

Tekel, A., Görer Tamer, N., Memlük, O. ve Ceylan Kızıltaş. A. (2015) Questioning the development process of students' visual perception skills in basic design education, art and design education symposium and workshop, interdisciplinary design, 22-27 Nisan 2015, Ankara.

Tokman, L. Y., 1999. Researching the effects of computer technology on architectural undergraduate education, doctoral thesis, Yıldız Technical University Department of Architecture, Architecture Program in Computer Environment, İstanbul, 74-75.

CHAPTER 3

A CASE STUDY IN MELEN WATERSHED FOR GETTING PRECISE DATA ACQUISITION IN CASE OF LIMITED DATA THROUGH THE COMBINED APPLICATION OF ARTIFICIAL NEURAL NETWORK AND SWAT WATER QUALITY MODEL

Asst. Prof. Dr. Muhammed Ernur AKINER[1]

[1] Akdeniz University, Vocational School of Technical Sciences, Environmental Protection and Control, Antalya, Turkey. ernurakiner@akdeniz.edu.tr . https://orcid.org/0000-0002-5192-2473

INTRODUCTION

In the area, emission sources are primarily diffuse/non-point sources. Pollution from a variety of sources is referred to as diffuse contamination in large areas. Diffuse sources make contaminants challenging to track and are typically transported with stormwater into the field (Keirle and Hayes, 2007).

Agriculture, forestry, urban, construction, mining, dams, land disposal, canals, saltwater intrusion, and runoff from city streets are typical non-point sources.

Broad discharges join the receiving waters at intervals depending on the weather conditions in a diffuse manner (Novotny and Olem, 1994; Novotny, 2003). Efficient region of dispersed contamination changes every year due to unregulated climate change and a new geological landscape (Novotny & Olem, 1994; Novotny, 2003):

There are issues in urban as well as rural areas. The critical difference between diffuse pollution and point emission is that diffuse air pollution is generally unpredictable, both temporally and spatially. Without local awareness, the runoff models and water quality literature can be used to determine the possible level of catchment contamination (Chiew and Mcmahon, 1999). The diffuse emission has no single solution. The treatment train principle, which integrates a range of complementary processing methods to improve water quality status or agricultural best management practices (BMP), is the most efficient. The most common BMPs may include the restoration of river banks to mitigate flooding, the restriction of animal access to the

creek, the development of alternate livestock water sources, the rehabilitation of stream canals, the creation of channel riffle sequences and dams, the formation of floodplain forest wetlands, the restructuring of vegetation buffers (Novotny and Hill, 2007).

The impact of nutrient enrichment in rivers will probably occur when the algals or plant growth cycles contain high levels of bioavailable nutrients (Bowes et al., 2005). Modeling and simulations are incredibly essential to forecast the long-term effect on current pollution patterns' water quality. The scientific, analytical, or statistics equations are generally used. Improvements in the water quality of waterways should be considered, and effective management procedures then suggested.

Nitrogen and phosphorus are two major plant nutrients. Although the nutrients in nature are deficient, they can contribute to the excessive growth of water plants (algal blooms). Humans can have high impacts. Blue-green algal blooms are plant, animal, and human harmful, cause decreased diversity of animals and plants and reduce dissolved oxygen levels, contributing to the killing of fish. The joint phosphorus compound is phosphate (PO_4). High levels of phosphorus will increase the number of algae and aquatic weeds in any waterway. The phosphate sources entering the rivers (PO4) include fertilizers, food waste, agricultural rotting matter, industry, sewage, soil degradation, and phosphate-based dry detergent under the Turkish Water Control Regulations (Official Gazette, 2004). Nitrates are such as phosphates, fertilizer, waste, and sewage-associated nutrients. Fertilizers and stock manure are significant sources of nitrate. Nitrates can, however, come

from the soil because of organically decomposing bacteria (ANZECC, 2000). Large quantities of nutrients resulting from human activity can contribute to eutrophication in rivers. Natural nutrients exist, but the background concentrations of various forms of the river are difficult to establish precisely.

The surface water quality can be classified by total phosphorus concentration, as provided in the Turkish Water Pollution Legislation (Official Gazette 2004) (see Table 1).

Furthermore, the following requirements for determining nutritive values are defined in Directive No. 4 of the EU Environmental Agency (EEA, 1999), as set out in Table 2. Furthermore, similar categorization is provided in the Environmental Agency Directive.

Table 1: Total Phosphorus Limitations for Surface Waters (Official Gazette, 2004).

Surface water quality class	Total Phosphorus (mg/L)
Class I (High quality)	0.02
Class II (Less polluted)	0.16
Class III (Polluted)	0.65
Class IV (Highly polluted)	>0.65

Table 2: Nutrient Values Evaluation by EEA in Marine, Coastal, and Transitional Waters (1999).

State of quality	Nitrite + nitrate (µmol L^{-1})	Phosphate (µmol L^{-1})
Good	<6.5	<0.5
Fair	6.5 to 9.0	0.5 to 0.7
Poor	9.0 to 16.0	0.7 to 1.1
Bad	>16	>1.1

A versatility strategy to set water quality goals for diverse habitats of differing health standards is suggested by the Council of Environment

and Conservation ANZECC of Australia and New Zealand. Their solution looks to the site-specific "cause principles" guideline. Nutrient, biodegradable organic matter and pH-based vulnerability in various environments are measured using trigger values. Table 3 indicates the causes shown in the ANZECC (2000) where Chl a stands for chlorophyll a, TP for total Phosphorus, FRP for filterable phosphate, TN for total nitrogen, NO_x for nitrogen oxides, NH_4^+ for ammonium, and DO for dissolved oxygen.

Table 3: Threshold Values Given by ANZECC (2000).

Ecosystem Type	Chl a ($\mu g\ L^{-1}$)	TP ($\mu g\ L^{-1}$)	FRP ($\mu g\ L^{-1}$)	TN ($\mu g\ L^{-1}$)	NO_x ($\mu g\ L^{-1}$)	NH_4^+ ($\mu g\ L^{-1}$)	DO (% saturation) Lower Limit	DO (% saturation) Upper Limit	pH Lower Limit	pH Upper Limit
Estuary	4	30	5	300	15	15	80	110	7.0	8.5
Upland river	N/A	20	15	250	15	13	90	110	6.5	7.5
Lowland river	5	50	20	500	40	20	85	110	6.5	8.0

Sustainable development firstly aims to conserve the environment; also, it addresses the needs of people. (Akiner and Akiner, 2010). Water quality depends entirely on sound ecosystems and sustainable land use in shorelines (Tezer, 2007). Chambers et al. (2000) shall summarize as follows fundamental sustainability considerations:

- The runoff from the built region shall not reduce the downstream watercourses or ecosystems of growth.
- Runoffs created by urban development within the development area should be managed in advance of disposal.

- The urban drainage schemes and methods should be developed regarding controlling water supplies and managing within the planning region.

The United Nations Food and Agriculture Organization, shortly the FAO, describes sustainable agricultural production as the protection and restoration of the natural resources base and an emphasis on technical and structural progress in a manner that ensures that human needs for current and future generations are achieved and fulfilled continuingly (FAO 1990). The management of watersheds had concentrated mainly on watersheds' physical features, such as erosion reduction, sedation reduction, further irrigation, and development. According to Campbell et al. (2004), the sustainable management of rural land has certain socioeconomic and environmental advantages as follows:

- Reduce the reliance on agricultural production subsidies.
- Increase the potential income from agri-environment schemes.
- Reduce inputs, save money, and reduce waste.
- Improve food quality and marketing opportunities reduce the risk of pollution.
- Improve landscape quality; increase the capital value of land and the tourism income.
- Improve habitat quality and help to generate income from shooting and fishing.

- Improve the relationship between the people who manage rural land.
- Reduce any health risks and improve the health of livestock and people.
- Increase biodiversity.
- Improve water quality and reduce the costs of water treatment.
- Improve soil stability, structure, and condition.
- Reduce the risk of pests and diseases and the need to use pesticides.
- Reduce runoff and increase aquifer recharge and dry-season flows.

1. METHODOLOGY

1.1. Estimate of Incomplete Rainfall Records by the Implementation of Artificial Neural Network (ANN)

For the intent of the preliminary precipitation value calculation, minimal recordings from Duzce rain gage stations were used. Data on precipitation is valid for just 23 months (from January 2008 to November 2009). There have therefore been comprehensive and continuing data for Bolu Province between 1995 and 2009. For years between 1995 and 2009, thus, missed precipitation values in Duzce were calculated using the data from Duzce as a dependent and Bolu as an independent parameter. ANN setup relies on trial and error and takes time.

The best result in terms of error minimization can be considered as an integrated ANN-architecture. In general, the training data collection builds and trains a limited network structure. By inserting neurons, the optimum network size that fulfills the training error resistance is reached. As Hung et al. (2009) stated earlier, a node number in the hidden layer ranges from (2n+1) to (2√n+m) where n is the number of the input node and m is the number of the output node (Fletcher and Goss 1993). A long trial-and-error stage has been performed to determine the optimal number of neurons and network architecture. There was a linear transfer function for missed values. However, the predicted precipitation projection was added to a non-linear model. Sigmoidal functions are the neurons' non-linear functions of the transfer. ANN model efficiency can be calculated in terms of the coefficient of correlation (R) and mean square error (MSE) (see Figure 1).

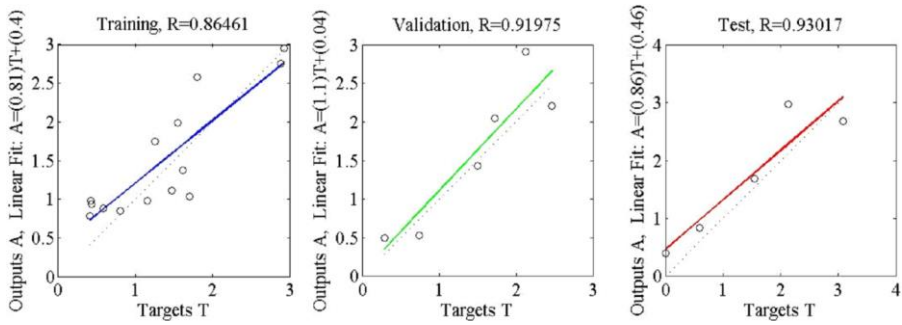


Figure 1: Training, validation, and test regression outputs of ANN for missing value estimation.

More specific neural networks can be created by smaller time series (Haofei et al., 2007). Therefore, 11 networks with a 12-month (12

elements) extended small sample was used instead of using one network with entire data between 2010 and 2020. The superior performance compared to the Levenberg-Marquardt algorithm is demonstrated by the gradient descent algorithm. The algorithm was used for ANN training in incremental mode. Linear, as well as non-linear transformations, were used (see Table 4). Hyperbolic tangents in the MLP network have continuous values between -1 and 1. The hyperbolic tangent function greatly accelerates weight learning compared to the logistic function (Fausett 1994).

For projections, both linear and non-linear models of transformation were added. On the other hand, since the linear model provides more substantial outcomes in the first half, the combined use of linear and non-linear transition functions avoids the projected data from getting trapped at the minimum peak values. This result is due to the effectiveness of linear and non-linear transfer functions together to avoid local minima.

Table 4: ANN Setup for the Projection of Missing and Potential Precipitation Value.

Parameter		**Value for missing precipitation**	**Value for forecasted precipitation**
Network configuration		1 (input) - 2 (hidden) - 2 (hidden) - 1 (output)	5 (input) - 6 (hidden) - 4 (hidden) - 1 (output)
Transfer Function		Linear transfer function at hidden 1 (Purelin) Linear transfer function at hidden 2 (Purelin) Linear transfer function at the output (Purelin)	Linear transfer function at hidden 1 (Purelin) Hyperbolic tangent sigmoid transfer function at hidden 2 (Tansig) Linear transfer function at hidden 1 (Purelin)
Training Algorithm		Levenberg-Marquardt algorithm (trainlm)	Gradient descent algorithm with adaptive learning rule (traingda)
Learning Rate		0.05	0.05
Ratio to increase the learning rate		1.05	1.05
Performance goal (MSE)		0.05	0.01
Maximum number of epochs to train		10	30
Performance	MSE	0.053	0.01 – 0.002
	R	0.865 (training), 0.919 (validation), 0.930 (test)	0.91 – 0.95

Stars signify model forecasts in Figure 2, and circles reflect the monthly average daily precipitation values observed for Duzce. The units are in millimeters/day. The ANN model has been used to estimate projected precipitation values for the province of Duzce for the years 2010 to 2020. All the projections from 1995 to 2020 can be seen in Figure 3.

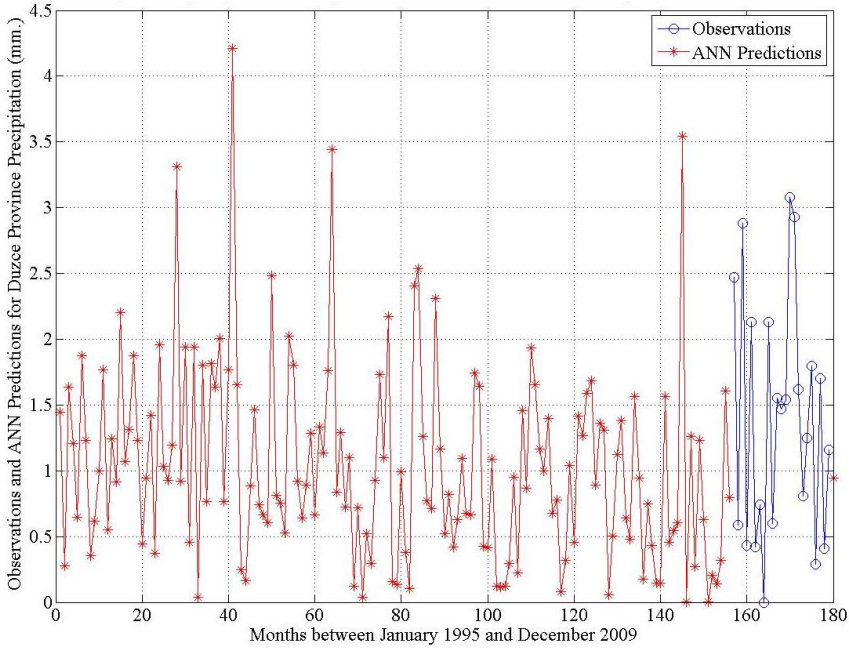


Figure 2: ANN forecasts for precipitation in Duzce from January 1995 to December 2009.

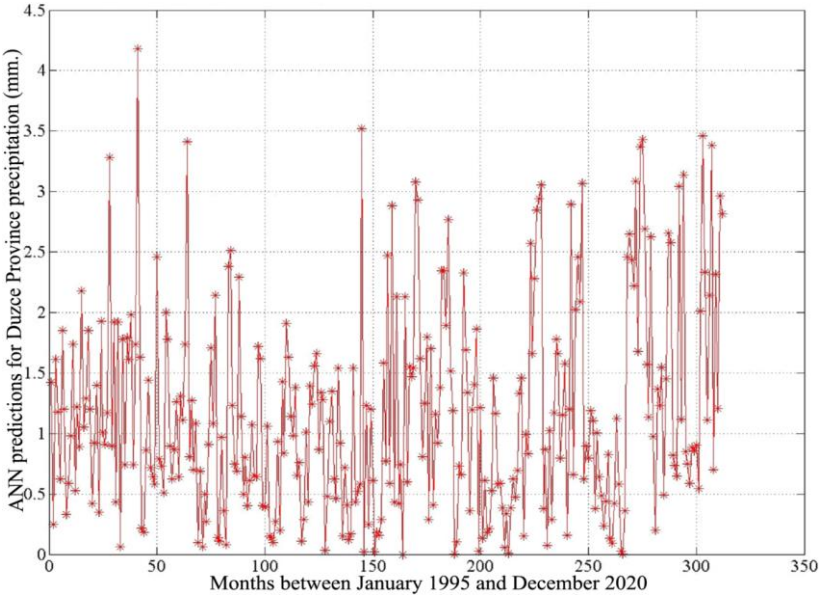


Figure 3: ANN forecast for precipitation in Duzce from January 1995 to December 2020.

The findings show that Duzce's monthly average daily precipitation rate ranges from 0 to 4.2 mm/day. This study shows that even using a small number of measurements; ANN can be an excellent method for long-term precipitation data prediction. On the other hand, the available data may not be a predictor of other environmental problems around the world; the findings of this paper indicate that this analysis may be an essential guide for evaluating, in particular, the efficacy and reliability of the precipitation prediction ANN models created.

1.2. Sensitivity Analysis

Monthly observed data of Flow, NO_3^-, NH_3, NO_2^-, CBOD, DO, Temp, TN, TKN, Total P, between 1995 and 2007 (DSI, 2011) was used for both sensitivity analysis and auto-calibration. Flow and nutrient load calibration of the model were done using available observed data for subbasin one since the outlet is located in this subbasin. Data sampled from "Büyük Melen Uğurlu Köyü" sampling point was used as an observed data since it is quite close, 4 km to the outlet of the whole watershed. Furthermore, this station has the most extensive data compared to all measuring stations (See Figure 4). The findings of the sensitivity analysis are as seen in Table 5.

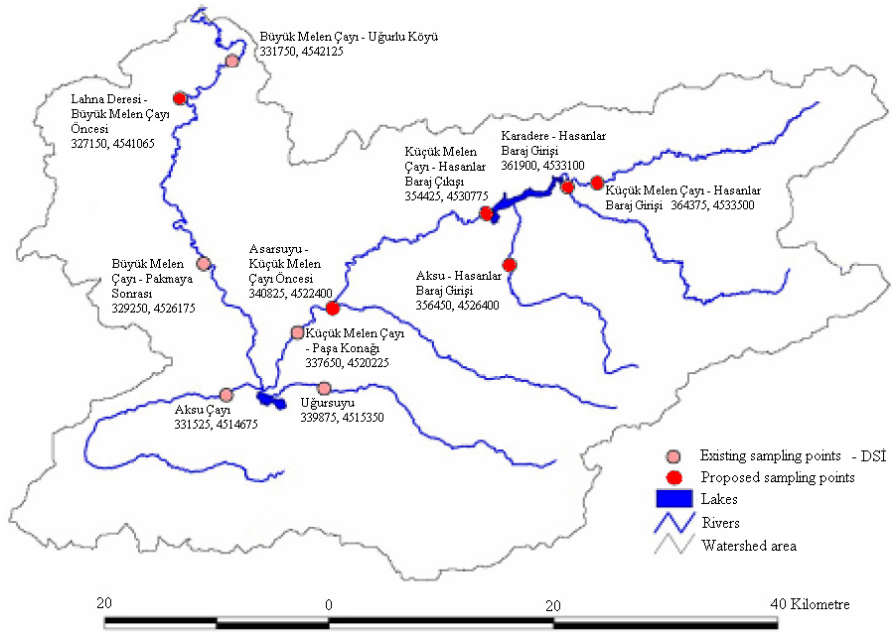


Figure 4: Sampling coordinates in WGS84 Datum UTM Coordinate Scheme 36N.

Note that parameters have the least importance among sensitivity analysis were not shown in Table 5. OF stands for Objective Function, that is, the sensitivity of that parameter compared to the observed values you provided. OUT is the parameter's sensitivity compared to the long term mean of outputs across the multiple trials. The numbers such as 12 and 13 refer to the sensitivity ranking. Usle_C falls pretty low on the ranking. A "19" on the "of 1" row, for instance, means that there are 18 other parameters that have a more significant impact on the sum of the squares of the residuals than does Usle_C. When we decide which parameters to use from the sensitivity analysis for autocalibration, we go by the "of 1" row rather than the "out 1" row, but we might be able to argue for using some kind of combination of rows. Moreover, if we re-do the sensitivity analysis with a different

seed number, we might get a different ranking. The numbers are ranking through sensitivity. So in our case, CN2 is the most sensitive parameter, Ch_K2 is the second most sensitive, and SOL_AWC follows them.

Table 5: Sensitivity Analysis Results From the Output.

	Ch_Cov	Ch_K2	Ch_N2	Cn2	Nperco	Sol_Awc	Sol_K	Sol_Z	Surlag	Usle_C	Usle_P
of 1	19	2	4	1	8	3	6	5	7	19	9
of 2	12	7	3	1	5	6	9	4	2	10	8
of 3	5	2	1	3	19	8	9	7	6	19	4
of 4	5	1	3	2	10	8	9	7	6	11	4
of 5	5	1	2	3	10	8	9	7	6	19	4
of 6	19	2	19	4	19	19	19	19	1	19	3
of 7	7	3	2	1	10	8	9	5	6	11	4
of 8	19	19	19	19	19	19	19	19	19	19	19
of 9	19	19	19	19	19	19	19	19	19	19	19
out1	19	4	7	1	8	2	5	3	6	19	9
out2	10	3	5	1	2	7	9	4	6	11	8
out3	2	5	1	4	10	8	11	6	7	9	3
out4	4	5	1	2	8	7	11	6	9	10	3
out5	2	5	1	4	9	10	11	7	6	8	3
out6	19	2	8	1	6	7	10	5	3	9	4
out7	2	6	1	4	8	10	11	7	5	9	3
out8	19	19	19	19	19	19	19	19	19	19	19
out9	19	19	19	19	19	19	19	19	19	19	19

Usually, the sensitivity analysis helps calibrate the model. That is, adjusting the parameters such that the simulated streamflow most nearly matches the observed streamflow. It's helpful because there are so many parameters to choose from; with the results (the ranking) from the sensitivity analysis, you narrow down your choices. For instance, if we need to decrease parameters need to use in calibration into 11, we can choose the following parameters, which are considered as the most effective, are selected for calibration: Cn2 (Integer Initial SCS runoff curve number for moisture condition II), Ch_K2 (Effective hydraulic conductivity in main channel alluvium

(mm/hr)), Sol_Awc (Available water capacity of the soil layer (mm H_2O/mm soil)), Ch_N2 (Manning's *n* value for the main channel), Sol_Z (Depth to bottom of soil layer (mm)), Sol_K (Saturated hydraulic conductivity of soil layer (mm/hr)), Surlag (Surface runoff lag time (days)), Usle_P (USLE equation support practice factor), Usle_C (Minimum value of USLE C factor for land cover/plant), Nperco (Nitrogen percolation coefficient), Ch_Cov (Channel cover factor) respectively.

In autocalibration, we might throw in the top ten parameters from the sensitivity analysis. An algorithm automatically finds the combination of parameter values that allows for the best or nearly best match between the simulated streamflow match to the observed streamflow. The autocalibration would then be used to "fine-tune" the parameters. However, we might have to be careful because sometimes two noticeably different sets of parameters can give nearly the same closeness of match to observed streamflow.

In terms of economic growth, agricultural development has been of crucial importance to the Melen Watershed. It is well known as a region under threat of significant degradation of natural resources if sustainability is not handled well. Water supplies compete for agricultural production, other innovations driven by humans, and local ecosystems' sustainability. Because of Istanbul's water demand, we can also expand the list of competitors.

According to the SWAT model's input criteria, daily climate data was sourced and prepared. There have been significant attempts to cover

temporary and spatial differences and identify some unmeasured climate variables from further measurements (e.g., inferior solar radiation at daytime, year-time, and latitude) in the climate data available. The hydrological model swat was run, and results are found favorable compared to the measured flow data.

1.3. Calibration Stage of the Model

Model calibration is a tool for modifying or fine-tuning model parameters to as much as possible reflect observable conditions while validating the calibrated model findings with separate data set without any further modification at various spatial and temporal scales (Neitsch et al., 2002). A three-step calibration approach has been suggested by Neitsch et al. (2002) in water quality modeling, first beginning with water balance and streamflow followed by sediment and nutrient consecutively. First of all, however, the data set's availability dictates whether the calibration should be performed on all or part of the proposed procedures.

2. FINDINGS

The following parameters were used to calibrate the SWAT model for the Melen Watershed (See Table 6). 10 subbasins and a few hundred hydrological response units (HRUs) are part of the Melen Watershed. Table 6 shows the values of the sensitive parameters given are average values specified for the entire watershed. However, it was found that there was a slight variation between the equivalent sensitive parameters defined for each HRU. After the calibration step, observed and simulated data were compared (see Figure 5-14).

Table 6: Parameters for Calibration in the SWAT Model.

Parameter	Description	Optimized Value
Parameters from sensitivity analysis		
CN2	Curve number	83 (71-86)
Ch_K2	Effective hydraulic conductivity in the main channel (mm/hr)	6-25
Sol_Awc	Available soil water capacity (mm H_2O/mm soil)	0.22
Ch_N2	Manning's *n* value for the main channels	0.05
Sol_Z(MX)	Maximum rooting depth of soil profile (mm)	500
Sol_K	Soil hydraulic conductivity (mm/hr)	460
Surlag	Surface runoff lag coefficient	4
Usle_P	USLE equation soil erodibility factor	0.7
Usle_C	Minimal value of USLE equation cover and management factor	0.001-0.03
Nperco	Nitrate percolation coefficient	0.9
Ch_Cov	Channel cover factor	0.595-0.95
Additional parameters adjusted for calibration		
ESCO	Soil evaporation compensation factor	0.95
GW_Delay	Groundwater delay time (days)	50

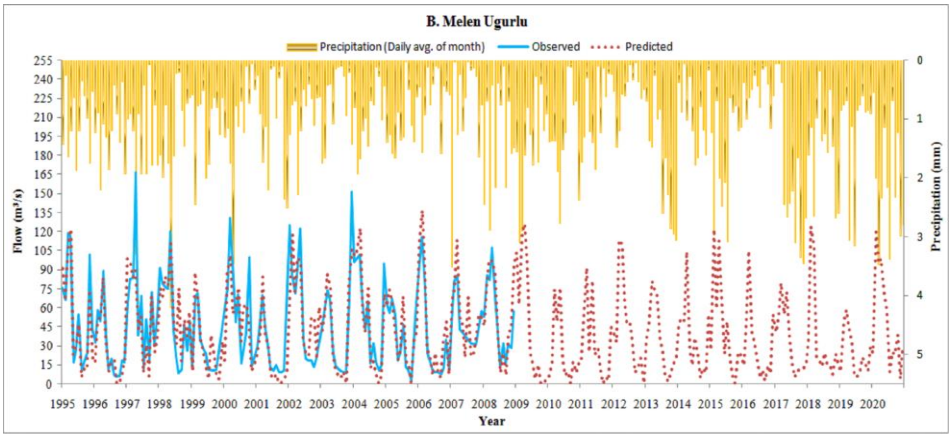


Figure 5: Flow intensity and precipitation rates simulated and measured values (daily month average value).

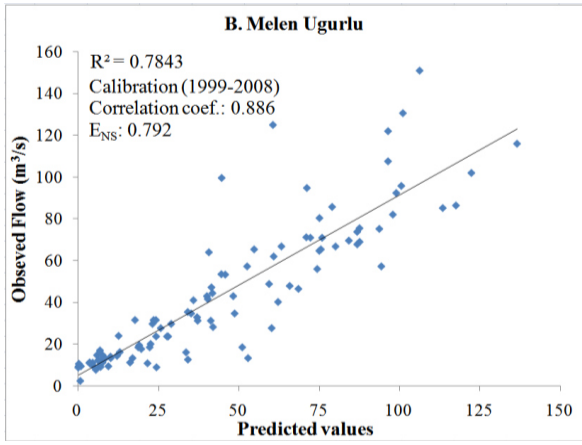


Figure 6: Watershed Outlet Calibration plots for flow rate – B. Uğurlu Köyü, Melen Çayı (1999-2008).

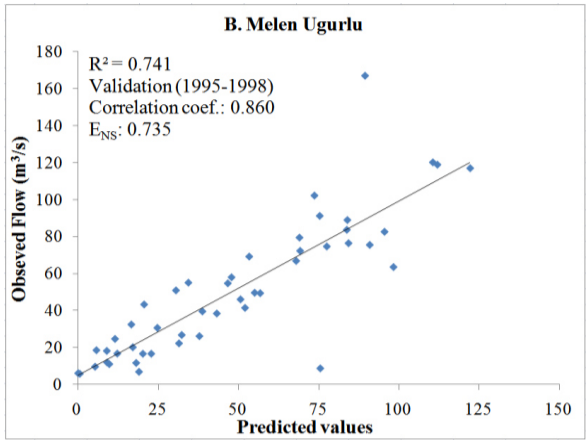


Figure 7: Validation plots at Watershed Outlet for flow rate - Büyük Melen Çayı Uğurlu Köyü (1995-1998).

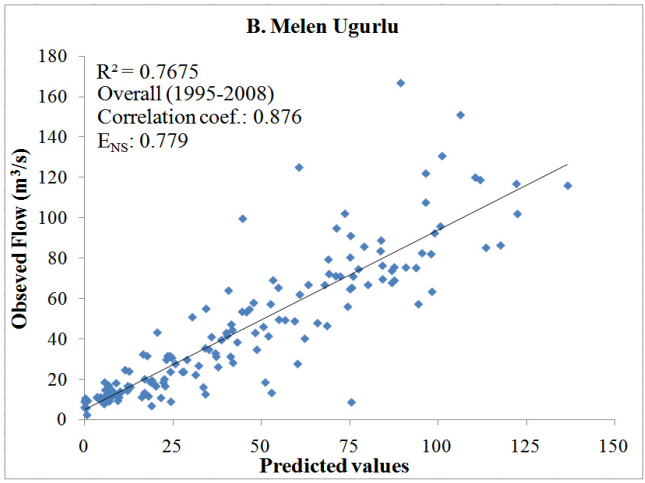


Figure 8: Watershed outlet scattering plot for observed and simulated flows – Büyük Melen Çayı Uğurlu Köyü (1995–2008).

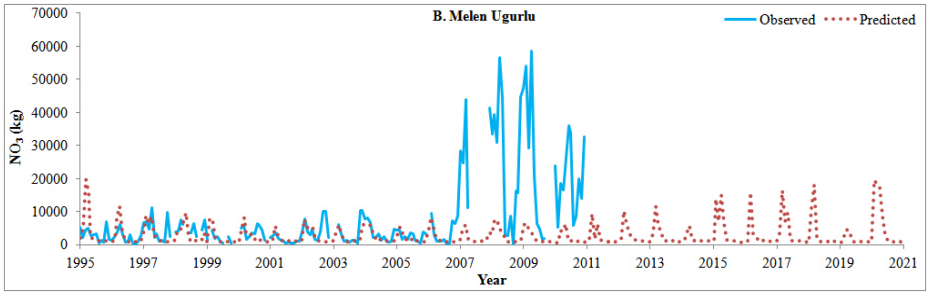


Figure 9: Values observed and simulated for average daily NO_3^- value of the month.

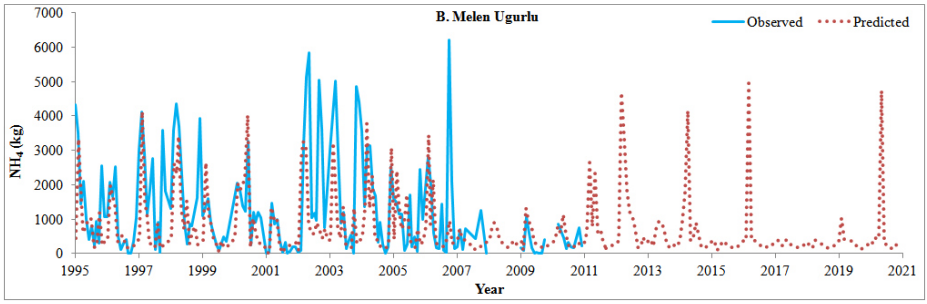


Figure 10: Values observed and simulated for daily average NH_4^+ value of the month.

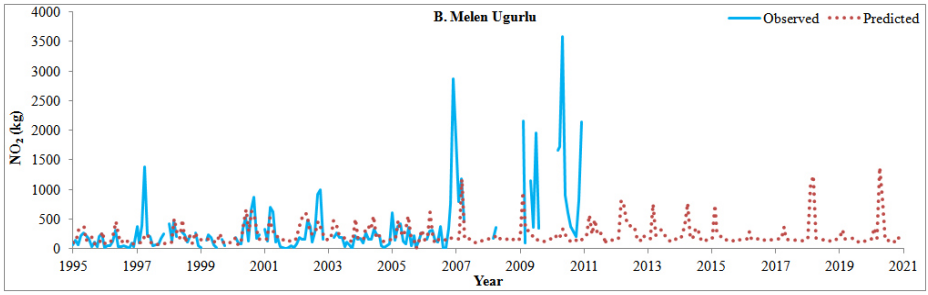


Figure 11: Values observed and simulated for average daily NO_2^- value of the month.

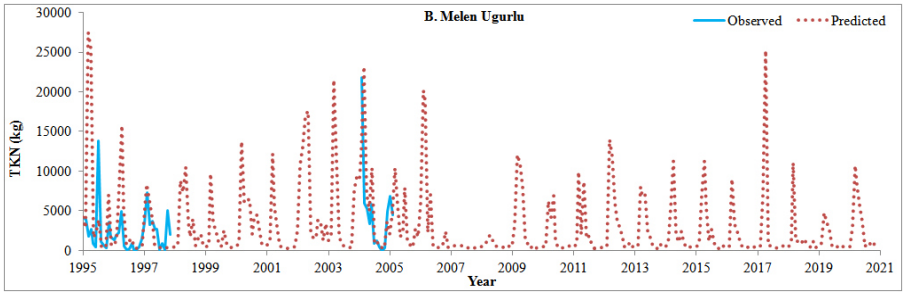


Figure 12: Observed and simulated values for daily average TKN value of the month.

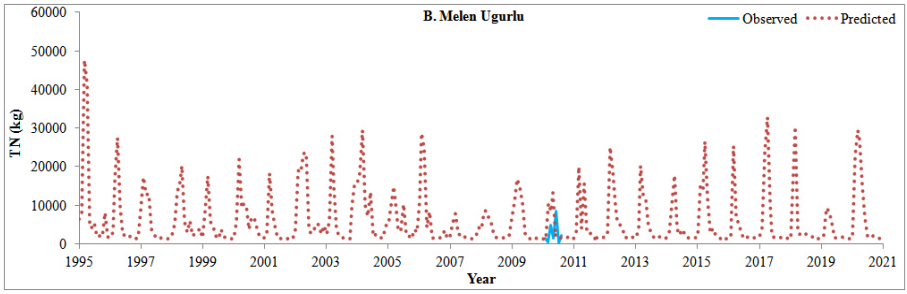


Figure 13: Values observed and simulated for daily average TN value of the month.

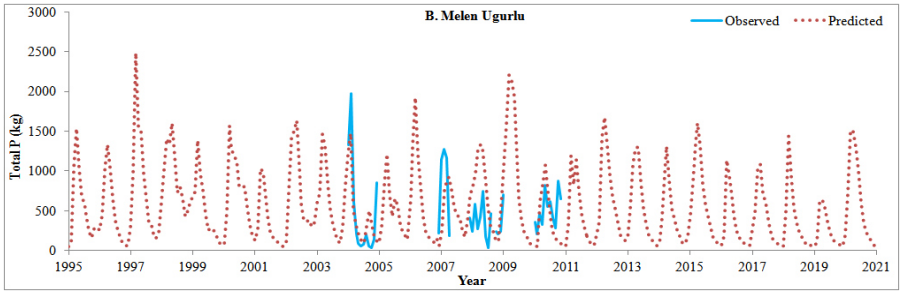


Figure 14: Values observed and simulated for daily average Total P value of the month.

CONCLUSIONS

This study revealed the Büyük Melen River's water quality and the entire basin in the Melen Basin, the source of drinking and utility water of Istanbul until 2050. Besides, a modeling study was carried out on the future regarding water quality with forwarding projections. Diffused sources have been identified as the most critical pollution source in the basin. There is almost no point source in the basin. The basin is very poor in terms of infrastructure. Made pollution in the basin analysis parameters, primarily worldwide, including Turkey's major water pollution control regulations, has been identified considering. Accordingly, nitrogen and phosphorus sources, especially nutrients, and several vital parameters were measured and analyzed. The forward projection was made through the SWAT Model (Soil and Water Assessment Tool) using temporal pollution and meteorological parameters.

The meteorological data required by the SWAT Model were partially provided, and missing data were completed by Artificial Neural Networks (ANN). Although there are water quality data measured between 1995 and 2009, a projection has been realized until 2020 through modeling. Of course, the reliability of the model has been statistically confirmed. The flow rate values obtained as a result of the modeling correspond to the measured values. The determination coefficient is around 0.8, and the model results are very reliable in this sense. Besides, nitrate, ammonium, nitrite, total Kjeldahl nitrogen, total nitrogen, total phosphorus values are also predicted by the SWAT Model that can be calibrated. In particular, some of the pollution parameters were very limited in terms of data. In this sense, the closest modeling could be done between 1995 and 2020 with the calibrated model. When the current data is compared with the model results, it is seen that the predicted and observed values overlap. This study constitutes an example that reliable results can be obtained with limited data, especially in our country, which has much trouble in terms of data supply.

REFERENCES

Akiner, I., & Akiner, E. (2010). Sustainable urban development: a case study on Goksu Delta. In *Conf. on Water Observation and Information System for Decision Support (Ohrid)*.

ANZECC, A. (2000). Australian water quality guidelines for fresh and marine waters. *Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia*.

Bowes, M. J., Hilton, J., Irons, G. P., & Hornby, D. D. (2005). The relative contribution of sewage and diffuse phosphorus sources in the River Avon catchment, southern England: implications for nutrient management. *Science of the Total Environment*, *344*(1-3), 67-81.

Chambers, B. J., Smith, K. A., & Pain, B. F. (2000). Strategies to encourage better use of nitrogen in animal manures. *Soil Use and Management*, *16*, 157-166.

Campbell, N. S., D'Arcy, B. J., Frost, C. A., Novotny, V. & Sansam, A. L. (2004). "*Diffuse Pollution: an Introduction to the Problems and the Solutions*", IWA Publishing, London.

Chiew, F. H. S., & McMahon, T. A. (1999). Modelling runoff and diffuse pollution loads in urban areas. *Water Science and Technology*, *39*(12), 241-248.

DSI (State Hydraulic Works) (2011). *1995-2010 Stream Pollution Parameters Measurements in the Melen Watershed*, Ankara, Turkey, (Unpublished report).

European Environmental Agency (EEA) (1999). *Nutrients in European Ecosystems*, Environmental Assessment Report No 4. EEA, Copenhagen.

FAO (1990). *Strategy on Water for Sustainable Agricultural Development*, United Nations, Food and Agriculture Organization, Rome.

Fausett, L. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms and Applications*, Prentice-Hall, Englewood Cliffs.

Fletcher, D., & Goss, E. (1993). Forecasting with neural networks: an application using bankruptcy data. *Information & Management*, *24*(3), 159-167.

Haofei, Z., Guoping, X., Fangting, Y., & Han, Y. (2007). A neural network model based on the multi-stage optimization approach for short-term food price forecasting in China. *Expert Systems with Applications*, *33*(2), 347-356.

Hung, N. Q., Babel, M. S., Weesakul, S., & Tripathi, N. K. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology & Earth System Sciences*, *13*(8), 1413-1425.

Keirle, R., & Hayes, C. (2007). A review of catchment management in the new context of drinking water safety plans. *Water and Environment Journal*, *21*(3), 208-216.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R., Williams, J. R., & King, K. W. (2002). *Soil and Water Assessment Tool Theoretical Documentation*, Version 2000, Grassland, Soil and Water Research Laboratory and Blackland Research Center, Temple.

Novotny, V. (2003). *Water Quality: Diffuse Pollution and Watershed Management*, John Wiley and Sons Publishing, New York.

Novotny, V., & Hill, K. (2007). Diffuse pollution abatement–a key component in the integrated effort towards sustainable urban basins. *Water science and technology*, *56*(1), 1-9.

Novotny, V. (1994). *Water quality: prevention, identification and management of diffuse pollution*. Van Nostrand-Reinhold Publishers.

Official Gazette (2004). *Water Pollution Control Regulation*, No: 25687, Ankara.

Tezer, A. (2007). Integrated Management of Aquatic Habitats: Urban Biosphere Reserve (UBR) Approach for the Omerli Watershed, Istanbul, Turkey, In: *Aquatic Habitats in Sustainable Urban Water Management: Science, Policy and Practice*, I. Wagner, J. Marshalek, P. Breil, eds., Taylor and Francis/Balkema: Leiden, pp. 200-208.

CHAPTER 4

PREDICTION OF GLIOMA GRADES USING MACHINE LEARNING ALGORITHM BASED ON RADIOMIC FEATURES

Öğr. Gör. Gökalp ÇINARER*

* Bozok University, Technical Sciences Vocational School, Computer Technologies Department, Yozgat, Turkey. gokalp.cinarer@bozok.edu.tr. https://orcid.org/0000-0003-0818-6746

INTRODUCTION

Radiomics features of MRI images are usually detected by expert radiologists. With the development of artificial intelligence applications recently, the pre-treatment use of quantitative patient imaging data has become widespread. It is an innovative method based on automatic acquisition of features of radiomics MRI images and used as a guide for the detection of prognostic factors of the disease (Forghani et al., 2019). In recent studies, it has been observed that tumor grades in different organs can be measured by radiomic analyze and the results obtained are supportive biomarkers in clinical decision making. Radiomic features have been used in the medical field such as lung cancer (Bousabarah et al., 2019; Khorrami et al., 2019), rectal cancer (Meng et al., 2019), prostate cancer (Abdollahi et al., 2019), head and neck cancer (Bogowicz et al., 2017; Mes et al., 2020), breast cancer detection (Leithner et al., 2019; Yurttakal et al., 2019), brain tumor classification (Feng, Tustison, Patel, & Meyer, 2020; Tixier, Um, Young, & Veeraraghavan, 2019).

An application known recently as radiomics has emerged as an effective factor in classification to quickly measure the properties of tumors. This article focuses on accurate identification of brain tumor grades using image processing techniques with radiomics features. The surgical timing for gliomas is related to the development of tumor-related symptoms, its rate and the patient's neurological status. Accurate identification of the brain tumor at an early stage is a key condition to provide an improved quality of treatment. In order to be

able to evaluate a brain tumor as a result of clinical examinations, it is necessary to radiologically assess its location, size, effect, properties and grade. In cases where glioma is considered radiologically; which of the surgical preferences such as stereotactic biopsy, subtotal resection, gross total resection or lobectomy, hemispherectomy depends on the surgeon's treatment strategy. The treatment method to be applied in the first stage of the tumor should be chosen correctly in order to increase the chance of survival of a tumor-exposed patient (Lara-Velazquez et al., 2017).

In this paper, there are four process are done to classification grades of gliomas. The first process is segmentation using Growcut Algorithm and wavelet , second stage is feature extraction using 3D slicer radiomic feature extraction modul, third stage is feature selection using SPSS program and fourth stage is the classification of brain tumors using Generalized Relevance Learning Vector Quantization (GRLVQ) machine learning algorithms.

The aim of this study is to determine the radiomics features of the tumor and to estimate its grades with machine learning algorithms accordingly, to help the treatment to be carried out more quickly and accurately.

1.MATERIAL AND METHOD

1.1. Dataset

The Cancer Image Archive (TCIA) is a service with public access and MRI of many types of cancer patients. There are images organized in the form of collections such as MR, CT modalities (Clark et al., 2013). LGG-1p19qDeletion (Erickson, Akkus, Sedlar, & Kofiatis, 2017) dataset was used in the study. Number of Images is 17519, while Number of Series is 478 in the datasetMRIs in the dataset are preoperative examinations in 159 people of Grade II and Grade III. All of these subjects have 1p / 19q results proven by biopsy. 121 images were used from the original dataset. 77 of them are Grade II and 44 of them are Grade III.

1.2. Segmentation

Segmentation is the most important stage of image analysis. It can be defined as dividing an image into meaningful regions. It should be noted that there is no general segmentation method that can be applied to all images and no segmentation method is perfect. In addition, methods designed for image segmentation as well as image enhancement and restoration problems and their performance vary depending on the image and application. Partitioning stops after the region of interest (ROI) of a particular application is separated and the edges of an image are acquire (Oak, 2016). It is not easy to distinguish between Grade II and III images and therefore segmentation should be done correctly.

In this study, ROIs of images of all tumor slices were extracted and classified with the software platform 3D Slicer, available for biomedical research. GrowCut segmentation algortihm used to segment tumors. GrowCut is an algorithm that uses cellular automaton theory that enables faster and more efficient work.

1.3. Feature Extraction

Wavelet transform is a system based on the main part of subband encoding. In this algorithm, low frequency subbands are enlarged to increase the image. In this way, productivity in image analysis is increased. Its characteristic features and high data processing speed are the most important factors in choosing a good technique in image processing. The 3D wavelet transform will decay the image to 8 sub-bands, which are LHH, LLL, LHL, LLH, HLL, HHL, HHH, HLH (Prochazka, Grafova, Vyšata, & Caregroup, 2011). For the current study, wavelet-based tissue identifiers from each ROI were calculated from HHH wavelet filter to classify brain tumor tissue according to density information using 3D wavelet transform. The first order features reflect the brightness, density, and color distribution of the tissues of the MR images with tumors. Eighteen first order intensity features were extracted from the enhancing lesion. Textural features are a group of features commonly used to differentiate tumor types. Seventy-five textural features were removed from five different groups Gray-Level Dependence Matrix (GLDM), Gray-Level Size Zone Matrix (GLSZM), Gray-Level Run Length Matrix (GLRLM), Neighborhood Gray Tone Difference Matrix (NGTDM) and Gray-

Level Co occurrence Matrix (GLCM). 93 features, namely first order feature (n = 18) and texture feature (n = 75), have been extracted with 3D slicer radiomic feature extraction modul.

1.4. Feature Selection

After removing all radiomic features, it is required to create a subset of the best feature associated with the result of the feature selection process. For all these processes, SPSS program was used. 12 features with $p < 0.005$ were selected from the wavelet HHH group by Mann Whitney-U test. We detected radiomics features that created significant semantics between tumor grades. In addition, properties with $p < 0.01$ are considered statistically significant (Holm, 1979).

Machine learning methods to be used for processing brain MR images can lead to improved classification quality, thereby increasing the accuracy of the diagnosis. First order and texture features are major effect for classification. A total of 12 features were selected: 2 features from the first order group, 10 features from the texture group (GLDM 4 features, GLRLM 2 features, GLSMZ 3 features, NGTDM 1 features) from 93 w-HHH radiomic features. Table 1 shows radiomics feature numbers and distribution to groups.

Table 1: Radiomics Feature Numbers and Distribution to Groups

Matrices	Selected Features	Total Features
GLCM	-	24
GLSZM	3	16
GLRLM	2	16
NGTDM	4	5
GLDM	1	14
First Order	2	18
	N=12	N=93

1.5. Classification

Classification is the process of moving each situation in a repository to a predefined set of classes or groups. Classification is an important technique used to differentiate the brain tumor grades. The purpose of the classification in this study is to accurately predict the target Grade II or Grade III for each patient from the radiomic features of the 3D images of brain tumors in the data set.

Linear Vector Quantization (LVQ) is artificial neural network architecture. It is a supervised learning method that is inspired by reinforced learning. Instead of modeling class densities, it models the discriminant functions determined by vectors and finds the closest neighbor between these vectors and data (Kohonen, 1995).

In this study, the classification was made with the Generalized Relevance Learning Vector Quantization (GRLVQ) algorithm. GLRVQ calculates each property with a relevance weight, so that the entire relevance is greater than 0 and the sum is 1. Relevance is optimized using Limited-memory Broyden–Fletcher–Goldfarb–Shanno (LBFGS) (Hammer & Villmann, 2002).

2. RESULTS

Prediction was done using the Python programming language. The main open source libraries used are scikit-learn, sklearn-lvq, seaborn, pandas, numpy, matplotlib. There are 121 patients in the dataset. 20% of the data set was reserved for testing. Table 2 shows the number of

samples in the training and test set. While 96 paitent were used for training and the other 25 patient used for testing.

Table 2: Number of Samples

Groups	Training Set	Test Set	Total
Grade II	61	16	77
Grade III	35	9	44
Total	96	25	121

Samples were normalized separately according to unit norms. Standardization was applied so that the mean value of the distribution was 0 and the standard deviation was 1. Later, GRLVQ was applied as the classifier. Table 3 shows GRLVQ hyper-parameters. Gtol is used for the successful termination of LBFGS. While the prototype per class value is 1, the beta parameter is 2.

Table 3: GRLVQ Hyper-Parameters

Prototypes per class	1
Initial prototypes	None
Initial relevancies	None
Regularization	0.0
Maximum iteration	2500
Gtol	1e-05
Beta	2

Figure 1 shows the confusion matrix resulting from the classification performed with test data. According to the figure, only 4 samples were misclassified in the test dataset. The algorithm has achieved a high rate when correctly classifying 15 of the 16 Grade II images it encountered for the first time.

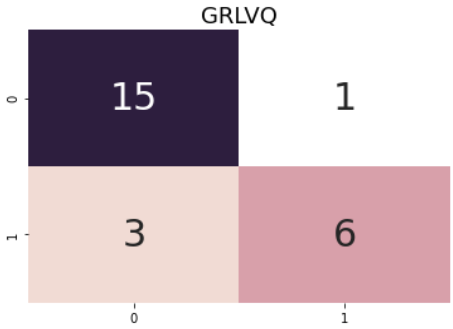


Figure 1: GRLVQ Confusion Matrix

In addition, the statistical analyses of the model was evaluated with sensitivity, specificity, precision, F1 Score. Table 4 shows other performance metrics. These values were obtained from the confusion matrix. In the table, True Positive(TP) shows True Negative(TN), while False Positive(FP) shows False Negative(FN).

Table 4: Performance metrics and derivations

Measure	Value	Derivations
Sensitivity	0.8333	TPR=TP / (TP + FN)
Specificity	0.8571	SPC=TN / (FP + TN)
Precision	0.9375	PPV=TP / (TP + FP)
F1 Score	0.8824	F1=2TP / (2TP + FP + FN)
Accuracy	0.8400	ACC = (TP + TN) / (P + N)

According to the obtained values, sensitivity, specificity, Precision values are 83.33%, 85.71%, 93.75%, while the accuracy value is 84%. In a difficult problem such as distinguishing grade types, a high accuracy rate has been achieved.

3. DISCUSSION

An experienced radiologist can differentiate tumor grades based on their experience, but an improved computer algorithm can differentiate tumor grades more quickly than the expert's performance. Besides, the developed computer classification algorithm and the method performed during its implementation can contribute to the development of the computer-aided diagnostic system.

In this article, features extracted from medical MRI images have an important role in the detection of brain tumor grades. In the literature, radiomic use in tumor detection in glioma patients is based on glioblastoma or benign disease. In a study based on neural networks, they classified normal and malignant tumors with wavelet transform properties and they achieved 0.83 accuracy (Zahran, 2014). However, in order to distinguish between normal and malignant tumors, higher results should be obtained in machine learning algorithms. In a brain tumor study examining the images of metastasis with glioblastoma, 71% classification performance was obtained with the Adaboost ensemble classifier (Abidin, Dar, D'Souza, Lin, & Wismüller, 2019). Metastasis is the spread of cancerous cells to other regions other than the tissue in which they are located, and the MRI image is different. It is easier to detect with high accuracy, but the results obtained in the study are rather low. In some studies, it was tried to estimate grades. They were generally examined as Grade (I, II, III, IV).

Zia et al. (Zia, Akhtar, & Aziz, 2018) achieved 0.85 accuracy by selecting features in determining the grade of brain cancer. Cho et al. (Cho, Lee, Kim, & Park, 2018) chose five radiomics features in glioma grading and three classifiers showed an average of 0.8866 AUC for support vector machine for test groups. Akkus et al. (Akkus et al., 2017) predicted the 1p / 19q status from MR images from image registration, tumor segmentation and Convolutional Neural Network (CNN) steps. As a result, they achieved 87.7% accuracy.

In this study is the first to perform a textural analysis of 3D MR images to estimate the grade of glioma as Grade II or Grade III. Radiomic features in glioma grading with GRLVQ machine learning method has led to higher classification success.

CONCLUSION

The use of computer technologies in medicine and the application of image processing methods have become very popular in recent years. There are many studies in this field in the literature, but in these studies, the detection of radiomic features from MR images and the active use of these features in treatment have not been fully realized. With the increase of this kind of studies, it will increase its importance in carrying out multidisciplinary studies in early detection and treatment of cancer. It is inevitable that program developers and researchers act in partnership with the medical community. The proliferation of artificial intelligence practices in the medical field will benefit humanity in the long term.

REFERENCES

Abdollahi, H., Mahdavi, S. R., Shiri, I., Mofid, B., Bakhshandeh, M., & Rahmani, K. (2019). Magnetic resonance imaging radiomic feature analysis of radiation-induced femoral head changes in prostate cancer radiotherapy. Journal of cancer research and therapeutics, 15(8), 11.

Abidin, A. Z., Dar, I., D'Souza, A. M., Lin, E. P., & Wismüller, A. (2019). Investigating a quantitative radiomics approach for brain tumor classification. Paper presented at the Medical Imaging 2019: Biomedical Applications in Molecular, Structural, and Functional Imaging.

Akkus, Z., Ali, I., Sedlář, J., Agrawal, J. P., Parney, I. F., Giannini, C., & Erickson, B. J. (2017). Predicting deletion of chromosomal arms 1p/19q in low-grade gliomas from MR images using machine intelligence. Journal of digital imaging, 30(4), 469-476.

Bogowicz, M., Leijenaar, R. T., Tanadini-Lang, S., Riesterer, O., Pruschy, M., Studer, G., . . . Lambin, P. (2017). Post-radiochemotherapy PET radiomics in head and neck cancer–the influence of radiomics implementation on the reproducibility of local control tumor models. Radiotherapy and Oncology, 125(3), 385-391.

Bousabarah, K., Temming, S., Hoevels, M., Borggrefe, J., Baus, W. W., Ruess, D., . . . Treuer, H. (2019). Radiomic analysis of planning computed tomograms for predicting radiation-induced lung injury and outcome in lung cancer patients treated with robotic stereotactic body radiation therapy. Strahlentherapie und Onkologie, 195(9), 830-842.

Cho, H.-h., Lee, S.-h., Kim, J., & Park, H. (2018). Classification of the glioma grading using radiomics analysis. PeerJ, 6, e5982.

Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., . . . Pringle, M. (2013). The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository. Journal of digital imaging, 26(6), 1045-1057.

Erickson, B., Akkus, Z., Sedlar, J., & Kofiatis, P. (2017). Data from LGG-1p19qDeletion. The Cancer Imaging Archive.

Feng, X., Tustison, N. J., Patel, S. H., & Meyer, C. H. (2020). Brain tumor segmentation using an ensemble of 3d u-nets and overall survival prediction using radiomic features. Frontiers in Computational Neuroscience, 14, 25.

Forghani, R., Savadjiev, P., Chatterjee, A., Muthukrishnan, N., Reinhold, C., & Forghani, B. (2019). Radiomics and artificial intelligence for biomarker and prediction model development in oncology. Computational and structural biotechnology journal, 17, 995.

Hammer, B., & Villmann, T. (2002). Generalized relevance learning vector quantization. Neural Networks, 15(8-9), 1059-1068.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian journal of statistics, 65-70.

Khorrami, M., Prasanna, P., Gupta, A., Patil, P., Velu, P. D., Thawani, R., . . . Fu, P. (2019). Changes in CT radiomic features associated with lymphocyte distribution predict overall survival and response to immunotherapy in non–small cell lung cancer. Cancer Immunology Research.

Kohonen, T. (1995). Learning vector quantization. In Self-organizing maps (pp. 175-189): Springer.

Lara-Velazquez, M., Al-Kharboosh, R., Jeanneret, S., Vazquez-Ramos, C., Mahato, D., Tavanaiepour, D., . . . Quinones-Hinojosa, A. (2017). Advances in brain tumor surgery for glioblastoma in adults. Brain sciences, 7(12), 166.

Leithner, D., Horvat, J. V., Marino, M. A., Bernard-Davila, B., Jochelson, M. S., Ochoa-Albiztegui, R. E., . . . Pinker, K. (2019). Radiomic signatures with contrast-enhanced magnetic resonance imaging for the assessment of breast cancer receptor status and molecular subtypes: initial results. Breast Cancer Research, 21(1), 106.

Meng, X., Xia, W., Xie, P., Zhang, R., Li, W., Wang, M., . . . Xie, Y. (2019). Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer. European radiology, 29(6), 3200-3209.

Mes, S. W., van Velden, F. H., Peltenburg, B., Peeters, C. F., Te Beest, D. E., van de Wiel, M. A., . . . Castelijns, J. A. (2020). Outcome prediction of head and

neck squamous cell carcinoma by MRI radiomic signatures. European radiology, 1-11.

Oak, R. (2016). A study of digital image segmentation techniques. International Journal of Engineering And Computer Science, 5(12).

Prochazka, A. s., Grafova, L., Vyšata, O., & Caregroup, N. (2011). Three-dimensional wavelet transform in multi-dimensional biomedical volume processing. Paper presented at the Proc. of the IASTED International Conference on Graphics and Virtual Reality, Cambridge.

Tixier, F., Um, H., Young, R. J., & Veeraraghavan, H. (2019). Reliability of tumor segmentation in glioblastoma: Impact on the robustness of MRI-radiomic features. Medical physics, 46(8), 3582-3591.

Yurttakal, A. H., Erbay, H., İkizceli, T., & Karaçavuş, S. (2019). Detection of breast cancer via deep convolution neural networks using MRI images. *Multimedia Tools and Applications*, 1-19.

Zahran, B. M. (2014). Classification of brain tumor using neural network. Int Rev Comput Softw (IRECOS), 9(4), 673-678.

Zia, R., Akhtar, P., & Aziz, A. (2018). A new rectangular window based image cropping method for generalization of brain neoplasm classification systems. International Journal of Imaging Systems and Technology, 28(3), 153-162.

CHAPTER 5

TESTING SPATIAL DEPENDENCE IN A MATRIX EXPONENTIAL SPATIAL SPECIFICATION

Prof. Dr. Anil K. BERA[1]
Dr. Osman DOĞAN[2]
Prof. Dr. Bülent GÜLOĞLU[3]

[1] University of Illinois at Urbana-Champaign, Department of Economics, Illinois, U.S.A., Email: abera@illinois.edu. ORCID ID: 0000-0003-3424-5041
[2] University of Illinois at Urbana-Champaign, Department of Economics, Illinois, U.S.A., Email: odogan@illinois.edu. ORCID ID: 0000-0002-5343-7682
3 Istanbul Technical University, Istanbul, Turkey, Email: guloglub@itu.edu.tr. ORCID ID: 0000-0002-5889-9616

INTRODUCTION

In this chapter, we develop robust Lagrange multiplier test for a matrix exponential spatial specification (MESS) that has spatial dependence both in the dependent variable and in the disturbance term. Our test can be used to test the presence of spatial dependence in the dependent variable or in the disturbance term. These tests are analogous to those developed in the Anselin et al., (1996) for the spatial autoregressive models. Our test statistic for testing the presence of spatial dependence in the dependent variable is valid irrespective of whether or not there is spatial dependence in the disturbance terms. Likewise, our test for testing spatial dependence in the disturbance term is valid in the presence of spatial dependence in the dependent variable. The computations of our tests only require the least square estimates from a linear regression model.

In the literature, it is well known that the maximum likelihood (ML) estimation of spatial autoregressive models is computationally intensive, because a Jacobian term in the likelihood function needs to be evaluated at each iteration (Anselin, 1988; Elhorst, 2010; LeSage & Pace, 2009). Estimators based on various instrumental variables (IV) and the generalized method of moments (GMM) have been considered for various spatial models as these estimators are relatively simple to compute (Doğan & Taşpınar, 2013; Kelejian & Prucha, 2010; Lee, 2007; Lee & Liu, 2010; Lin & Lee, 2010). LeSage & Pace (2007) suggest the MESS model as a computationally simple alternative to the spatial models that have spatial autoregressive terms. The

likelihood function of the MESS model is free from any Jacobian terms and can be easily evaluated. Debarsy et al. (2015) investigate the large sample properties of the MESS model under both the homoskedastic and heteroskedastic cases. They show that the quasi maximum likelihood estimator (QMLE) of the MESS model considered in LeSage & Pace (2007) is consistent even under heteroskedasticity, which is a property not shared by the QMLE of spatial models that have spatial autoregressive lags. The panel data versions of the MESS model considered in LeSage & Pace (2007) are also considered in the literature (Figueiredo & da Silva, 2015; Zhang et al., 2019). Recently, LeSage & Chih (2018) extend the MESS model by allowing the parameters to be heterogeneous and develop an efficient Bayesian Markov Chain Monte Carlo sampler for the estimation of the extended model.

Besides the computational burden, the MESS model differs in some other aspects from the spatial autoregressive models. One important difference is about the parameter space defined for the spatial parameters in these models. The spatial autoregressive models are usually considered as the empirical counterparts for the equilibrium outcome of theoretical models of interacting spatial units. In order to obtain the reduced form of the spatial autoregressive models, we need to restrict the parameter space of spatial parameters such that the inverse of certain matrices exists. On the specifications of the parameter space for spatial autoregressive models, among others, see LeSage & Pace (2009), Kelejian & Prucha (2010) and Elhorst, (2010). On the other hand, a MESS model always has a reduced form (it

always leads to a stable spatial process), i.e. we do not need to impose any restrictions on the parameter space of the spatial parameters in the MESS model (LeSage & Pace, 2009; LeSage & Pace, 2007, Debarsy et al., 2015; LeSage & Chih, 2018). Another difference is about the interpretations of parameter estimates in these models. In both MESS and spatial autoregressive models, the marginal effects accounting requires the analysis of k different $n \times n$ matrices, where k is the number of explanatory variables and n is the number of spatial units. To ease the interpretation and presentation of marginal effects, LeSage & Pace (2009) suggest three summary measures: (i) direct effects, (ii) indirect effects and (iii) total effects. In computing these effects, in the case of spatial auturegressive models, the influence of higher order neighbors decay geometrically, whereas in the case of a MESS model, it will decay exponentially (LeSage & Pace, 2009; LeSage & Pace, 2007; Debarsy et al., 2015; LeSage & Chih, 2018).

The bulk of literature on the MESS models focuses on the estimation issues as well as the identification and marginal effect accounting issues (LeSage & Pace, 2007, Debarsy et al., 2015; LeSage & Chih, 2018). The ML and Bayesian estimation approaches and related issues are considered in (LeSage & Pace, 2007, 2009), the QML and GMM approaches are suggested in Debarsy et al. (2015). It is hard to find any studies on the parameter testing and specification search issues for the MESS models. On the other hand, there is a large literature on testing and specification search exercises for spatial autoregressive models. Among others, see Anselin et al. (1996), Baltagi & Yang (2013b, 2013a), Bera, Doğan, et al. (2019), Doğan et al. (2018), Florax

et al. (2003), Taşpınar et al. (2017, 2018) and Yang (2010, 2015). The purpose of the current study is to show how robust LM test statistics can be systematically derived for the MESS models.

The rest of this study proceeds in the following way. In Section 1, we specify a first order MESS model and state its likelihood function. In Section 2, we show how the robust LM tests for testing spatial spatial dependence in the dependent and the disturbance terms can be systematically derived. Section 3 ends our study with some concluding remarks and some directions for future studies.

1. Model Specification and Estimation

We consider the following cross-sectional MESS model that has spatial dependence in the dependent and the disturbance term (for short MESS(1,1)).

$$e^{\alpha_0 W} Y = X\beta_0 + U,\ e^{\tau_0 M} U = V, \qquad (1.1)$$

where $Y = (y_1, y_2, \cdots, y_n)'$ is the $n \times 1$ vector of a dependent variable, X_n is the $n \times k$ matrix of non-stochastic exogenous variables with the matching parameter vector β_0 , W and M are the $n \times n$ spatial weights matrices of known constants with zero diagonal elements. The scalar parameters α_0 and τ_0 are called the spatial autoregressive parameters. We can call $U = (u_1, \cdots, u_n)'$ as the $n \times 1$ vector of regression disturbance terms and $V = (v_1, \cdots, v_n)'$ as the $n \times 1$ vector of disturbances (or innovations). The model is stated with the true parameter vector $\theta_0 = (\alpha_0, \tau_0, \beta_0')'$ and we use $\theta =$

$(\alpha, \tau, \beta')'$ to denote any other arbitrary parameter value in the parameter space.

Under the assumption that $V \sim N(0, \sigma_0^2 I_n)$, where I_n is the $n \times n$ identity matrix, we have $Y \sim N(X\beta_0, \sigma_0^2 e^{-\alpha_0 W} e^{-\tau_0 M} e^{-\tau_0 M'} e^{-\alpha_0 W'})$. Then, the likelihood function of the model can be expressed as

$$L(\theta, \sigma^2) = -\frac{n}{2} \ln 2\pi\sigma^2 + \ln|e^{\alpha W}| + \ln|e^{\tau W}| - \frac{1}{2\sigma^2}(e^{\alpha W}Y - X\beta)' e^{\tau M'} e^{\tau M}(e^{\alpha W}Y - X\beta). \quad (1.2)$$

The Jacobian terms $\ln|e^{\alpha W}|$ and $\ln|e^{\tau W}|$ drop from the likelihood function since $\ln|e^{\alpha W}| = e^{tr(\alpha W)} = 0$ and $\ln|e^{\tau M}| = e^{tr(\tau M)} = 0$, where $tr(\cdot)$ is the trace operator. Therefore, the likelihood function simplifies to

$$L(\theta, \sigma^2) - \frac{n}{2} \ln 2\pi\sigma^2 - \frac{1}{2\sigma^2}(e^{\alpha W}Y - X\beta)' e^{\tau M'} e^{\tau M}(e^{\alpha W}Y - X\beta). \quad (1.3)$$

The likelihood function in (1.3) indicates that we can define the MLE of θ_0 in the following way.

$$\hat{\theta} = argmin_\theta Q(\theta),$$

where $Q(\theta) = (e^{\alpha W}Y - X\beta)' e^{\tau M'} e^{\tau M}(e^{\alpha W}Y - X\beta)$. Under some assumptions (see Debarsy et al. (2015)), it can be shown that

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, \sigma_0^2 \Sigma),$$

where $\Sigma = \lim_{n\to\infty} \mathbb{E}\left(\frac{1}{n} \frac{\partial^2 Q(\theta_0)}{\partial\theta\partial\theta'}\right)$.

2. Robust Lagrange Multiplier Tests

In this section, we develop robust LM tests for testing the null hypotheses $H_0^{\alpha}: \alpha_0 = 0$ and $H_0^{\tau}: \tau_0 = 0$. The local misspecification in our setting will be denoted by the local alternative hypotheses $H_1^{\alpha}: \alpha_0 = \delta_{\alpha}/\sqrt{n}$ and $H_1^{\tau}: \tau_0 = \delta_{\tau}/\sqrt{n}$, where δ_{α} and δ_{τ} are non-stochastic bounded scalars. Our suggested test for testing $H_0^{\alpha}: \alpha_0 = 0$ will be valid irrespective of whether $H_0^{\tau}: \tau_0 = 0$ or $H_1^{\tau}: \tau_0 = \delta_{\tau}/\sqrt{n}$ holds. Similarly, we will derive a test statistic for testing $H_0^{\tau}: \tau_0 = 0$, and it will be valid under both $H_0^{\alpha}: \alpha_0 = 0$ and $H_1^{\alpha}: \alpha_0 = \delta_{\alpha}/\sqrt{n}$.

Let $\tilde{\theta} = \left(0,0,\tilde{\beta}'\right)'$ be the restricted MLE obtained under the joint null hypothesis $H_0: \alpha_0 = \tau_0 = 0$. The computation of our tests require the score functions evaluated at $\tilde{\theta}$, and a plug-in estimator of Σ obtained under $H_0: \alpha_0 = \tau_0 = 0$. Using (1.3), the required score functions evaluated at $\tilde{\theta}$ can be derived as

$$Q_{\alpha}\left(\tilde{\theta}\right) = \frac{1}{n}\frac{\partial Q(\tilde{\theta})}{\partial \alpha} = \frac{2}{n}Y'W'U\left(\tilde{\theta}\right), \tag{2.1}$$

$$Q_{\tau}\left(\tilde{\theta}\right) = \frac{1}{n}\frac{\partial Q(\tilde{\theta})}{\partial \tau} = \frac{2}{n}U'\left(\tilde{\theta}\right)MU\left(\tilde{\theta}\right), \tag{2.1}$$

where $U\left(\tilde{\theta}\right) = Y - X\tilde{\beta}$. Let $\tilde{\Sigma}$ be a consistent estimator of Σ under the joint null hypothesis $H_0 = \alpha_0 = \tau_0 = 0$. Using $\tilde{\theta}$, we can formulate a plug-in estimator of Σ. Assume that $\tilde{\Sigma}$ is partitioned into sub-matrices $\tilde{\Sigma}_{aa}$ with dimensions$k_a \times k_a$ for $a \in \{\alpha, \tau, \beta\}$, where k_a is the dimension of a, i.e.,

$$\tilde{\Sigma} = \begin{pmatrix} \tilde{\Sigma}_{\alpha\alpha} & \tilde{\Sigma}_{\alpha\tau} & \tilde{\Sigma}_{\alpha\beta} \\ \tilde{\Sigma}_{\tau\alpha} & \tilde{\Sigma}_{\tau\tau} & \tilde{\Sigma}_{\tau\beta} \\ \tilde{\Sigma}_{\beta\alpha} & \tilde{\Sigma}_{\beta\tau} & \tilde{\Sigma}_{\beta\beta} \end{pmatrix}.$$

Then, it can be shown that

$$\begin{aligned}
\tilde{\Sigma}_{\alpha\alpha} &= \tilde{\sigma}^2 tr(W^s W^s) + 2\big(WX\tilde{\beta}\big)'(WX\tilde{\beta}), \\
\tilde{\Sigma}_{\alpha\tau} &= \tilde{\sigma}^2 tr(M^s W^s), \\
\tilde{\Sigma}_{\alpha\beta} &= -2\tilde{\beta}X'W'X, \\
\tilde{\Sigma}_{\tau\tau} &= \tilde{\sigma}^2 tr(M^s M^s), \\
\tilde{\Sigma}_{\tau\beta} &= 0, \\
\tilde{\Sigma}_{\beta\beta} &= 2X'X,
\end{aligned}$$

where $A^s = A + A'$ for any square matrix A and $\tilde{\sigma}^2 = \big(Y - X\tilde{\beta}\big)'\big(Y - X\tilde{\beta}\big)/n$.

Before we state our suggested test statistics, we need to introduce some notation. Define

$$\begin{aligned}
\tilde{\Sigma}_{a\cdot c} &= \tilde{\Sigma}_{aa} - \tilde{\Sigma}_{ac}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{ca} \\
\tilde{\Sigma}_{ab\cdot c} &= \tilde{\Sigma}_{ab} - \tilde{\Sigma}_{ac}\tilde{\Sigma}_{cc}^{-1}\tilde{\Sigma}_{cb}
\end{aligned}$$

for $a, b, c \in \{\alpha, \tau, \beta\}$. Our suggested test for testing H_0^α takes the following form:

$$LM_\alpha = \frac{n\left(Q_\alpha(\tilde{\theta}) - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}_{\tau\cdot\beta}^{-1}Q_\tau(\tilde{\theta})\right)^2}{\tilde{\Sigma}_{\alpha\cdot\beta} - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}_{\tau\cdot\beta}^{-1}\tilde{\Sigma}'_{\alpha\tau\cdot\beta}}. \tag{2.3}$$

Changing the role of α and τ in (2.3), we will get the following robust test for testing H_0^τ:

$$LM_\tau = \frac{n\left(Q_\tau(\tilde{\theta}) - \tilde{\Sigma}_{\tau\alpha\cdot\beta}\tilde{\Sigma}^{-1}_{\alpha\cdot\beta}Q_\alpha(\tilde{\theta})\right)^2}{\tilde{\Sigma}_{\tau\cdot\beta} - \tilde{\Sigma}_{\tau\alpha\cdot\beta}\tilde{\Sigma}^{-1}_{\alpha\cdot\beta}\tilde{\Sigma}'_{\tau\alpha\cdot\beta}}. \tag{2.4}$$

The numerator in LM_α is the squared *adjusted score function*, and the denominator is the variance of the adjusted score function. It is clear that when $\tilde{\Sigma}_{\alpha\tau\cdot\beta} = 0$ holds, the adjusted score function $Q_\alpha(\tilde{\theta}) - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}^{-1}_{\tau\cdot\beta}Q_\tau(\tilde{\theta})$ reduces to the score function $Q_\alpha(\tilde{\theta})$, and the variance term $\tilde{\Sigma}_{\alpha\cdot\beta} - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}^{-1}_{\tau\cdot\beta}\tilde{\Sigma}'_{\alpha\tau\cdot\beta}$ to $\tilde{\Sigma}_{\alpha\cdot\beta}$. This argument shows that a non-robust test for testing H_0^α can be obtained as

$$LM^a_\alpha = \frac{n\left(Q_\alpha(\tilde{\theta})\right)^2}{\tilde{\Sigma}_{\alpha\cdot\beta}}. \tag{2.5}$$

It is well known that LM^a_α is an invalid test statistic when $\tilde{\Sigma}_{\alpha\tau\cdot\beta} \neq 0$ holds, because its asymptotic null distribution is a non-central chi-squared distribution (Bera, Bilias, et al., 2019; Bera & Yoon, 1993). In the context of our model, it can be shown that $\tilde{\Sigma}_{\alpha\tau\cdot\beta} = \tilde{\sigma}^2 tr(W^s M^s) \neq 0$ indicating that the non-robust statistic LM^a_α is an invalid test statistic. The same argument also applies to LM_τ for testing H_0^τ. In particular, $\tilde{\Sigma}_{\tau\alpha\cdot\beta} = \tilde{\sigma}^2 tr(M^s W^s) \neq 0$, suggesting that the non-robust test $LM^a_\tau = \frac{n\left(Q_\tau(\tilde{\theta})\right)^2}{\tilde{\Sigma}_{\tau\cdot\beta}}$ is an invalid test statistic, i.e., its asymptotic null-distribution is a non-central chi-squared distribution.

Given our results on the score functions and the elements of $\tilde{\Sigma}$, we can determine the explicit expressions of the test statistics in (2.3) and

(2.4). The adjusted score function with respect to α in the numerator of (2.3) is

$$Q_\alpha(\tilde{\theta}) - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}_{\tau\cdot\beta}^{-1}Q_\tau(\tilde{\theta}) = \frac{2}{n}(Y'W'U(\tilde{\theta}) - \frac{tr(W^sM^s)}{tr(M^sM^s)}U'(\tilde{\theta})MU(\tilde{\theta})$$

The variance term in the denominator of (2.3) can be derived as

$$\tilde{\Sigma}_{\alpha\cdot\beta} - \tilde{\Sigma}_{\alpha\tau\cdot\beta}\tilde{\Sigma}_{\tau\cdot\beta}^{-1}\tilde{\Sigma}'_{\alpha\tau\cdot\beta} = \tilde{\sigma}^2 tr(W^sW^s) + 2(WX\tilde{\beta})'(WX\tilde{\beta}) - 2\tilde{\beta}X'W'X(X'X)^{-1}X'WX\tilde{\beta} - \tilde{\sigma}^2\frac{tr^2(W^sM^s)}{tr(M^sM^s)}.$$

The required terms for the computation of LM_τ in (2.4) are $\tilde{\Sigma}_{\tau\alpha\cdot\beta} = \tilde{\sigma}^2 tr(M^sW^s)$, $\tilde{\Sigma}_{\tau\cdot\beta} = \tilde{\sigma}^2 tr(M^sM^s)$ and

$$\tilde{\Sigma}_{\alpha\cdot\beta} = \tilde{\sigma}^2\mathrm{tr}(\mathrm{W^sW^s}) + 2(\mathrm{WX\tilde{\beta}})'(\mathrm{WX\tilde{\beta}}) - 2\tilde{\beta}\mathrm{X'W'X(X'X)^{-1}X'WX\tilde{\beta}}.$$

Then, the explicit forms of the terms in (2.4) are

$$Q_\tau(\tilde{\theta}) - \tilde{\Sigma}_{\tau\alpha\cdot\beta}\tilde{\Sigma}_{\alpha\cdot\beta}^{-1}Q_\alpha(\tilde{\theta}) = \frac{2}{n}U'(\tilde{\theta})MU(\tilde{\theta}) - \frac{2}{n}\frac{\left(\tilde{\sigma}^2\, tr(M^sW^s)Y'W'U(\tilde{\theta})\right)}{(\tilde{\sigma}^2 tr(W^sW^s) + 2(WX\tilde{\beta})'(WX\tilde{\beta}) - 2\tilde{\beta}X'W'X(X'X)^{-1}X'WX\tilde{\beta}}$$

$$\tilde{\Sigma}_{\tau\cdot\beta} - \tilde{\Sigma}_{\tau\alpha\cdot\beta}\Sigma_{\alpha\cdot\beta}^{-1}\tilde{\Sigma}'_{\tau\alpha\cdot\beta}$$
$$= \tilde{\sigma}^2 tr(M^s M^s)$$
$$- \frac{\tilde{\sigma}^4 tr^2(M^s W^s)}{\tilde{\sigma}^2 tr(W^s W^s) + 2\left(WX\tilde{\beta}\right)'\left(WX\tilde{\beta}\right) - 2\tilde{\beta}X'W'X(X'X)^{-1}X'WX\tilde{\beta}}.$$

Under H_0^α, it can be shown that $LM_\alpha \xrightarrow{d} \chi_1^2$, where χ_1^2 is the chi-squared distribution with one degrees of freedom. Similarly, we have $LM_\tau \xrightarrow{d} \chi_1^2$ under H_0^τ. Here, it is important to note that the asymptotic null distribution of LM_α is χ_1^2 irrespective of whether H_0^τ or H_1^τ holds. Similarly, the asymptotic null distribution of LM_τ is χ_1^2 irrespective of whether H_0^α or H_1^α holds. Thus, LM_α and LM_τ tests are respectively robust to the local parametric misspecification in the forms of H_1^α and H_1^τ. These theoretical properties can be formally shown under the set of assumptions adopted in Debarsy et al. (2015) by following the asymptotic argument given in Bera et al. (2019).

Remark 1: *In terms of computation, our tests stated in (2.3) and (2.4) only requires* $\tilde{\theta} = \left(0,0,\tilde{\beta}'\right)'$, *which is the restricted MLE obtained under the joint null hypothesis* $H_0: \alpha_0 = \tau_0 = 0$. *In our case,* $\tilde{\beta}$ *is the OLS estimator obtained from the linear regression* $Y = X\beta_0 + U$, *i.e.,* $\tilde{\beta} = (X'X)^{-1}X'Y$.

Remark 2: *The analogous test statistics in* Anselin et al. (1996) *are derived for the following first-order spatial autoregressive model.*

$$Y = \lambda_0 WY + X\beta_0 + U, \qquad U = \rho_0 MU + V, \qquad (2.7)$$

where λ_0 *and* ρ_0 *are scalar spatial parameters. Let* $\tilde{\beta} = (X'X)^{-1}X'Y$ *and* $\tilde{\sigma}^2 = \tilde{V}'\tilde{V}/n$, *where* $\tilde{V} = Y - X\tilde{\beta}$. *Define*

$$\widetilde{D} = \frac{\left[\left(WX\tilde{\beta}\right)'(I_n - X(X'X)^{-1}X')WX\tilde{\beta} + \tilde{\sigma}^2 tr(W^s W)\right]}{\tilde{\sigma}^2}.$$

Then, the robust test statistic suggested by Anselin et al. (1996) *for* $H_0^{\rho}: \rho_0 = 0$ *in the local presence of* λ_0 *is given by*

$$LM_{\rho} = \frac{\left(\frac{\tilde{V}'M\tilde{V}}{\tilde{\sigma}^2} - \frac{tr(M^s W)\widetilde{D}^{-1}\tilde{V}'WY}{\tilde{\sigma}^2}\right)^2}{tr(M^s M) - tr^2(M^s W)\widetilde{D}^{-1}} \qquad (2.8)$$

The robust test statistic for $H_0^{\lambda}: \lambda_0 = 0$ *in the local presence of* ρ_0 *suggested by* Anselin et al. (1996) *is given by*

$$LM_{\lambda} = \frac{\left(\frac{\tilde{V}'WY}{\tilde{\sigma}^2} - \frac{tr(W^s M)\tilde{V}'M\tilde{V}}{\tilde{\sigma}^2 tr(M^s M)}\right)^2}{\widetilde{D} - tr^2(M^s W)/tr(M^s M)} \qquad (2.9)$$

Under the respective null hypothesis, both test statistics have chi-squared distribution with one degrees of freedom (Anselin et al., 1996). *It is clear that these test statistics cannot be used for our model in (1.1).*

CONCLUSION

In this study, we provided robust test statistics in the ML framework for testing the presence of spatial dependence in the MESS(1,1) model. Our robust test statistics are valid when there is a local parametric misspecification in the alternative model that used to construct the test statistics. Our suggested test statistics are simple and can be computed by the OLS estimator obtained from a simple linear

regression model. Though we developed our tests only for the MESS (1,1) model, our approach can be easily adopted for formulating similar tests for higher-order MESS models. Similarly, our approach can be used to develop test statistics for the panel data and heterogeneous coefficients versions of MESS models considered in the literature. We leave these extensions for future studies.

REFRENCES

Anselin, L. (1988). *Spatial econometrics: Methods and Models*. Springer.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics*, *26*(1), 77–104.

Baltagi, B. H., & Yang, Z. (2013a). Heteroskedasticity and non-normality robust LM tests for spatial dependence. *Regional Science and Urban Economics*. https://doi.org/10.1016/j.regsciurbeco.2013.05.001

Baltagi, B. H., & Yang, Z. (2013b). Standardized LM tests for spatial error dependence in linear or panel regressions. *Econometrics Journal*. https://doi.org/10.1111/j.1368-423X.2012.00385.x

Bera, A. K., Bilias, Y., Yoon, M. J., Taşplnar, S., & Doğan, O. (2019). Adjustments of Rao's Score Test for Distributional and Local Parametric Misspecifications. *Journal of Econometric Methods*. https://doi.org/10.1515/jem-2017-0022

Bera, A. K., Doğan, O., Taşpınar, S., & Leiluo, Y. (2019). Robust LM tests for spatial dynamic panel data models. *Regional Science and Urban Economics*. https://doi.org/10.1016/j.regsciurbeco.2018.08.001

Bera, A. K., & Yoon, M. J. (1993). Specification testing with locally misspecified alternatives. *Econometric Theory*. https://doi.org/10.1017/S0266466600008021

Debarsy, N., Jin, F., & Lee, L. (2015). Large sample properties of the matrix exponential spatial specification with an application to FDI. *Journal of Econometrics*, *188*(1).

Doğan, O., & Taşpınar, S. (2013). GMM estimation of spatial autoregressive models with moving average disturbances. *Regional Science and Urban Economics*, *43*(6), 903–926. https://doi.org/10.1016/j.regsciurbeco.2013.09.002

Doğan, O., Taşpınar, S., & Bera, A. K. (2018). Simple tests for social interaction models with network structures. *Spatial Economic Analysis*. https://doi.org/10.1080/17421772.2017.1374550

Elhorst, J. P. (2010). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, *5*(1), 9–28.

Figueiredo, C., & da Silva, A. R. (2015). A matrix exponential spatial specification approach to panel data models. *Empirical Economics*, *49*(1), 115–129. https://doi.org/10.1007/s00181-014-0862-2

Florax, R. J. G. M., Folmer, H., & Rey, S. J. (2003). Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics*, *33*(5), 557–579. https://doi.org/https://doi.org/10.1016/S0166-0462(03)00002-4

Kelejian, H. H., & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, *157*, 53–67.

Lee, L. (2007). {GMM} and {2SLS} estimation of mixed regressive, spatial autoregressive models. *Journal of Econometrics*, *137*(2), 489–514.

Lee, L., & Liu, X. (2010). Efficient {GMM} Estimation Of High Order Spatial Autoregressive Models With Autoregressive Disturbances. *Econometric Theory*, *26*(01), 187–230.

LeSage, J., & Chih, Y.-Y. (2018). A Matrix Exponential Spatial Panel Model with Heterogeneous Coefficients. *Geographical Analysis*, *50*(4), 422–453. https://doi.org/10.1111/gean.12152

LeSage, J., & Pace, R. K. (2007). A matrix exponential spatial specification. *Journal of Econometrics*, *140*, 190–214.

LeSage, J., & Pace, R. K. (2009a). Introduction to spatial econometrics. In *Introduction to Spatial Econometrics* (pp. 1–341). https://doi.org/10.1111/j.1467-985x.2010.00681_13.x

LeSage, J., & Pace, R. K. (2009b). *Introduction to Spatial Econometrics (Statistics: A Series of Textbooks and Monographs*. Chapman and Hall/CRC.

Lin, X., & Lee, L. (2010). GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics*, *157*(1), 34–52. https://doi.org/10.1016/j.jeconom.2009.10.035

Taşpınar, S., Doğan, O., & Bera, A. K. (2017). GMM gradient tests for spatial dynamic panel data models. *Regional Science and Urban Economics*. https://doi.org/10.1016/j.regsciurbeco.2017.04.008

Taşpınar, S., Doğan, O., & Vijverberg, W. P. M. (2018). GMM inference in spatial autoregressive models. *Econometric Reviews*, *37*(9), 931–954. https://doi.org/10.1080/00927872.2016.1178885

Yang, Z. (2010). A robust LM test for spatial error components. *Regional Science and Urban Economics*. https://doi.org/10.1016/j.regsciurbeco.2009.10.001

Yang, Z. (2015). LM tests of spatial dependence based on bootstrap critical values. *Journal of Econometrics*. https://doi.org/10.1016/j.jeconom.2014.10.005

Zhang, Y., Feng, S., & Jin, F. (2019). QML estimation of the matrix exponential spatial specification panel data model with fixed effects and heteroskedasticity. *Economics Letters*, *180*, 1–5. https://doi.org/https://doi.org/10.1016/j.econlet.2019.03.034

CHAPTER 6

DETERMINATION OF EFFECTIVE ATTRIBUTES IN THE EARLY DIAGNOSIS OF GENDER-BASED DIABETES

Asst. Prof. Dr. Cüneyt YÜCELBAŞ [1]
Asst. Prof. Dr. Şule YÜCELBAŞ [2]

[1] Corresponding Author .Electrical and Electronics Engineering Department, Hakkari University, Hakkari, TURKEY.cuneytyucelbas@hakkari.edu.tr. ORCID ID: 0000-0002-4005-6557
[2] Electrical and Electronics Engineering Department, Hakkari University, Hakkari, TURKEY.suleyucelbas@hakkari.edu.tr. ORCID ID: 0000-0002-6758-8502

1. INTRODUCTION

Diabetes mellitus is a high sugar level in the blood and the presence of sugar in the urine. The main factor that causes this disease is that the pancreas cannot produce enough insulin hormone or that this hormone cannot be used effectively (Idf, 2020; Turkdiab, 2020). Type 1 and Type 2 are two common types of this illness (Idf, 2020; Turkdiab, 2020). Type 1 diabetes usually occurs in children and adolescents. For this reason, this type of diabetes is also named as "youth diabetes" in the literature (Turkdiab, 2020). In this type, little or no insulin hormone is produced (Idf, 2020; Turkdiab, 2020). Another and most common type of diabetes is type 2. This type is usually seen in overweight people over the age of 40 (Idf, 2020; Turkdiab, 2020). This type of diabetes accounts for 90% of the patients (Idf, 2020; Turkdiab, 2020). Apart from these, another type that occurs in pregnant women is gestational diabetes (Idf, 2020; Turkdiab, 2020). In addition, people who do not fit the known diabetes types but have high blood sugar levels are classified as pre-diabetes (occult diabetes) (Turkdiab, 2020).

Acute (hypoglycemia, hypoglycemia, ketoacidosis, bacterial / fungal (fungal) infections, etc.) and chronic injuries (cardiovascular, retinopathy, nephropathy, neuropathy, etc.) can occur in people with both Type 1 and Type 2 diabetes (Turkdiab, 2020). For any reason, it is very important in the field of medicine to be able to early detect life-threatening diseases such as diabetes. Evaluations made by specialist for early detection of diseases may be insufficient for some

reasons. At this stage machine learning algorithms, which can evaluate the data like doctors and obtain extremely high results, can play an important role. Artificial intelligence-based expert systems have been used and accepted by researchers in many similar fields and purposes, especially in the early diagnosis of diseases. In (Islam et al., 2020), the data obtained from the answers given to the questions directly asked to the patients were converted into numerical expressions. Later, these data were analyzed with Naive Bayes (NB), Logistic Regression (LR), J48, and Random Forest (RF) data mining techniques. As a result, the best performance was obtained with Random Forest algorithm (Islam et al., 2020). In another study (Hossain et al., 2020a), feature analysis was performed on the data and the results were evaluated using NB, LR, j48 and RF algorithms. In the study (Hossain et al., 2020a), computational feature selection (CFS) with breadth first search (BFS) and greedy stepwise (GS) forwarding search, PCA with ranking search method, wrapper subset evaluation with GS and InfoGain with ranking search methods were applied as feature selection. In addition, the researchers developed a formula according to the risk status of the patients with the system they created and tried to select a feature with it (Hossain et al., 2020a). Thanks to the proposed system, while the classification error was obtained as 1.5%, the classification error between 23% and 27% was obtained in the other feature selection methods (Hossain et al., 2020a). Briefly in this study (Hossain et al., 2020a), it is emphasized that feature selection methods do not work. In study (Ogedengbe & Egbunu, 2020), the new data created after the feature analysis was

performed on the Pima Indian Diabetes data set with the Classifier Subset Evaluator (CSE) feature selection method, was presented to the decision tree (DT) system for training and testing. NB, Support vector machine (SVM) and DT algorithms were used to compare this hybrid system (CSE-DT). As a result, the best classification rate was obtained in CSE-DT system with 81.64% (Ogedengbe & Egbunu, 2020). In (Hossain et al., 2020b), data mining and machine learning algorithms were tested on data sets in various fields, including diabetes data. The main purpose of the authors has been to identify and apply algorithms that can be used in the field of Internet of Things (IoT) (Hossain et al., 2020b). In briefly, the verification of machine learning algorithms for obtaining the best results in studies using different data sets was performed (Hossain et al., 2020b).

The process steps of this study are seen in Figure 1. The data used in this study were taken from the University of California Irvine (UCI) Machine Learning Repository (Dua, 2019). The effect of gender factor was investigated in detail by presenting diabetic data with 14 attributes (Islam et al., 2020) to the K-Nearest Neighbor (KNN) classifier for early detection of this disease. The aim of this study was to determine which features by gender are effective in the early diagnosis of diabetes. Wrapper subset evaluator feature selection algorithm was applied to the data according to the best first search algorithm. Before the feature selection process, the accuracy rates of 97.86% and 96.35% were obtained for male and female, respectively. As a result of the procedures applied later, a classification rate of 96.87% was achieved for female with only the

features of "Polyuria", "delayed healing" and "Alopecia". For male, it was determined that the "weakness" feature was not effective on the result of the classification, and the same rate (97.86%) was obtained this time with 13 attributes.

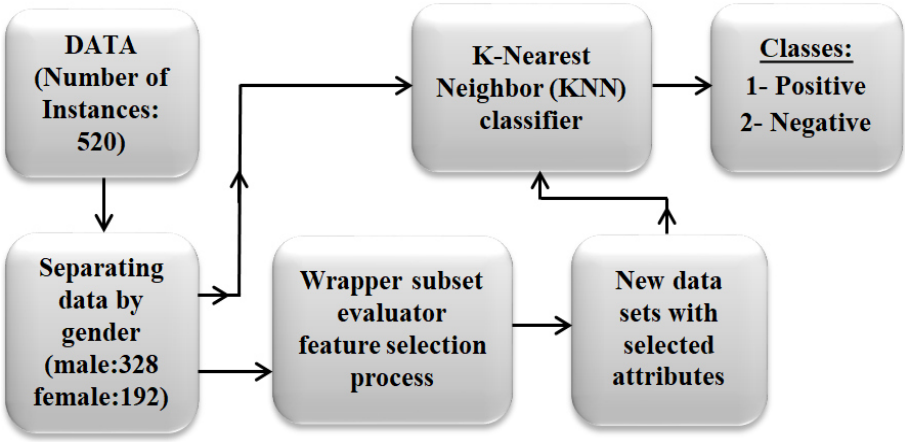


Figure 1. The block diagram of the study

2. MATERIAL AND METHODS

2.1. Used Data

In this study, "Early stage diabetes risk prediction dataset" (Islam et al., 2020) obtained from UCI Machine Learning Repository (Dua, 2019) was used for early detection of diabetes. This data set consists of the answers of 14 different questions asked to a total of 520 people, 328 of which are males and 192 females. Fourteen different questions were asked to participants whose age range was between 20 and over 65 (Islam et al., 2020; Dua, 2019). Questions with 'YES' or 'NO' answers are 'polyuria', 'polydipsia', 'sudden weight loss', 'weakness', 'polyphagia', 'genital thrush', 'visual blurring', 'itching', 'irritability',

'delayed healing', 'partial paresis', 'muscle stiffness', 'alopecia', and 'obesity' (Islam et al., 2020; Dua, 2019). The answers given to the questions within the scope of the study were converted into numeric expressions as "1" and "0".

2.2. Feature Selection Process

In this study, Wrapper subset evaluator was applied to the data according to the best first search algorithm to find effective features for the best statistical performance results. Firstly, the classifier to be used must be determined to start the discriminant operation of the Wrapper feature selection process. Then, it is searched within the feature list according to the determined classifier and the selected search algorithm. Finally, the least number of features that can provide maximum accuracy are determined (Cornforth et al., 2004).

2.3. Classification Process

In this stage of study, KNN classifier was selected to classify determined data. Weka 3.8.3 Software (Witten et al., 1999) was used for both feature selection and classification processes. K-fold cross validation method was preferred to divide data into 10-fold and as a result, training-testing data were created for the classifier. Classes were labelled as “Positive” and “Negative”. In the interpretation of the results, true-positive rate (TPR), false-positive rate (FPR), precision (P), F-measure (F), Matthew correlation coefficient (MCC), receiver operating characteristic area (ROC), precision-recall area (PRC) and classification accuracy rate (ACC) statistical criteria were selected and interpreted.

3. EXPERIMENTAL RESULTS

In this study, the diabetic data set (Islam et al., 2020; Dua, 2019) which consists of the answers of 14 different questions asked to a total of 520 people (328 of which are males and 192 females) was used. The answers of these 14 different questions were presented to KNN classifier separately for each gender. After that, Wrapper subset evaluator was run to select effective features for both genders. Lastly, selected features were classified again with KNN classifier and the results were interpreted. Table 1 contains the statistical classification results made with features before and after feature selection for female group. If Table 1 was analyzed, it could be seen that much better performance was obtained with three effective features ("Polyuria", "delayed healing" and "Alopecia") determined after selection of attributes. Besides, it could be said that a sufficient success was achieved in terms of both time and speed as a result of feature selection process.

Table 1. Statistical values obtained as a result of classification with KNN of all features and the chosen for female group (TNI: Total number of instances, CCI: Correctly classified instances)

Used Data	TNI	CCI	TPR	FPR	P	F	MCC	AUC	PRC	ACC (%)
All features	192	185	0.96	0.28	0.96	0.96	0.77	0.97	0.98	96.35
The chosen features ("Polyuria", "delayed healing" and "Alopecia" features)	192	186	0.96	0.28	0.97	0.96	0.81	0.95	0.97	96.87

The ROC curve obtained as a result of the classification made with 3 attributes for female group, which was determined to be more effective after the feature selection process, was shown in Figure 2. As is known, one of the most important evaluation criteria to check the performance of any classification model is the ROC curve.

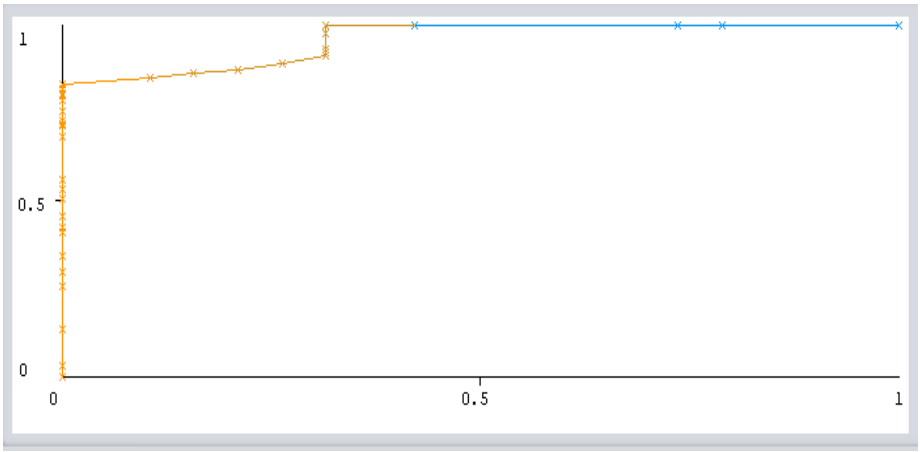


Figure 2. The ROC curve obtained as a result of the classification made with the selected features for female group

The area under this curve (AUC) is one of the most common metrics used to evaluate the performance of machine learning algorithms, especially when working with data with incompatible values. Classification success is directly proportional to the closeness of AUC value to 1. As seen in Table 1 and Figure 2, although the ACC increased, the AUC value decreased. However, since this decrease was 0.02, it did not affect the graph result much. While the results for the female group were like this, Table 2 contains the statistical classification results made with features before and after feature selection for male group.

Table 2. Statistical values obtained as a result of classification with KNN of all features and the chosen for male group (TNI: Total number of instances, CCI: Correctly classified instances)

Used Data	TNI	CCI	TPR	FPR	P	F	MCC	AUC	PRC	ACC (%)
All features	328	321	0.97	0.02	0.97	0.97	0.95	0.98	0.98	97.86
The chosen features (Other features except "weakness" feature)	328	321	0.97	0.02	0.97	0.97	0.95	0.99	0.99	97.86

When Table 2 was examined, it was seen that the classification results obtained before and after feature selection were almost the same for this group. Besides, as a result of the feature selection process, it was understood that only "weakness" feature was ineffective for the male group. Finally, the ROC curve obtained as a result of the classification performed with selected features for male group was shown in Figure 3. As seen in Table 2 and Figure 3, the AUC value increased by 0.01

although the ACC did not change. As a result of the classification, the value of 0.99 AUC was reached and the data were separated almost perfectly.

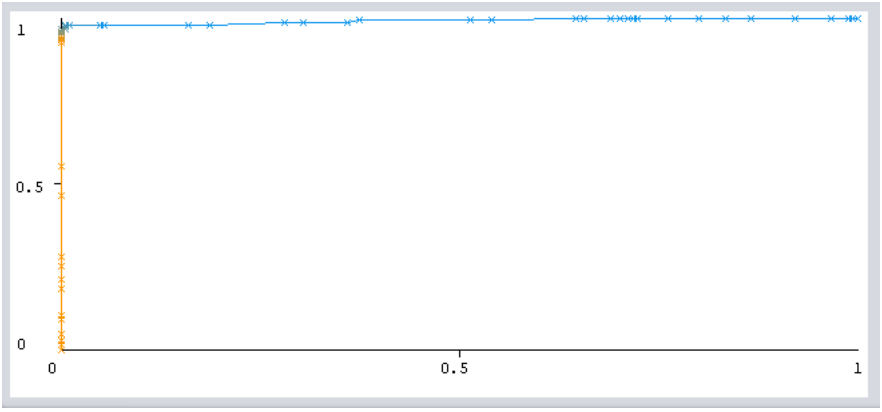


Figure 3. The ROC curve obtained as a result of the classification made with the selected features for male group

4. DISCUSSION AND CONCLUSION

The aim of the study was to determine which features by gender are effective in the early diagnosis of diabetes. K-Nearest Neighbor as a classifier and Wrapper subset evaluator according to the best first search algorithm as a feature selection method were applied to the data set to perform this operation. Before the feature selection process, the accuracy rates of 97.86% and 96.35% were obtained for male and female, respectively. As a result of the procedures applied later, a classification rate of 96.87% was achieved for female with only the features of "Polyuria", "delayed healing" and "Alopecia". It was determined that the "weakness" feature was not effective on the result

of the classification, and the same rate (97.86%) was obtained this time with 13 attributes for male subjects.

When the results are examined, it is understood that the feature selection process makes a slight improvement on the performance values obtained. This situation should not mislead researchers. Because even if the same result is obtained for many systems, it is very important to reach this result with the least feature. In this way, the system that performs the operations will be able to perform better in a shorter time without getting tired. Researchers can apply different feature selection methods simultaneously in this field and compare their performance.

REFERENCES

Cornforth, D. J., Jelinek, H. F., Teich, M. C., & Lowen, S. B. (2004). *Wrapper subset evaluation facilitates the automated detection of diabetes from heart rate variability measures.* Paper presented at the Proceedings of the international conference on computational intelligence for modelling control and automation (CIMCA'2004), University of Canberra, Australia.

Dua, D. a. G., C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. from Irvine, CA: University of California, School of Information and Computer Science.

Hossain, M. A., Ferdousi, R., & Alhamid, M. F. (2020a). Knowledge-driven machine learning based framework for early-stage disease risk prediction in edge environment. *Journal of Parallel and Distributed Computing, 146*, 25-34.

Hossain, M. A., Ferdousi, R., Hossain, S. A., Alhamid, M. F., & El Saddik, A. (2020b). A Novel Framework for Recommending Data Mining Algorithm in Dynamic IoT Environment. *IEEE Access, 8*, 157333-157345.

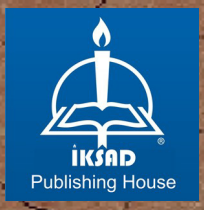
Idf (The International Diabetes Federation) (2020). What is diabetes., https://www.idf.org/aboutdiabetes/what-is-diabetes.html (Accessed date: 30 September 2020).

Islam, M. F., Ferdousi, R., Rahman, S., & Bushra, H. Y. (2020). Likelihood prediction of diabetes at early stage using data mining techniques *Computer Vision and Machine Intelligence in Medical Image Analysis* (pp. 113-125): Springer.

Ogedengbe, M. T., & Egbunu, C. O. (2020). CSE-DT Features Selection Technique for Diabetes Classification. *Applications of Modelling and Simulation, 4*, 101-109.

Turkdiab (Türkiye Diyabet Vakfı) (2020). Diyabet nedir?, https://www.turkdiab.org/bilgiler.asp?lang=TR&id=67 (Accessed date: 30 September 2020).

Witten, I. H., Frank, E., Trigg, L. E., Hall, M. A., Holmes, G., & Cunningham, S. J. (1999). Weka: Practical machine learning tools and techniques with Java implementations.



ISBN: 978-625-7687-10-2