EJFE

EUROPEAN JOURNAL OF FORMAL SCIENCES AND ENGINEERING

July - December 2022

Volume 5, Issue 2

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)

REVISTIA

PUBLISHING AND RESEARCH

EUROPEAN JOURNAL OF FORMAL SCIENCES AND ENGINEERING

July - December 2022

Volume 5, Issue 2

Typeset by Revistia

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)

Address: 11, Portland Road, London, SE25 4UF, United Kingdom

Tel: +44 2080680407

Web: https://ejfe.revistia.org

Email: office@revistia.org

Indexed in Elsevier's Mendeley, WorldCat, RePEc & Ideas, Google Scholar, Microsoft Academics, Index Copernicus, Crossref

International Editorial and Scientific Advisory Board

TABLE OF CONTENTS

LAND USE/LAND COVER CHANGES IN A DISTURBED RIVER WATERSHED KENYA 1

ZACHARY GICHURU MAINURI
JOHN M. MIRONGA
SAMUEL M.MWONGA

QUALITATIVE RESEARCH AS A TOOL TO CARRY OUT ARCHITECTURAL AND INDUSTRIAL DESIGN PROJECTS: A VISION FROM THE ACADEMIC PERSPECTIVE 13

LUCIA CONSTANZA IBARRA CRUZ
LEYDA MILENA ZAMORA SARMIENTO

SPECIFICS OF ALGORITHMIZATION IN DATA CULTURE 24

TADEUSZ MICZKA

THE COMPARISON OF SOME METHODS IN ANALYSIS OF LINEAR REGRESSION USING *R* SOFTWARE 38

ILIR PALLA

USING UNMANNED AERIAL VEHICLES (UAVS) TO ANALYZE THE URBAN ENVIRONMENT 50

EFTHIMIOS BAKOGIANNIS
CHARALAMPOS KYRIAKIDIS
VASILEIOS ZAFEIRIS

THE DESIGN OF A TOMATO POWDER PROCESSING PLANT WITH A CAPACITY OF 75,000 TONNES PER ANNUM IN OYO STATE, NIGERIA, WEST AFRICA 62

OSARETIN N.I. EBUEHI

DETECTION OF ANOMALIES IN THE COMPUTER NETWORK BEHAVIOUR 79

DANIJELA PROTIĆ
MIOMIR STANKOVIĆ

DESIGN AND APPLICATION OF OPTIMALLY-TUNED VARIABLE PARAMETERS PID CONTROLLER FOR NONLINEAR ENGINEERING SYSTEMS 90

MAGDY A. S. ABOELELA

A STUDY OF GAS DIFFUSION CHARACTERISTICS ON NANO-STRUCTURED CERAMIC MEMBRANES 98

PRISCILLA OGUNLUDE
OFASA ABUNUMAH

EDWARD GOBINA

COMPARING DIFFERENT EQUIPMENT AND APPLICATIONS IN PAVEMENT DATA COLLECTION AS PART OF ROAD MANAGEMENT SYSTEM 103

ARJOL LULE
SHKELQIM DAJA

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Land Use/Land Cover Changes in a Disturbed River Watershed Kenya

Zachary Gichuru Mainuri

Egerton University, Crops, Horticulture and Soils Department

John M. Mironga

Egerton University, Geography Department

Samuel M.Mwonga

Egerton University, Crops, Horticulture and Soils Department

Abstract

Drivers of land use change were captured by the use of DPSIR model where Drivers (D) represented human needs, Pressures (P), human activities, State (S), the ecosystem, Impact (I) services from the ecosystem and Response (R), the decisions taken by land users. Land sat MSS and Land sat ETM+ (path 185, row 31) were used in this study. The Land sat ETM+ image (June 1987, May, 2000 and July, 2014) was downloaded from USGS Earth Resources Observation Systems data website. Remote sensing image processing was performed by using ERDAS Imagine 9.1. Two land use/land cover (LULC) classes were established as forest and shrub land. Severe land cover changes was found to have occurred from 1987-2000, where shrub land reduced by -19%, and forestry reduced by -72%. In 2000 – 2014 shrub land reduced by-45%, and forestry reduced by -64%. Forestry and shrub land were found to be consistently reducing.

Keywords: watershed. Land use\land cover change, Landsat imagery, Geographic Information System

Introduction

Land use/Land cover change (LULCC) is continuously changing the Middle part of the River Njoro watershed, thereby threatening sustainability and livelihood systems of the people. Biodiversity is facing widespread competition with humanity as human population increases, resulting in increasing conflict between economic development and the need for biodiversity conservation. These environmental problems are often related to LULC changes. LULCC and human/natural modifications have largely resulted in deforestation, biodiversity loss, global warming and increase of natural

disasters like flooding (Fan *et al.*, 2007, Dwivedi, *et al*, 2005). LULCC plays a major role in the study of global Land use/land cover change. Coexistence between local land uses and conditions for environmental, social, and economic sustainability has not been adequately addressed. Land use/land cover change is dynamic. It is mainly driven by natural phenomena and anthropogenic activities. Seto, *et al.*, 2002, has reported that pressure from growing population and increasing socio-economic necessities results in unplanned and uncontrolled changes in LULC. Therefore, available data on LULC changes can provide critical input to decision-making of environmental management and planning the future (Fan, *et al.*, 2010, Prenzel, 2004).

Drivers, pressure, State, Impact and Response (DPSIR) model as a decision making tool, has been applied in numerous research efforts; including Water Resources Management at various scales. It has also been used in a series of international and multidisciplinary research projects as the main analysis tool (Tscherning *et al.*, 2012). The demand for agricultural land, energy, water, food, transport and housing can serve as examples of driving forces (Giupponi, 2002; Kristensen, 2004; Wood and van Halsema, 2008). Pressures consist of the driving forces' consequences on the environment such as the exploitation of resources (land, water, minerals, and fuels), pollution and the production of waste or noise (Wood and van Halsema, 2008). As a result of pressures, the 'state' of the environment is affected; that is, the quality of the various natural resources (air, water, and soil) in relation to the functions that these resources fulfill. The 'state of the environment' is thus the combination of the physical, chemical and biological conditions. The support of human and non-human life as well as the depletion of resources can serve as pertinent examples (Kristensen, 2004). Changes in the state may have an impact on human health, ecosystems, biodiversity, amenity value and financial value. Impact may be expressed in terms of the level of environmental harm and finally, the responses demonstrate the social efforts to solve the problems identified by the assessed impacts, e.g. policy measures, and planning actions (EEA, 1999; Giupponi, 2002; Kristensen, 2004; Wood and van Halsema, 2008).

Remote sensing and Geographical Information Systems (GIS) as a resource management tool is powerful to derive accurate and timely information on the spatial distribution of land use/land cover changes over large areas (Guerschman, *et al*., 2003, Rogana and Chen, 2004, Zsuzsanna, *et al*., 2005). GIS provides a flexible environment for collecting, storing, displaying and analyzing digital data necessary for change detection (Yomralıoğlu, *et al*., 2000, Demers, 2005, Wu *et al.,* 2006). The aim of land cover change detection process is to recognize LULCC on digital images that change features of interest between two or more dates (Muttitanon and Tiıpathi, 2005).This change in land use has exposed the land to various pressures resulting from poor management, low cost technologies for soil fertility management, continued use of inappropriate technologies and intensive cultivation. Therefore, there is a need to understand how land use changes had affected the environmental sustainability of the area.

1. Study Area

The area of study covers about 8,170 ha and lies between latitudes 0º 15´ S and 0º 25´ S and longitudes of 35º 50´ E and 36º 00´ E (Figure 1). The whole watershed has a population of about three hundred thousand (300,000) people with more than three thousand (3000) individual farm holding units (Baldyga, *et al.,* 2003). However, according to Kenya National Bureau of Statistics, Njoro Sub County registered a population of 23,551 people having grown by 3% from a population of 22, 845 people in 1999 (KNBS, 2009). Based on the same growth rate, the watershed population may have also grown to 309, 000 people with may be 3100 households. Due to the heavy settlement in the middle part of the watershed, it is estimated to be home to about 2000 farm holding units in an area of more than 8,000 ha with slopes ranging from < 2 to > 18 % and soils that are predominantly volcanic clay loam except near the lake where silt clay is found (Mainuri and Owino, 2013).

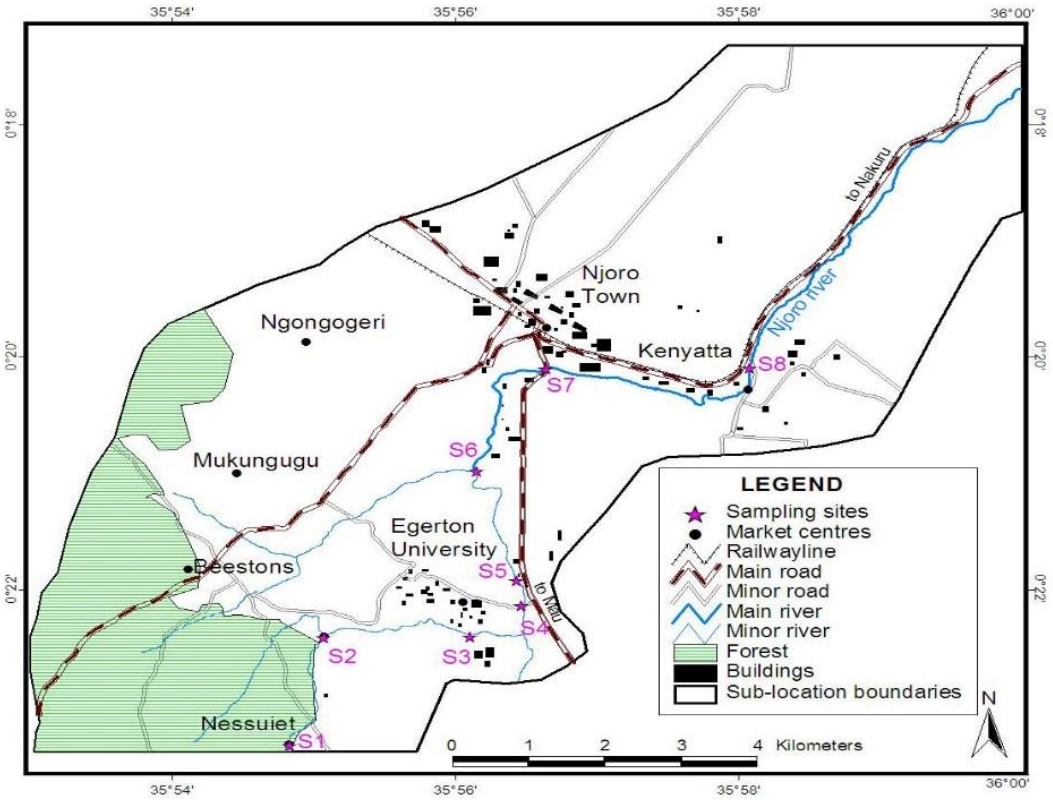

Figure 1: Middle River Njoro Watershed (Source: Mainuri and Owino, 2014)

2. Methods

A baseline survey at household-level encompassing socio-economic changes and impacts of land use activities in the middle part of the River Njoro Watershed was

established. Additionally, information on factors influencing land use decisions, productivity factors and change in economic activities were sought through use of a questionnaire. The middle part of the River Njoro Watershed household survey was to target an area of approximately 8000ha.The Landsat scenes were selected (1987, 2000 and 2014) for this study. These dates captured the major excision and settlement changes that have taken place in the watershed. Efforts were made to acquire imagery that corresponds with major land use/land cover changes within this period.

The study utilized 200 questionnaires which were administered to homesteads that were initially identified at random on both sides of the river. The questionnaires were subjected to scrutiny for completeness and consistency in question answering and the way they addressed the various issues intended to be captured. The questionnaires were sorted out and entered into the SPSS (version 20) work sheet. With the descriptive and categorical nature of most of the questions, simple descriptive analysis was done using SPSS and inferential statistics performed based on the results.

2.1 Image classification

Land sat MSS and Land sat ETM+ (path 185, row 31) were used in this study. The Land sat ETM+ images (June 1987, May, 2000 and July, 2014) were downloaded from USGS Earth Resources Observation Systems data. The dates of both images were chosen to be as closely as possible in the same vegetation season. All visible and infrared bands were included in the analysis. Remote sensing image processing was performed using ERDAS Imagine 9.1.Five LULC classes were established as commercial farms, forest, settlement, subsistence farms, and shrub land. Three dated Land sat images (1987, 2000, and 2014) were compared using supervised classification technique. In the supervised classification technique, three images with different dates were independently classified. A Supervised classification method was carried out using training areas. Maximum Likelihood Algorithm was employed to detect the land cover types in ERDAS Imagine 9.1.

3. Results

3.1 Nature and status of Land Use/ Cover during acquisition time

The study established that most of the land was under cultivation when the current owners acquired it, as the majority (31.7%) of the responses portrays it. This was closely followed by grass cover which formed 26.6% of the total responses, with 19% reporting that the land area was under indigenous trees when they initially moved in, while a 15.4% response exhibited presence of exotic trees. However, only 7.3% of the total responses reported the presence of soil and water conservation structures on the land during initial settlement period (Table 1).

Table 1: Nature/ state and extent of Land cover during acquisition by current owners

Land Use/ Cover	Responses on Land use		Percent of Cases (interviewed)
	N (Number of Respondents Interviewed)	Percent (observed Land use change)	
Presence of soil and water conservation structures	24	7.3%	12.9%
Under cropping	105	31.7%	56.5%
Under grass cover	88	26.6%	47.3%
Under indigenous trees	63	19.0%	33.9%
Under exotic trees	51	15.4%	27.4%
Total	331	100.0%	178.0%

3.2 Land use activities and factors influencing decisions

An interview was carried out on some key informants concerning the land use activities. They reported that the main environmental impacts were a general increase in agricultural activities on riparian zones. The main economic activity creating impacts to the ecosystem that was reported by these people was usually farming which resulted in the reduction of natural vegetation. However, the state of the ecosystem has remained a bit stable due to agro forestry that has contributed to planted forest which is thriving in some parts of the ecosystem. The response from those interviewed indicated that 88 per cent of those interviewed were farmers, 3 percent were business persons, 3 percent masons, and 3 percent crafts men and 3 percent teachers. Respondents' level of education refers to the actual number of years spent in school. The interview showed that 50 percent of the respondents had obtained up to primary education, while 20% percent have not obtained any formal education. A lower proportion (33%) had obtained secondary and post secondary level of education. Generally, 70 percent of the respondents had primary level education and below. The finding indicates that most of the respondents in the middle part of the river Njoro watershed had low formal education and this may have affected the way in which they responded to new information on resource conservation and how they also received innovative ideas.

The respondents were interviewed on the changes in natural vegetation. A huge portion of the respondents (93%) have observed massive land use changes taking place with 7% not feeling that there has been any noticeable change in land use. This possibly could be that they have recently settled in the area and since they settled

there has been no change. The pressures exerted by the society through deforestation may have led to unintentional or intentional changes in the state of the ecosystem. As a result of no proper land ownership, most people are shy to invest in long term development activities and majorities are sluggish or unable to take any resource conservation measures. Assessment of driving forces behind land use change was done to capture past patterns and also be able to forecast future patterns. Driving forces on land use included most of the factors that influenced human activity that exert pressure on the ecosystem, including population increase, poverty, land tenure and markets. Also other underlying factors that drive actions like food preference demand for specific products, financial incentives and environmental state indicators such as soil quality, terrain and moisture availability played a great role in affecting the natural vegetation as shown in Table 2.

Table 2: Change detection

Class Type	1987 Area in Hectares	2000 Area in Hectares	2014 Area in Hectares	Percent change in area (2000-1987)	Percent change in area (2014-2000)
Forest	1460.898	405.351	145.712	(-1055.55)-72%.	(-259.64) -64%
Shrub land	849.281	687.820	373.150	(-161.46) -19%,	(-341.67) -45%

Increasing land use/cover changes were observed in the middle part of the river Njoro watershed ecosystem over the last twenty seven (27) years. These changes resulted from a number of factors, but mainly related to habitat loss due to various human activities. Information about changing patterns of land use/cover through time and the factors influencing such changes have been captured in the change detection maps shown in Figure 2, 3 and 4 below.

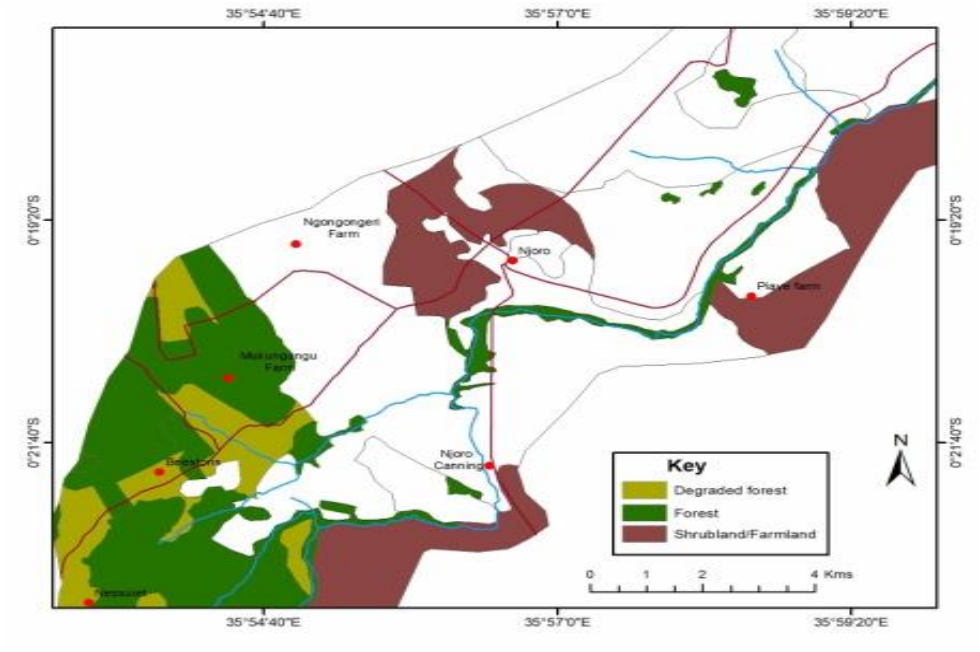

Figure 2: Forests and Shrub Lands cover in 1987

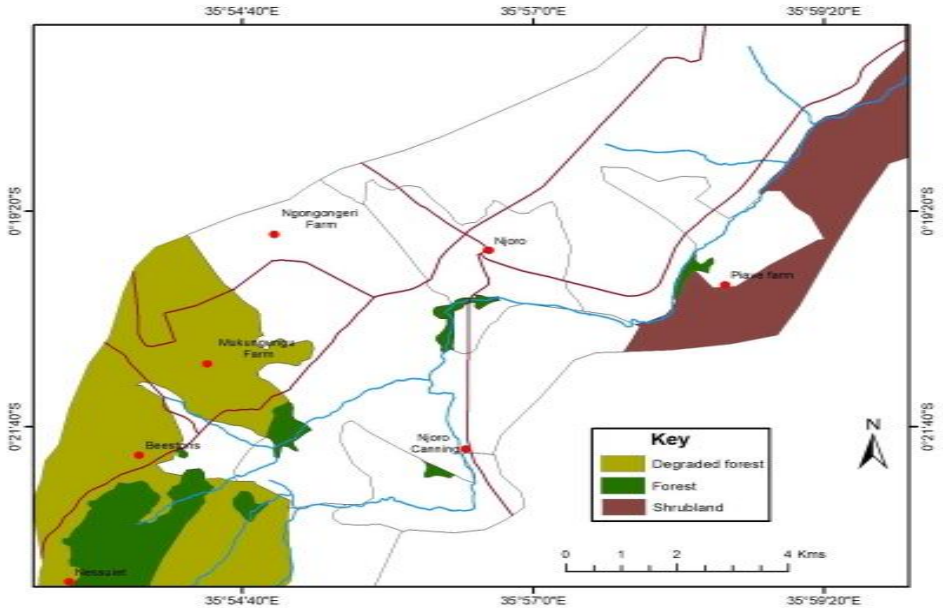

Figure 3: Reduction of Forests and Shrub Lands in the year 2000

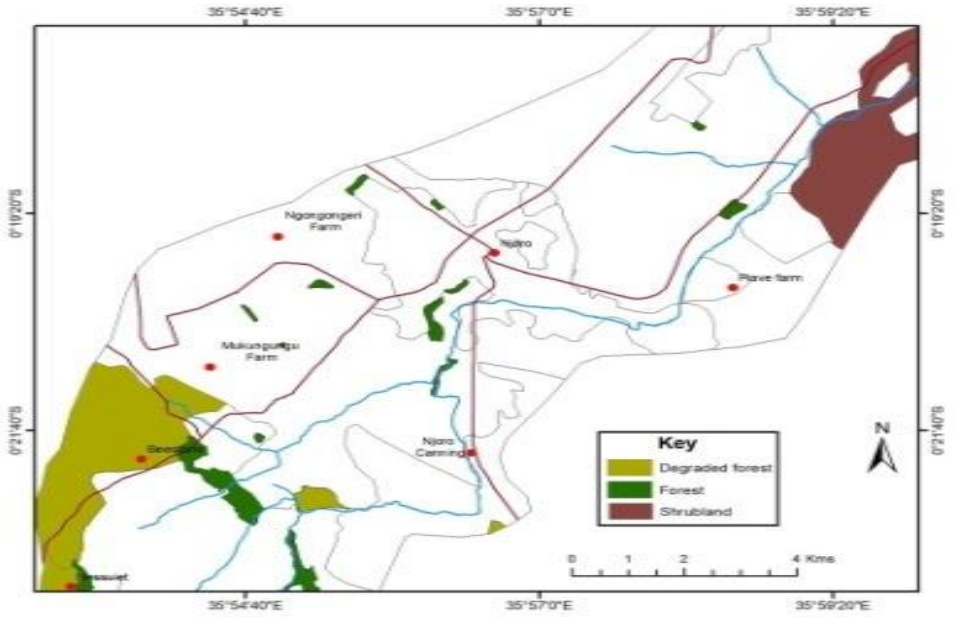

Figure 4: Reduction of Forests and Shrub Lands in the year 20014

3.3 Natural Vegetation Cover

From the study, it is evident that natural vegetation which was indicated by forest and shrub land (Table 3) has reduced over the period the respondents have resided in the area. The results from image processing and analysis for the years 1987, 2000 and 2014 portray a general reduction in both forests and shrub lands within the study area. We can therefore say that deforestation has been witnessed in the study area for the last two decades due to land use patterns.

Table 3 Respondents' view on Natural Vegetation

Year	Forest Area(ha)	Shrub land Area (ha)	Natural Vegetation Change	Frequency (Number interviewed)	Percent of respondents interviewed
1987	1460.898	849.281	Decrease	32	20.6
2000	405.351	687.820	Decrease	123	79.4
2014	145.712	373.150	Total	155	100

3.4 Reasons for Reduction in Natural Vegetation

Several activities and their impact on reducing natural vegetation were identified during the study. From Table 4, cultivation stood out to be the major driving force that led to the reduction in natural vegetation cover in these areas as reported from the respondents. This constituted 33% of the total responses. Other activities included charcoal burning (11.2%), infrastructural development (10.4%) and grazing (9.9% and commercial timber production (4.7%). Collectively, these have led to deforestation in the area under study.

Table 4: Responses for change in natural vegetation

Reasons for change	Responses on Land cover change		Percent of Cases
	N	Percent(observed)	
Commercial timber production	18	4.7%	11.1%
Cultivation	127	33.0%	78.4%
Infrastructural development	40	10.4%	24.7%
Charcoal burning/ firewood	43	11.2%	26.5%
Grazing	38	9.9%	23.5%

4. Discussions and Conclusions

In order to determine the current land use and factors that influence land use decisions in the middle part of the River Njoro watershed the study sought to established the kind of land use before the occupation of the current inhabitants. It was found that 32 % of the land was under cultivation when the current owners acquired it as confirmed by the interviewee. 27% of the respondents indicated that they occupied land that was under grass cover with 19% reporting that the land area was under indigenous trees when they initially moved in, while a 15% response exhibited presence of exotic trees. Driving forces on land use included most of the factors that influenced human activity that exert pressure on the ecosystem, including population increase, poverty, land tenure and markets.

Alongside determining the land use and factors influencing land use decisions, the study also looked at land use/land cover changes that were as a result of land use decisions that the people made. It was noted that there were increasing land use/cover changes observed in the middle part of the river Njoro watershed over the period of study. These changes resulted from a number of factors that included increase in population, change in lifestyle and the need to provide food for the increasing numbers of people.. Several activities and their impact on reducing natural vegetation were identified during the study with cultivation being the major driving forces that has led to the reduction in natural vegetation cover in these areas

constituting 33% of the total responses. Other activities that contributed to land use/land cover change included charcoal burning, infrastructural development and grazing and commercial timber production. Collectively, these have led to deforestation in the area under study.

Land degradation by overgrazing and intensive agriculture on marginal lands is a major driver of land cover loss in the middle part of the river Njoro watershed. In this rapidly industrializing area with dense populations, demand for land for industry and residential use is driving the transformation of some of the most productive agricultural land out of production in the watershed. Policy efforts to avoid this loss of production are there but, their effectiveness in the face of economic demand is often limited. The effectiveness of these efforts and other national efforts to reduce the negative impacts of LULCC remain to be seen. The need for greater efforts and new methods to monitor and mediate the negative consequences of LULCC remains acute and we have to sustain current and future human populations under desirable conditions. This can be realized by putting in place policies like reafforesttion of natural forests, mandatory planting of trees in homestead, controlled tree harvesting and restricting encroachment into the forests.

Conclusion

The factors driving land use decisions in the middle part of the River Njoro watershed include demographic and economic developments in the watershed community, and the corresponding changes in lifestyles, overall levels of consumption and production patterns. These drivers have exerted pressure to the ecosystem in form of waste disposal, over cultivation, overgrazing and deforestation. These pressures have caused negative changes to the watershed which have caused heavy impacts mainly through removal of natural vegetation. The removal of natural vegetation (LULCC) in the middle part of the River Njoro watershed has resulted in the decrease of the forest area by 1314 ha and shrub land by 475 ha in the last 27 years. The integration of remote sensing and GIS was found to be effective in monitoring and analyzing land cover patterns and also in evaluating impacts of land use change for future land development projects by the residents of study areas.

The residents are therefore recommended to develop responses to rehabilitate the degraded environment through re-afforestation, soil and water conservation and reduction of land use/land cover change (LULCC) in order to mitigate the negative outcomes of the ecosystem changes.

References

[1] Baldyga, T.J. (2003). Assessing Land cover Change Impacts in Kenya's River Njoro Watershed using remote sensing and Hydrological Modeling. MSc Thesis (unpublished) submitted to the Department of Renewable Resources at the University of Wyoming. Laramie, U.S.A.

[2] Demers, M. N. (2005). Fundamentals of Geographic Information Systems, John Wiley & Sons, Inc., Newyork, USA, 2005.
[3] Dwivedi, R.S.; K.Sreenivas; K.V.Ramana, (2005). Land-use/land-cover change analysis in part of Ethiopia using Landsat Thematic Mapper data. *International Journal of Remote Sensing* 2005, 26 (7), 1285-1287.
[4] European Environment Agency (1999). Environmental Indicators: Typology and Overview. Technical Report No. 25, Copenhagen.
[5] Fan, F.; Weng, Q.; Y. Wang, (2007). Land use land cover change in Guangzhou, China, from 1998 to 2003, based on Landsat TM/ETM+ imagery. *Sensors* 2007, *7*, 1323-1342.
[6] Fan, F., Q.Weng, and Y.Wang, (2010). Land use and land cover change in Guangzhou, china, from 1998 to 2003, based on Landsat TM /ETM+ imagery Sensors 7, 1323-1342
[7] Government of Kenya, Kenya National Bureau of Statistic (KNBS) (2009). Kenya National Population and housing census, 2009.
[8] Guerschman J.P.; J.M.Paruelo,; C.D.Bela,; M.C.Giallorenzi,; F.Pacin, (2003). Land cover classification in the Argentine Pampas using multi-temporal Landsat TM data. *International Journal of Remote Sensing* 2003, *24*, 3381–3402.
[9] Giupponi C. (2002). From the DPSIR reporting framework to a system for a dynamic and integrated decision making process. International MULINO Conference on "European policy and tools for sustainable water management" Island of San Servolo, Venice, Italy, November 21-23.
[10] Kristensen P. (2004). The DPSIR Framework. National Environmental Research Institute, Denmark. European Topic Centre, European Environment Agency.
[11] Mainuri, Z.G., and J.O.Owino, (2013). Effects of land use and management on aggregate stability and hydraulic conductivity of soils within River Njoro Watershed in Kenya. *International Journal of Soil and wate conservation research. Vol.1 No.2 pp. 8o -87.*
[12] Muttitanon W. N.K . Tiıpathi, (2005). Land use/land cover changes in the coastal zone of Ban Don Bay, Thailand using Landsat 5 TM data. *International Journal of Remote Sensing* 2005, *26* (11), 2311- 2323.
[13] Prenzel, B.(2004). Remote sensing-based quantification of land-cover and land-use change for planning.*Progress in Planning* 2004, *61*, 281–299.
[14] Rogan J, D.Chen, (2004). Remote sensing technology for mapping and monitoring land-cover and landuse change. *Progress in Planning* 2004, *61*, 301–325.
[15] Seto, K.C., C.E.Woodcock, C.Song, X.Huang, J.Lu, R.K. Kaufmann, (2002).Monitoring land use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing* 2002, *23*, (10), 1985-2004
[16] Tscherning K., K.Helming, B.Krippner, S.Sieber and Y.Gomez, S.Paloma, (2012).Does research applying the DPSIR framework support decision

making?, *Land Use Policy*, 29(1), 102–110. United Nations (UN), 1996. Indicators of sustainable development: Framework and Methodologies. Report. 428

[17] Wood A. and van Halsema G. (2008), Scoping agriculture–wetland interactions Towards a sustainable multiple-response strategy, FAO Water Reports No 33

[18] Wu, Q.; H. Q.Li,; R.S.Wang,; J.Paulussen,; H.He,; M.Wang,; B.H.Wang,; Z.Wang, (2006). Monitoring and predicting land use change in Beijing using remote sensing and GIS. *Landscape and Urban Planning* 2006, *78*, 322–333.

[19] Yomralıoğlu, T. Coğrafi Bilgi Sistemleri Temel Kavramlar ve Uygulamalar, Seçil Ofset, Istanbul, Turkey, 2000.

[20] Zsuzsanna, D.; J.Bartholy,; R.Pongracz,; Z.Barcza, (2005). Analysis of land-use/land-cover change in the Carpathian region based on remote sensing techniques. *Physics and Chemistry of Earth* 2005, *30*, 109-115. © 2008 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Qualitative Research as a Tool to Carry out Architectural and Industrial Design Projects: a Vision from the Academic Perspective

Lucia Constanza Ibarra Cruz

PhD, Full-time Research Professor,

Autonomous Metropolitan University Xochimilco

Leyda Milena Zamora Sarmiento

MDI. Full-time Research Professor,

Autonomous Metropolitan University Xochimilco

Abstract

In recent years, qualitative research has gained ground as one of the main tools to analyze reality and propose innovation projects.This qualitative methodology has strengthened an innovative pedagogical structure, within the teaching of design disciplines.This is the specific case of our educational activity in the Modular System (UAM X), which facilitates the application of these tools in the design process. Reviewing the data in the analysis in a more qualitative way, promotes a more credible, reliable and valid vision to propose solutions to the problems detected.Thus, achieving a comprehensive understanding of the phenomenon and considering the factors involved in order to generate more relevant solutions.Within architecture, some projects that are addressed, such as the historical and cultural heritage of spaces and cities, use qualitative research techniques to discover and interpret the real vision of the inhabitant/ visitor of these spaces. It is essential to understand how the city is lived from the point of view of the different actors that interact in the space, in order to generate plans and projects that reinforce and enrich social interaction.In industrial design, the students take these tools and apply them to the first phase of their project in search of insights that will allow them to bring forward innovative products with a real impact on society.This paper addresses these two cases, highlighting the importance of these tools being internalized by our students completing their training with a research profile focused on design.

Keywords: qualitative research, knowledge, skills, Modular System, teaching of architecture and industrial design.

Introduction

Qualitative research as a tool for industrial design and architecture projects: a vision from the academic perspective

Qualitative research methods have been widely developed within the social sciences since long ago, but in areas such as architecture and design, their implementation has been recent.

The professional training of architects and designers has been traditionally within the methods focused on creativity, however the new situations proposed by the context make these disciplines approach methods related to social sciences that allow them to have a qualitative vision of context and the user or inhabitant.

In the Xochimilco UAM, this condition is very frequent because the proposed pedagogical model, the Modular System, implies starting from reality itself; therefore, the involvement of the qualitative vision is not foreign, however, it has not been given much emphasis due to the paradigms of design and architecture in its traditional teaching and the free professorship that is given in the institution.

Because the Modular System is the basic teaching model at the Universidad Autónoma Metropolitana Unidad Xochimilco (from now the university will refer as UAM X), it is important we give context to this model.

UAM Xochimilco and the Modular System

In 1974, the UAM opens its doors. One of its main purposes is the training of professionals in Mexico City. The Xochimilco unit, unlike the other four units that make up the UAM; undertakes its activities with the proposal of an innovative pedagogical project, with the aim of training professionals with a vision of reality and a greater social commitment to their country and its context.

The Modular System emerged, then as a pedagogical alternative that faced the student and his commitment to social problems, seeking greater sensitivity and making him the architect of his own training.

Starting from the principles of structuring and construction of Jean Piaget's knowledge, cognitive development based on the dialectic between Vygotsky's individual and society and combining it with the concepts of interdisciplinary and operational groups, the Modular System challenges the student to he himself, from his own experience, becomes a transforming subject of reality and thus his knowledge takes on a true meaning.

The Modular System, as stated in the conceptual bases, allows the formation of agents of social transformation, with “capacity for critical thinking, capacity for creative action and capacity for global understanding of the facts” (UAM, 1974-78).

As a complement to the postulates of the Modular System, the activities that guide the daily work in the institution are the three substantive functions of our University: teaching, research and service:

1. Research as knowledge production based on specific social objectives.

2. Teaching as communication and practical confrontation of knowledge.

3. Service, as the social application of such knowledge" (UAM, 1974-1978).

In general, these functions allow the university's daily work "to contribute to the attention of problems, the generation of research, development, innovation, production, commercialization and sustainability, among others" (UAM, University Legislation, p. 243).

The basic postulate for the construction of knowledge, according to the Xochimilco Document (UAM, 2012) takes as a tool the application of scientific research, this was due to the fact that the Divisions within which the academic activities began were those of Social Sciences and Humanities, and Biological and Health Sciences. In these areas, from the processes of systematic observation, measurement, experimentation and hypothesis formulation, the objective was to minimize the subjectivity of the knowledge construction process.

However, the processes in the Division of Science and Arts for the Design of this Unit, hereinafter CyAD, required a more linked method with subjective aspects such as human creation, with a different approach to address the problems that society poses.

Beyond the traditional pedagogical training by disciplines, the Modular System proposes to link the teaching-learning process to a relevant social problem to be approached in an interdisciplinary way. Currently, we are immersed in a process of resignification and updating of the Modular System, it is recognized that although it has some critical points, it has been more successful in relation to its link with reality itself.

This approach is applied in the three UAM-X Divisions, which 45 years after its foundation has 14,007 students in 18 degrees and 972 students in 29 Postgraduate plans.

Vocational training consists of 12 modules or teaching-learning units, proposed from the objects of transformation, which are socially relevant problems to solve. The educational character is concretized in the development of three fundamental concepts in the student's formation: the capacity for critical thinking, creative action and the global understanding of the facts. All focused on social transformation.

An important element is the rethinking of the objects of transformation and disciplinary and interdisciplinary fields that come to resignify and transform professional practice. The Bachelor of Architecture and Industrial Design are not foreign to this phenomenon.

Today, at UAMX, the Architecture and Industrial Design degree require a different dynamic derived from the new challenges of education and context. As Bauman (2007, p. 33) tells us: "the world, as it is lived today, seems more an artifact designed to forget than a place for learning." The challenge is to rethink the relevance and significance that academic structures have for the world as we live it.

Qualitative research in the Modular System

For Taylor and Bogdan (1987) there are two theoretical positions to conduct research: the positivist who seeks "the facts or causes of social phenomena regardless of the subjective states of individuals" (p.15) and the phenomenological one that "wants to understand social phenomena from the actor's own perspective. Examine the way in which the world is experienced" (p.15) Given the fundamentals of the Modular System, the second position is the one that works most for the purpose of starting from a context and a reality of its own and thus identifying the elements and situations that they become important to that person or community that is being investigated.

For Denzin and Lincoln (2012, p. 3) qualitative research "is a situated activity that locates the observer in the world. It consists of a set of interpretive, material practices that make the world visible. These practices transform the world. They turn the world into a series of representations, including field notes, interviews, conversations, photographs, recordings, and memos to the self." The Modular System places us in concrete reality and asks us to transform reality from the deep knowledge of your problems and needs. Identifying, how, from our specific position in the world you can contribute to a reality as complex as the one we live.

According to Creswell (2013), qualitative research is based on assumptions and uses theoretical and interpretative frameworks to be able to approach studying specific problems or situations. The possible approaches, from the qualitative position (narrative research, phenomenology, grounded theory, ethnography and case study) provide a particular perspective throughout the research process, which begins with the approach itself, through the collection of data, the analysis process, until reaching the final report.

An element that stands out in this research process is "the final written report or the presentation [which] includes the voices of the participants, the researcher's reflexivity, a complex description and interpretation of the problem" (Creswell, 2013, p. 44). By linking it with the purposes of the Modular System, this type of "reports" are the ideal products for the construction of knowledge based on their own experiences.

Starting from reality to return to reality itself with relevant proposals, is the path that research proposes to travel. This means, as Denzin and Lincoln (2012) say that qualitative researchers study things in their natural environments, trying to make

sense or interpret phenomena in terms of the meanings that people give them, and making sense from their own interpretative frameworks.

It is necessary that from interdisciplinarity, concept raised early in the teaching model of the Modular System, and with the knowledge and the possibility of implementing "a wide range of interconnected interpretive practices" (Denzin and Lincoln, 2012, p.4) that facilitate the understanding of the context, making possible a different vision from each particular practice, enriching the interpretation of events.

Qualitative research is then a perfect complement within the teaching model of the Modular System. And taking place within CyAD, it is the opportunity to make our architecture and industrial design students more sensitive to the reality around us. Qualitative research offers us a variety of possibilities to walk in the framework of research. A method that has strengthened the path of the different teaching-learning processes in the different disciplines that are related to design, specifically in CyAD.

The great challenge and value in the transformation objects proposed for the teaching of design and architecture in our unit, is to work with real problems or situations, which demands actions different from those proposed by traditional design methods.

The training of architects and designers in CyAD

In response to the multiple visions that occur within our academy, the freedom of professorship allows each teacher to choose the design method they prefer to share with their students and with which they intend to approach to solve the problems of reality.

This includes traditional design methods where the approach is given by the same teacher and the student only has to "be inspired" to generate a large number of possibilities to solve the specific problem bounded by the teacher. Two situations are obvious: the first has to do with the process that the teacher has followed to identify the same problems, needs and problems to solve, the second has to do with giving the student the questions and not allowing him to distrust of that approach.

In this situation, the research phase of the design process is limited to inquiring about production technologies, and existing references, preventing seeing beyond the approach previously established by the teacher.

Other design methods contemplate research as a starting point to be able to frame the problem statement. It is here that qualitative research becomes the ideal tool to place the teaching-learning process in reality itself.

The direct approach of the student with reality leads him to question his value judgments and prejudices around the topics and contexts addressed, to try to understand in an empathic way the specific situations. Once the realities are lived and experienced, the student is more sensitive to ask questions and explore possible solutions.

Approaching reality allows the student to identify more needs than the same user or inhabitant is aware of. These needs are called implicit, latent or non-articulated needs, which are the result of detailed observation and in-depth analysis.

When the user articulates the need, the project remains to please the user, when the designer identifies the unarticulated need, the project can scale the limits of the proposals and reach true objects or systems that solve that need.

It is in this sense that qualitative research methods provide those perspectives that allow identifying true needs, or latent needs and allow them to propose innovative objects and / or spaces.

Qualitative research in the teaching of Architecture in CyAD

In the training of architecture students at UAM X, integration is considered as a fundamental element for their analysis and study. In this sense qualitative research, ethnography, psychology and many other disciplines have a great acceptance among this field of teaching; what has allowed us to find in these qualitative research methods a valuable instrument to approach and know the reality of society and its habitable spaces.

In many cases, quantitative and statistical approaches are not entirely sufficient for the analysis and study of human behavior, since they exclude the subjectivity of the interaction between people and the researcher's vision to obtain situated elements and objectives.

Within the qualitative research applied to living spaces, the researcher is aware of the influence he can exert when collecting information. Being aware of this makes him sensitive to his influence on his object of study and will generate a more complex interpretation of the phenomenon.

To work with undergraduate and graduate students in architecture, it is necessary to analyze and verify directly with reality. This is being built in relation to the different problems present in the social, spatial, environmental, architectural and territorial fields. Addressing reality from this systemic perspective allows us to build a complex interpretation of everyday life and contribute to the generation of knowledge in this area.

With qualitative research methods, architects are permanently trained in the interpretation of reality through interaction. Life stories, participant observation and case studies, allow to gather the necessary information for this.

In this sense, qualitative research proposes methods and procedures that enrich the architect's vision. That is, a new paradigm, a new model that supports an operation to obtain certain results. It can be said that it is a search for knowledge that is very committed to reality and to the people who inhabit it.

The search for knowledge in this way, provides a new point of support for the search for knowledge in living spaces. This qualitative search promotes elements for the development of architectural proposals, which is constantly evolving and in the search for social innovation.

Qualitative research in the teaching of industrial design in CyAD

The degree in industrial design is no stranger to the situation of the free professorship described in Architecture. Each teacher decides the method with which their students will develop their design projects. These methods range from the traditional ones that start from the approach of the problem, for example the one proposed by Munari (2004), to methods where students are asked to approach reality to understand it as the Design Thinking proposed by Brown (2008), but at the time of implementing it in the classroom it is reduced to creativity techniques.

Since the industrial design began to be formally taught, the human aspect has been fundamental, however, the approach was based on intuitions about what users did users prefer. This generated processes focused on the creative moment and unlinked from the deep knowledge of the user, in this way, the design process begins by posing a problem and goes on to generate solutions, resulting in an almost pre-established response.

A few years ago, when the user went from being a group of specific people with established demographic characteristics, to an individual with qualities, designers began to identify opportunities in the multidisciplinary design and work method. Here comes qualitative research because it allows us to understand the user in a more empathetic, more real way.

Finding different views on the same user helps to understand it in a more comprehensive way, identifying elements that are not obvious, allowing the questions to be framed and generating the answers in a different and innovative way.

This recovers and reorders the design practice, from another level, integrating a stage where emphasis is placed on context and user research. It is important that the designer take on the challenge of carrying out certain activities to limit the scope of the problem, understand the context and know its user.

Some techniques that have been proposed to design students to get closer to reality, have to do with techniques that involve interaction and dialogue, such as semi-structured interviews, or shading, and techniques that privilege observation, such as participant observation, visual ethnography and user footprint analysis.

The registration of activities to gather information is a fundamental element in this process. The designer is afraid to approach someone he does not know, to observe him, but this is understandable due to the distrust and fear that exists in societies like ours. However, one of the best ways to approach reality and users is to explain the purpose of interacting with them.

The designer is used to propose solutions, but it is very difficult for him to ask himself the questions. Qualitative research and the approach to reality promote this type of professional training, make the designer assume roles of researcher, but requires many conditions so that it can be carried out.

The training as researchers of the architects and industrial designers of CyAD UAMX

Something that differentiates us from graduates of other schools of industrial design and architecture is that our professionals leave with the sensitivity and training to approach reality and identify problems and opportunities for innovation.

The new professional field implies that the designer begins to train as a researcher. Beyond transforming the environment through objects or spaces, both the architect and designer will have the ability to identify and raise pertinent questions to interact with the context.

Moving towards more complex solutions demands an interdisciplinary approach and a deep understanding of the context. It is essential, without a doubt that based on qualitative research tools, designers and architects can collect the information properly to continue with the different stages of the design processes, without losing sight of the fact that design is a tool that allows improving quality of life of beings inhabit the planet.

The importance of the researcher (architect and / or designer) in qualitative research is fundamental, since it promotes a relative perspective, where the filters granted by the researcher will allow a better understanding of the phenomena to investigate. For the analysis, this new researcher provides interpretive frameworks based on his previous knowledge and experiences, which, complemented with those of other researchers, will allow him to approach the complex interpretation of the facts.

The possibilities of this research are unlimited, because as the world continues its progress and transformation in the historical course, the phenomena are reconsidered and allow new interpretations. This research is able to adjust to the events and problems that the researcher faces day by day.

The project practice, typical of architecture and design, implies a process of reflection based on the origins of the project. Where, beyond the creative process, the designer assumes himself as a researcher and plays an important role in the interpretation of social events and events in order to achieve more complex and pertinent approaches to the reality that is happening.

The designer, beyond his creative posture, where he only produces the answer to the questions that others have asked, assumes a much more reflective, participatory and critical posture, which allows him to integrate from the very origin of the project, identifying opportunities for Innovation in social contexts.

Beyond the traditional profession of architects and designers, the researcher profile demands a professional with specific skills and abilities, participating in systematically organized processes which increases their knowledge, obtaining better results.

According to Buchanan (1995) the characteristics that develop in the formation of design thinking are: curiosity, decision making, interaction with others, empathy and argumentation, all in relation to the development of design projects. In addition to the five characteristics Buchanan mentions as essential in the designer's thinking, his profile as a researcher must be strengthened. This implies that the investigator's own characteristics are integrated with those of the designer allowing a designer / architect profile more sensitive to reality.

There are skills and qualities that are already in the individuals and it is only to promote, encourage and induce them to achieve a higher level of development and sensitivity, as Sánchez mentions: “for man, knowing is task and undertake. It is a program of life in terms of being historical and social” (2010, p. 60).

In this sense, six operations are proposed that will encourage the growth of designers as researchers: opening operations such as observation and reading. Expression operations in reference to the exchange and production of signs. Operations of creativity and rigor, operations related to socialization, construction and conceptualization operations, and strategy operations that involve teamwork, planning and decision-making capacity (Sánchez, 2010).

The design student is called to leave his role as a formalizer and to assume a new role of researcher that develops specific skills and allows him to position himself in the current design landscape worldwide with an innovative profile. The use of tools from interdisciplinarity, contextualization of referents and arduous analysis processes are the keys to successfully linking research processes with the results of the design process.

Qualitative research then becomes a basic tool of the entire design process and interdisciplinary linking is essential to carry out design projects. Thinking becomes a basic skill and conducting analysis leads directly to the process of generating strategies for innovation.

Conclusions

Speaking of this "new" method that allows a vision of practice integrated to the designer, neglects the responsibility of the shape of the objects that make up the artificial environment of man, to think more broadly and participate in the whole process that It involves generating a design project.

For this it is essential that the designer develops tools as a researcher, in the first phase of the design; as a designer whose skills are given in the materialization and

execution of the design project, and also as a communicator to be able to enter the deliverable phase and be able to adequately transmit the results of the design process.

The designer and the architect then become more participatory and less intuitive, managing to consolidate an integrative, analytical and critical thinking in search of the best answers to concrete problems of reality. The challenge would have to be to integrate content into your training that will help you assume this new role and allow you to have a greater vision and critical thinking ability.

The teaching of design and architecture within the Modular System requires a more qualitative approach, because from a concrete reality, students must be more sensitive and receptive to the phenomena that occur in their reality. But they must also be able to analyze and synthesize these concepts in more relevant and real solutions.

The architect and designer graduated from UAM X, develops research skills that allow him to differentiate himself from graduates of other institutions that use traditional teaching methods. Some developed features have to do with the capacity for analysis and reflection, critical thinking and diversity of paradigms, models and procedures that give them a complex vision to interpret reality, in an open search for knowledge, which make the human realities better understood and its problems in context.

As Bauman would conclude in his reflection on education: "We still have to learn the art of living in an oversaturated world of information. And we must also learn the even more difficult art of preparing the next generations to live in such a world" (Bauman, 2007, p.46).

References

[1] Bauman, Z. (2007) Los retos de la Educación en la modernidad líquida. Barcelona: Gedisa.

[2] Bermúdez, L., Zamora, L. (2016) Design thinking ≠ pensamiento creativo: Repensando un modelo para la innovación en torna al contexto. En Modelos clave para el diseñador ante los escenarios de cambio. (pp. 409-433). Ciudad de México: Newton-UAM Cuajimalpa

[3] Brown. T. (2008) Design Thinking. Harvard Bussines Review. América Latina. Recuperado en el 3 de noviembre de 2011 en https://emprendedoresupa.files.wordpress.com/2010/08/p02_brown-design-thinking.pdf

[4] Buchanan, R. (1992). Wicked Problems in Design Thinking. Design Issues, 8(2), 5-21. Cambridge: The MIT Press.

[5] _____. (1995). Rhetoric, Humanism, and Design. Discovering design: explorations in design studies (pp. 23-66). Chicago: University of Chicago Press.

[6] Creswell, J. (2013). Qualitative inquiry & research design: Choosing among five traditions. Thousand Oaks: Sage.

[7] Denzin, N., Lincoln, Y. (2011). En Cap 1 Introducción general, la investigación cualitativa como disciplina y como práctica. En: El campo de la investigación cualitativa. Volumen I. (pp. 43-101). Barcelona: Gedisa.

[8] Denzin, N., Lincoln, Y. (2012). Introduction: The Discipline and Practice of Qualitative Research. The Sage handbook of qualitative research. 4th ed., pp. 1 – 19. Thousand Oaks, CA: Sage. Recuperado el 16 de septiembre de 2019 de https://www.sagepub.com/sites/default/files/upm-binaries/40425_Chapter1.pdf

[9] Munari, B. (2004). ¿Cómo nacen los objetos?: Apuntes para una metodología proyectual. 10ª reimpresión. Barcelona: GG Diseño.

[10] Sánchez, R. (2010). Enseñar a investigar. Ciudad de México: IISUE UNAM.

[11] Taylor, S., Bogdan, R. (1987). Introducción a los métodos cualitativos: la búsqueda de significados. Barcelona: Paidós ediciones.

[12] Universidad Autónoma Metropolitana (1978) Informe de la Rectoría 1974-1978 Recuperado el 28 de septiembre de 2019 de http://www.uam.mx/sah/pre-pa/tema04/villarreal74-78/sistemamodular.htm

[13] Universidad Autónoma Metropolitana (2012), Xochimilco Document, 10ª reimpresión. Ciudad de México: UAM.

[14] Universidad Autónoma Metropolitana (2013), University Legislation: General Policies. Ciudad de México: UAM

[15] Zamora, L., Bermúdez, L. (2018) Investigación cualitativa, etnografía y diseño. Los aportes de otras miradas en el proceso de diseño. En Métodos cualitativos de investigación en educación superior (pp. 143-168). Ciudad de México: Newton – UAM Xochimilco

[16] Zamora, M. (2016) El nuevo rol del diseñador industrial. En el Diseño a debate: Diversidad y aprendizaje. (pp. 336-347) Mexico D.F: UAM-Comaprod. Disponible en http://www.comaprod.com/wp-content/uploads/2018/06/Diseno_a_Debate.pdf

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Specifics of Algorithmization in Data Culture

Tadeusz Miczka

Professor at the Institute of Culture Studies,

University of Silesia in Katowice, Poland

Abstract

Information societies effectively transform existing cultures. New cultures are variously defined, but because of the fact that they are dominated by information, the term “data cultures” seems to be the most relevant name for them. Although, it is possible to create a single global data culture in the future and such predictions prevail in academic and non-academic reflections on this subject, so far in addition to global trends there occur local data cultures, what also dynamizes and enriches both individual and collective identities. As Kazimierz Krzysztofek aptly notes, in this situation: "The greatest contradiction of the civilization of the 21st century is drawn. On the one hand, a continuous imperative: be creative and innovative, on the other hand, an increasing pressure on prediction of people’s behaviour, because unpredictability causes chaos, which cannot be managed". In other words, one of the most important social issues today is to create some order in data culture / cultures (often pictured by columnists and researchers as a "magnetic storm"), to reduce its / their infinite complexity, i.e. simply the algorithmization process. With regard to culture, it is not possible to use unequivocally a mathematical algorithm that is the most precise, or a genetic or hormonal algorithm that functions in nature, because accustoming cultural chaos is always strongly ideologized. The algorithm should be treated as a metaphor used to explain cultural phenomena, especially their developmental tendencies. For the researcher of contemporary societies, it is very important to answer the question: what proportions of structure and network are the most beneficial for the survival of data culture / cultures? This answer also directs reflection on the quality of life of individuals and societies, limiting or promoting individualism and collective intelligence in the era of hyper-digitization. These considerations are limited to the initial characterization and evaluation of the information algorithmization of man. The author of the study refers to the concepts of researchers from different countries, highlighting the specificity of today's algorithmization, among others the model of ambient perception, which facilitates participation in the networked information environment, scope and reach of the big data phenomenon, forms of data visualization, personalization of content, Isotype visual language,

network custody, data journalism and others. In conclusion, it is pointed out that the information algorithmization of man is constantly growing, which proves that data management strategies weaken the phenomenon of information overload through the logic of numerical civilization, which limits diversity, seeking to count, record and globalize everything.

Keywords: information society, data culture, algorithmization, distraction, strategy, mobile application, authority

Introduction

The development of every society is an extremely complex process, and therefore its description must be limited to presenting only its dominant features. Modern society is most often referred to as the information society, because information is one of its essential products of important utility and culture-forming value for it, largely determining its progression. Hence, information is the axis of our attention and the following considerations. More precisely, we will focus on high technology, which for half a century, transforming mostly into objects of everyday use (so-called soft technology), makes further breakthroughs in interpersonal communication and all spheres of life, i.e. when compared to the past it extremely speeds up production, storage, processing, sending and exchanging various types of information. In the 21st century, almost every member of society, regardless of the diversity of these societies in terms of economic and cultural development, has a sense of increasing information redundancy and experiences difficulties in solving many problems related to it.

Information societies transform, quite radically, all - small, large, highly developed, and underdeveloped - local identity cultures, making them strongly dependent on a very dynamic and expansive global culture, a non-identity culture in the traditional sense of personal identification. Many boundaries and principles that have conditioned previous social orders are blurred. First of all - as the media sociologist aptly emphasizes –

In the modern world, hierarchical structures have become completely outdated. The “up-bottom” hierarchy of information flow gradually rots - it can be metaphorically said that it has not stood the test of time. A knowledge and information based society even forces us to function in more flexible, egalitarian structures, with a rapid circulation of information and other resources, which is ensured only by the network structure. [...] Networks, however, are not amorphous, you can determine their number of connections, orientation, reciprocity, transitivity, density, strength and other elements thanks to which we can analyze what position individual units have in the network and how the network affects interpersonal interactions. However, life in a network society is not free from disadvantages – taking into account the high complexity of the structure, there may occur problems with coordination of activities [...] (Szpunar, 2005, pp. 82-83).

The problems mentioned above, regarding not only actions, but also the coordination of thinking, have become very complicated and multiplied over the last several years for many reasons, not only those related to the expansion of technology. Referring to the classical typology of social time, by Georges Gurvitch, it can be stated that at the end of the second decade of the 21st century, information societies live in *le temps de l'incertitude* (uncertain times), in which events develop in a very variable rhythm, and the present shows a pronounced advantage over the past and which - in the common sense - are increasingly taking the form of *le temps explosive de la création* (extremely explosive time of creation), based primarily on discontinuity, instability and randomness. According to the French researcher, the uncertain time still accelerates its course, while the explosive time increases the risk and forces strenuous effort beyond one's capabilities (Gurvitch, 1961, pp. 37-39). Some tired and frightened by the pace of life members of information societies even believe that the order of time sequence embraced in the Latin sentence: *Animus meminit praetoritorum, praesentia cernit, future providet* (Mind remembers the past, sees the present, predicts the future) has already been unduly violated, while others even say about the "magnetic storm" (rapid and intense changes in the physical and symbolic parameters of the information stream) that is already underway in the communication space. Even if such a radical position is rejected, there is no doubt that one of the most important social issues is to create some order in this space.

This space, i.e. the culture produced by the information society, is today referred to as the data culture, because in the public and private sectors, both employees and decision-makers most often shape their knowledge and make decisions based on the principles of the so-called datafication, i.e. the use, mostly in the form of analysis, of an increasing number of numerical data. Datafication, already dominating in business, industry, economy and more and more often in politics, also enters colloquial communication, art and fun, it is everywhere and involves managing an enormous amount of information that on the one hand increases the sense of redundancy, but – on the other - is also a the source of the wealth of cultures that can expand the possibilities of human cognition and develop the creativity of individuals and social groups.

Researchers generally use the terms "information" and "data" interchangeably, although colloquially, "data" is treated as raw facts, and "information" as contextualized data having a specific meaning. However, in communication practices,

millions of books stored on library shelves contribute to overload in the same way as terabytes of databases on server magnetic tapes. As long as the Web user does not interact with them in a real communication situation, there is basically no difference whether we are talking about data or whether we use the concept of information (Piekarski, 2017, p. 17).

In other words, in the pragmatic approach that we are interested in this text, the idea of the data culture is the concretization of a new and higher form of culture based on

the network structures of functioning of the information society. The data reinforce the hitherto endless complexity of life in contemporary individuals and societies, and cast a stronger, partly new light on the issues of the adopted principles of social orders, i.e. procedures of conduct in specific circumstances. These processes receive a powerful impulse for data production via the "Internet of Things" and the so-called smart homes, what researchers pointed out a few years ago. For example, one of them wrote:

It is estimated that there are about 40 billion devices integrated into the network, which means that there are already 5 times more devices connected to the network than people. By 2017, the ratio of devices to people will increase as 1 to 130. Their number in the long term [i.e. today - T. M.] will exceed a billion. All these objects emit an avalanche of data (Krzysztofek, 2015, p. 59).

The same researcher several years earlier pointed out the greatest contradiction of the 21st century civilization. On the one hand, a constant imperative: be creative and innovative, on the other, an increasing emphasis on predictable human behavior, because unpredictability causes chaos that cannot be managed. This was relatively less important when the technology of the era of mechanics imposed certain procedures on human muscles and senses. However, today's technologies replace some of the brain's functions by imposing their algorithms on it (Krzysztofek, 2004-2005, p. 61).

Therefore, the purpose of this research reflection is to identify the most important algorithms that are developed by the modern information societies as well as the specific features of this algorithmization. We assume that information algorithmization of a human being is exemplary (model) for many other ways of solving current social dilemmas, which, among other things, explains why it is also worth to consider its specificity.

When one talks about different types of recipes used in a world where more and more electronic calculations need to be made, obviously a mathematical algorithm seems to be the most accurate recipe. The very term "algorithm" comes from the Latinized name of the scholar Muhammad ibn Musa al-Khuwarizmi (Latin Algorismus), who in the 9th century in his treatise entiled *Al-kitab al-muchtasar fi hisab al-Jabr wa-al-mukabala* (*A short book on calculating by complementing and balancing*) wrote "rules for performing basic decimal arithmetic operations" (after Knuth, 1968, p. 3).

And at the beginning of our century, the famous mathematician, the creator of "A New Kind of Science", unambiguously announced that "at the beginning of everything there was an algorithm" (Wolfram, 2002), thus expressing the conviction that all scientific and man-made algorithms imitate algorithms of nature. However, the algorithms associated with the latest inventions based on high technology are not created on the basis of simple similarities, neither to mathematical programming nor to natural "recipes", e.g. genetic or hormonal algorithms, although sometimes, as in the case of

ant colony optimization algorithm programs or neural networks, they imitate them a little.

In addition:

Computerization of intelligence, its algorithmization, is something that we know very little about. The very term “artificial intelligence” or [...] algorithmic intelligence seems to be contradictory. We understand intelligence as our, essentially unprogrammed, and therefore non-algorithmic, abilities. For many people, the very idea of an intelligent machine does not sound sensible (Harel, 2002, p. 163).

Regardless of how strongly and widely doubts in these matters are dispelling and deepening, one belief gains absolute certainty in every person today, namely that culture (and everything else also), despite its increasing complexity and openness at the same time, is growing more ordered, i.e. algorithmized. Therefore, it is best to treat the algorithm as the most important - at least for now - metaphor of modern culture, with the help of which one can describe, analyze and explain numerous changes that occur in our lives.

Methodology

The short history of the data culture and research on it unambiguously confirms the need to use multi-methodology in analyses of human algorithmization. It has been proofed by the works by Geoffrey C. Bowker and Susan Leigh Star devoted to the classification and standardization of various aspects of life, both those that are self-steering and those that are subject to external control (1999) and Bowker's article on layering biodiversity processes in "local data cultures" (2000). Furthermore, information algorithms are elements of the organization of systems of all kinds and almost all spheres of human activity, as demonstrated by, among others Microsoft's 2014 incorporation of data culture into everyday life through - primarily company-oriented - services such as Office 365, Azure or SQL Server (in the scope of device and database management as well as network security, in the field of cloud calculating services and others).

The constant increase in the number of algorithms and their increasing range of functioning are factors that somehow automatically impose on researchers a broad, increasingly wider multidisciplinary perspective. The analysis of this cultural phenomenon, always strongly ideologized and politicized, which in these considerations is not the subject of attention, therefore requires parallel and comparative approaches, based on knowledge from various fields and disciplines of science, mainly sociology, anthropology, history, information systems and design.

This is indeed the logic of numerical civilization created and developed by - as Arthur Kroker and Michael A. Weinstein called its authors in the previous century - "virtual class" (1994). According to it, everything should be *ponderabilium*, *calculabilium* and *metrabilium*, which means counted, weighed, measured, recorded, and today it often

means simply - globalized, i.e. mainly algorithmized. The dream of mathematization and computation, which can be derived from the seventeenth-century philosophy of Gottfried Wilhelm Leibniz (the idea of Calculemus) still remains just a dream, because the scale of data production and processing is constantly increasing. New spirals of complexity are constantly emerging, expanding the information spaces arranged by an increasing number of algorithms. Currently, these spirals are wound up by self-recording and self-displaying technologies as well as the machines that communicate with each other by their algorithms (Arthur, 2009).

The geometric increase in new algorithms makes the area of research we analyze extremely dynamic and forces us to engage in some kind of research games, i.e. contamination of various ideas and methods that consolidate the fairly common belief in the fluidity of the studied information reality itself and all descriptive and evaluative approaches to it. Today, basically, one can only point to certain solid elements in this reality, knowing that they will evolve quickly, becoming more perfect, or be replaced by completely new ones.

Algorithms, or strategies to prevent information overload

The phenomenon of information overload has been known since antiquity and was dealt with by representatives of various sciences in each of the past eras (see e.g. Blair, 2010). That is why it is known that the most important commodity causing commonly felt communication disruptions is not information at all, but it is the human attention needed to capture its sense, and more precisely the distraction increasing along with the development of our civilization (North, 2012).

Digital technologies make – as Chad Wellmon convincingly states - the network become more accessible because it seems much smaller and easier to use than we usually have imagined. [...] the discussion about the infinity of information is rather pointless, because we do not experience information as pure data regardless of whether it is a byte or yottabyte, but as data filtered and shaped by the keyboards, screens and touchpads of our digital technologies. Regardless of how impressive the astronomical amounts of information seem to us, our bewilderment and concern about the enormity of data obscure only the image of how we actually interact with them and the world that data and we are part of (2012, p. 67).

Following this trail of thought, it can be stated that also in the networked information economy, in which multitasking currently enjoys great interest (simultaneous use of many communication devices), the most noticeable is the distraction of attention that multimedia users, that is almost all members of the modern society, are constantly trying to defeat.

Information treatment, that is, such a focus of attention, thanks to which data gains meaning, is most often referred to as filtration. The more information we have, the more we need to filter it, and therefore it is very important to have knowledge about filters and on how to use them. Algorithmization involving the use of specific

information selection strategies is one of the most effective filters available to the information society. Its effectiveness largely depends on the members of society with several basic media competences, especially the basics of programming. Depending on the level of mastery of these skills, algorithms of this kind are more or less "friendly" (and useful) to humans, which somewhat jokingly, but with conviction illustrates Douglas Rushkoff's postulate: "program or be programmed" (2010).

In communication practice, we already use many strategies to prevent information overload. In this text, I will mention only those that the previously cited researcher Karol Piekarski considers fundamental and presents them in the chapter of his book devoted to Data Culture, entitled: *Algorithms to the rescue: strategies for selecting content on the Web* (2017, pp. 141-256). The author treats strategies as hybrid phenomena (combinations of technological and design solutions, as well as social content management mechanisms), arising in response to information overload, arising both in a bottom-up way and independently of the intentions of the information society's members as a necessary condition for their adaptation to real needs.

In his opinion:

The development of the World Wide Web has been a logical consequence of phenomena that have been taking place in the knowledge system since at least the beginning of the 20th century - defragmentation and globalization have forced the creation of a universal management system that would allow to reorganize dispersed knowledge using the intellectual capital of various social groups. The overriding principle of the new system was the standardization of protocols, enabling the smooth joining of various pieces of information. Due to this type of interoperability and interconnectivity, it became possible to find and create patterns and principles in a (permanently) disordered information environment (Piekarski, 2017, p. 143).

In other words, the realization in 1989 of the vision of "the potential possibility of connecting anything with anything" (Berners-Lee, 2000, p. 162), i.e. the emergence of a hypertext web, consisted in the wide availability of tools enabling quick combining of content from various sources and developing new communication and cultural standards to maintain the open nature of the Web and to deal with the data overload.

Except for these standards of combining and receiving information, the basis of today's algorithms is also the need to accept the chaotic nature of knowledge and the use of new, much more than before, democratic information processing strategies. New knowledge systems are created thanks to a new type of taxonomy, known as folksonomy, i.e. the categorization of content by spontaneously cooperating people within unhierarchized communities by using arbitrarily selected keywords and bottom-up metadata creation, enabling information selection (Maj, 2009 & 2014, pp. 24-45). These folksonomic strategies are not intended to create a holistic and finite picture of reality, but they can effectively manage information chaos by using partial

solutions. The essence of this form of social filtering is the combination of algorithmization and collective intelligence: "With the right algorithm, we are able to contextualize a disordered set of meta-information, just like search engines or tag and link management systems do" (Piekarski, 2017, p. 172).

In practice, there already exist many forms of algorithmization. The Big Data model, which is a fundamentally new approach to constantly growing information resources, is very popular and still expansive. It consists in replacing existing ways of understanding the reality, primarily in formulating hypotheses, determining - on the basis of an avalanche of information taking into account the preferences of the Web users - the degree of dependence between completely different phenomena, the degree of probability of occurrence of events, risk assessment, etc. (Mayer-Schonberger, Cukier, 2013). Algorithmization based on data visualization also becomes more and more popular, which goes far beyond the existing specialized diagrams development thanks to the use of the universal visual language Isytope. The creator of this language, based on transformation mechanisms that transform almost all figures into images, is Otto Neurath. Transformation mechanisms are creatively adapted to new technology environments, take into account the huge potential of interactivity and virtuality, personalize messages and generate completely new symbols (Neurath, 2010). The language of visualization plays an increasingly important role in the so-called data journalism, practiced by interdisciplinary teams of journalists, programmers and designers searching and using infinite databases.

Among other popular algorithms, there are two alternative ways to filter content: automatic personalization and curated web. The first is based on the automatic adjustment of messages to the preferences of multimedia users, while the second - on the strict selection of information made by network curators. The first generates overproduction of profiles, the second - intermediaries in access to information who replace former gatekeepers. Both methods of algorithmization currently strongly "compete with each other", but most researchers believe that although they have a high potential for creating order, none of them strengthens the credibility of information sources. Karol Piekarski believes that associated with them in attempts to overcome information overload limitations and threats can be reduced with the help of a new model of perception, called ambient perception (2017, pp. 257-284). As it happens in the world of everyware, it covers with algorithmization not only traditional media products, but also everyday objects (Greenfield, 2010). Simply said, various extensions of the human perception apparatus make information easily available at the place and time required by media users, making it easier for them to make specific decisions. The ambient perception model, which is a kind of synthesis of all previously described forms of algorithmization, is a model open to new information phenomena, open to the future.

Algorithms, or utility software programmed for portable devices

All the forms and methods of algorithmization mentioned so far are widely known to the participants of network communication and used by them, even if they are not very educated people and do not deal with programming. Some are primarily used for practicing various professions, arts and sciences, sports and games of all kinds and highly specialized works, but the vast majority, in addition to those profession-related associations and conditions, which rightly let us treat algorithms as tools of power (this is, however, a broad topic for a separate studies), is more or less consciously, useful and effectively used by all members of the information society. The algorithms achieved by means of mobile applications are much closer to them, especially since mobile phones have been replaced by smartphones (since 1992) and the popularity of tablets and phablets (hybrids of both devices) is increasing. We will make a brief review of their huge potential for algorithmizing modern man by following the though path of the researcher, who defines them as "the tissue of everyday life", defining and cataloging all spheres of human life, which is "updated according to the marketing principles of survival on the market" (Orzeł, 2017, p. 13) .

Mobile applications are primarily utility software that kaleidoscopically meet the emerging needs of individuals and social groups, more and more often referred to by programmers as *Multi-Screen Consumer* (combinations of three screens: laptop, tablet and smartphone are the main matrix for obtaining information), and by sociologists as *homo mobilis.* The increase in mobility is associated with the increase in the decision-making of the multimedia user and, therefore, each subsequent update of operating systems entails enrichment and harmonization of their experience. However:

What would a smartphone be without permanent internet access? Undoubtedly, permanent connectivity is the existential foundation of the “smart” prefix and the mobile applications that are a part of it. This system of connected vessels becomes an indicator of a new cultural order. Why not use this fact for effective (and impressive) product promotion? (Orzeł, 2017, p. 79).

Research on these forms of algorithmization conducted around the world clearly shows that the main goal of the expansion of the new mobile order is to develop new consumer behavior. Through the application, producers of various goods try to get into the consciousness of customers, and consumers try to manifest their unique "I" to force them to meet their own needs. That is why current trends on the mobile application market are extremely mobile: the words of Steve Jobs, Apple president, many years ago stating that "people do not know what they want until they are shown" often are true. However, opposite beliefs, a strong firmness expressed by consumers, seem true similarly often too. In addition, the dynamically understood principle of "one size does not suit all people" is spreading in new marketing. Of course, marketers are more effective than customers, but customers have more and

more opportunities and chances to articulate their needs for which they want to pay providing they are met.

Mobile marketing still has huge development perspectives ahead, because it uses increasingly complex strategies for transforming multimedia users into potential consumers. For example, the SoLoMo trend has recently become popular (an acronym for Social, Local and Mobile), based on a combination of social media, geolocation and a rich mobile staffage, offering recipients information that steer their expectations, firmly embedded in their locality and encapsulated in additional up-to-date content (Kelly, 2014). Other trends in this area are illustrated by, among others application versions of branch online and brick-and-mortar stores, a *Brandomesticator* loyalty application, or finally application hybrids promoting famous places and institutions important for some reason.

A separate, very popular and multi-genre group of applications is associated with computer games ubiquitous in the contemporary culture, which are also used on other devices, such as portable consoles, tablets and smartphones. Telephones for new media players have been created since 2003, the pocket games segment is rapidly growing. Furthermore, the tactile trend is intensively developing due to the use of "Augmented Reality". Mobile gaming is constantly diversified and enters complex relationships with new technological inventions, social phenomena and consumption trends.

On the one hand, mobile applications satisfy and even enhance the culture of individuality, on the other, they express the strenuous pursuit of information society members to rationalize and organize their lives. Since 2013, when first selfie and later self-tracking made the smartphone a manifest of human personality, there was an invasion of mobile start-ups, widely known since 2011, which propose ways of dealing with risk through innovation. This type of algorithmization accurately reflects the nature of one of the basic directions of development of modern societies, most often referred to as "controlled destruction", about which an outstanding British sociologist wrote twenty years earlier: "The possibility of destroying the established order of things and opening new roads, and thus colonizing a new fragment of the future, is a feature of the disturbing nature of modernity" (Giddens, 1991, p. 184).

Summing up this fragment of considerations, it can be stated that many everyday devices organize our lives, because they have become ICT hybrids, screens with components of a computer connected to the internet and logged in to the appropriate "cloud". Televisions, telephones, watches, and even ovens, washing machines, refrigerators and coffee machines become "intelligent" and "cross-linked" thanks to mobile applications and effectively facilitate everyday management. Effectively, because members of the information society not only change most of their current communication behaviors, but also acquire completely new competences in the field of information production and processing. These are, among others digital reading, multitasking or recording high quality videos with the use of drones. Some of these

behaviors, however, are quite controversial, e.g. according to the American Psychiatric Association, the constant desire to take pictures and post them on social networking sites can be a form of obsessive-compulsive disorder (*Selfitis* is on the list of personality disorders), and non-reflective checking of everything in mobile applications can lower the level of individual meanings and decision making.

Conclusion

Not only researchers of the information society, but also its ordinary members are at the same time satisfied with information algorithmization, but also do not hide anxiety about the role of computer algorithms in their lives. They more often control their daily thoughts and actions and determine their future.

Algorithms, invisible pieces of code forming the construction and mechanics of the modern era of machines - writes the English mathematician, Hannah Fry - gave everything to the world - from subscribing to information channels in social media, through search engines and satellite navigation, to the systems of recommendation of musical works - and are part of our modern infrastructure on a par with bridges, buildings and factories. We installed them in hospitals, courtrooms and cars. They are used by the police, supermarkets and film studios. They got to know our likes and dislikes; they tell us what to watch, what to read and who to date. At the same time, they have hidden possibilities, **due to which they slowly change criteria of humanity**" [emphasis T.M.] (2019, pp. 12-13).

Though this diagnosis made to the modern world may resound strongly and menacingly, but it is repeatedly backed up by multidisciplinary research conducted on all continents and in the everyday life of several billion users of multimedia and other electronic devices.

H. Fry thoroughly characterizes and analyzes the mechanisms and ways of functioning of algorithms in various areas of life. First, in the chapter with an eloquent title: *Power,* she refers to a chess game played in 1997 by Grandmaster Garry Kasparov with a chess computer, Deep Blue. Man's loss became an opportunity to document the thesis that "the power of the algorithm is not a simple function of what was written in the lines of his code" (Fry, 2019, p. 17), it is "our human tendency to look zero-one - recognizing the algorithms as either all-powerful or useless – that is a serious problem in the age of advanced technologies" (2019, p. 38). The author looks at the algorithms that the police use to create silhouettes of potential criminals, judges to formulate judgments, doctors to challenge previous diagnoses (including problems of overdiagnosis and unnecessary treatment), passengers of autonomous cars to clarify their moral principles, and contemporary artists to predict popularity. Her findings accurately summarize the current state of affairs in the areas of reality we analyze, and contain clear and specific postulates for the further development of algorithmization of the information society and scientific research on it. According to H. Fry, the most important is the attempt to achieve a balance between the strengths

of man and machine: "the best algorithms are those that take into account man at every stage of their operation" (2019, p. 258).

So we should never allow machines to have authority over us, to force us to adapt to them thoughtlessly. However, achieving this goal is not easy, because the algorithms enjoy numerous successes in replacing many human skills and difficult tasks. Although we know that not every aspect related to the human being can be quantified, today at the peak of the extremely explosive time of creation (this old Gurvitch's metaphor gained more and more relevance at the end of the second decade of the 20th century), when almost everyone experiences the ubiquity of high technology, most popular ideologies are based on the dogma of growth, which means that quantity is treated as a measure of the better quality of various processes and human life (Miczka, 2015, pp. 13-17). Quantity has become an obsession for individuals and social groups, and of course this also applies to algorithmization. But do biological nature and structure, today mainly network, and communication mechanisms cope with the solution of this *coincidentia oppositorium* shaping the development of the information society?

The answer to this question is (for now?) negative. Most experts on this subject, like cited earlier Kazimierz Krzysztofek, believe that there is no time perspective that would set boundaries for the growing pressure of information algorithmization of man, on the contrary, technologies producing algorithms are accelerating (2005, pp. 60-73). It is hard not to wonder when one considers the redundancy of information - after all, breaking through such reality requires increasingly better information bypasses, but the point is that adapting technology to man, which today involves the increasingly intensive algorithmization of them and their surroundings, does not dehumanize our species.

Reference

[1] Arthur, B. (2009). *The Nature of Technology: What it is and How it Evolves*. Cambridge-London: The Free Press and Penguin Books.

[2] Berners-Lee, T. (2000). *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. San Francisco: HarperBusiness.

[3] Blair, A. (2010). *Too Much to Know: Managing Scholarly Information Before the Modern Age*. New Haven: Yale University Press.

[4] Bowker, G.C.; Star, S.L. (1999). *Sorting things out. Classification and its consequences (Inside Technology)*. Cambridge, MA: The MIT Press.

[5] Bowker, G.C. (2000). Biodiversity Datadiversity. *Social Studies of Science*, Vol. 30, Issue 5, pp. 643-683.

[6] Fry, H. (2019). *Hello world. Jak być człowiekiem w epoce maszyn.* Kraków: Wydawnictwo Literackie.

[7] Giddens, A. (1991). *Modernity and Self-identity. Self and Society in the Late Modern Age*. Cambridge: Polity Press.

[8] Greenfield, A. (2010). *Everyware: The Dawning Age of Ubiquitous Computing*. Berkeley: New Riders.

[9] Gurvitch, G. (1961). *La multiplicité des temps sociaux*. Paris: Centre de documentation universitaire.

[10] Harel, D. (2002). *Komputery – spółka z o. o. Czego komputery naprawdę nie umieją robić*. Warszawa: Wydawnictwo Naukowo-Techniczne.

[11] Kelly, S.M. (2012). *SoLoMo Revolution Picks Where Hyperlocal Search Left Off* : http://mashable.com/2012/01/12/solomo-hyperlocal-search [12.11.2018].

[12] Knuth, D. E. (1968). *The Art of Computer Programming.1:Fundamental Algorithms*. Boston: Addison-Wesley Professional.

[13] Kroker, A.; Weinstein, M. A. (1994). *Data Trash: The Theory of the Virtual Class*. New York: St. Martin Press.

[14] Krzysztofek, K. (2004-2005). @lgorytmiczne społeczeństwo. *Transformacje. Pismo interdyscyplinarne*, 3-4 (41-42) & 1-4 (43-46), pp. 60-73.

[15] Krzysztofek, K. (2015). Zmiana – transformacja – Big Data: wzajemne sprzężenia. *Transformacje. Pismo interdyscyplinarne*, 1-2 (84-85), pp. 33-67.

[16] Maj, A. (2009). *Folksonomia jako nowy model wiedzy. Komunikacyjne i kulturowe aspekty Web 2.0*, [online:] http://annamaj.wordpress.com/2009/11/04/folksonomia-jako-nowy-model-wiedzy-komunikacyjne-i-kulturowe-aspekty-web-2-0 [10.07.2019]

[17] Maj, A. (2014). The Role of Retention and Data Analysis in the Ubicomp Paradigm. *Transformations – An Interdisciplinary Journal*, 3-4 (82-83), pp. 24-45.

[18] Mayer-Schonberger, V.; Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt.

[19] Miczka, T.(2015). Imperative: "The More the Better" as a Danger for Proper Development of Information Society. *Academic Journal of Interdisciplinary Studies*, Special Issue, Vol.1, No 3, S.1:13-17.

[20] Neurath, O. (2010). *From Hieroglyphics to Isotype: A Visual Autobiography*. London:Hyphen Press.

[21] North, P. (2012). *The Problem of Distraction*. Stanford: Stanford University Press [epub].

[22] Orzeł, B. (2017). *Aplikacja mobilna jako zjawisko kulturowe*. Katowice:Wydawnictwo Uniwersytetu Śląskiego.

[23] Piekarski, K. (2017). *Kultura danych. Algorytmy wzmacniające uwagę*. Gdańsk: Wydawnictwo Naukowe Katedra.

[24] Rushkoff, D. (2010). *Program Or Be Programmed: Ten Commands for a Digital Age*. New York: OR Books.

[25] Szpunar, M. (2005). Od struktury hierarchii do struktury sieci jako dominującej formy porządku społecznego. In: M. Sokołowski (ed.). *Oblicza Internetu – Internet a globalne społeczeństwo informacyjne*. Elbląg: Wydawnictwo Państwowej Wyższej Szkoły Zawodowej: 75-84.

[26] Wellmon, Ch. (2012). Why Google I snt Making Us Stupid…or Smart. *The Hedgehog Review*, Vol.14 No 1:62-78.

[27] Wolfram, S. (2002). *A New Kind of Science*: Wolfram Media.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

The Comparison of Some Methods in Analysis of Linear Regression Using *R* Software

Ilir Palla

Department of Mathematics and Physics, Faculty of Natural and

Human Sciences, University of Korça, Albania

Abstract

This article contains the OLS method, WLS method and bootstrap methods to estimate coefficients of linear regression and their standard deviation. If regression holds random errors with constant variance and if those errors are independent normally distributed we can use least squares method, which is accurate for drawing inferences with these assumptions. If the errors are heteroscedastic, meaning that their variance depends from explanatory variable, or have different weights, we can't use least squares method because this method cannot be safe for accurate results. If we know weights for each error, we can use weight least squares method. In this article we have also described bootstrap methods to evaluate regression parameters. The bootstrap methods improved quantile estimation. We simulated errors with non constant variances in a linear regression using R program and comparison results. Using this software we have found confidence interval, estimated coefficients, plots and results for any case.

Keywords: homoscedasticity, heteroscedasticity, studentized errors, ncvTest.

Introduction

The processes of model fitting are: model selection, parameters estimation, checking the adequacy of the model, calculating confidence intervals and testing hypotheses about the parameters in the model and interpreting the results. In the chosen model we have to consider: the assumption of observation independence or at least unrelated, a single error term in the model, the choice of scale for the analysis, the choice of variables that will be included in the model.

If the variability comes only from variability of errors and variance of the errors term does not depend from the values of the explanatory variable X, thus have equal variance. This property is called homoscedasticity.

A test that controls the dispersion of errors in a regression is called "Breusch-Pagan test" by Breusch, and Pagan, 1979, which control the hypothesis if the variance of

errors is constant versus the alternative the error variance changes with the level of response. With command "ncvTest", which is into the R program we compute a score test of the hypothesis of constant error variance.

In issue 1.1 is treated least squares method to estimate the parameters of regression, hat matrix, leverages, studentized residuals, variance of parameters and the standard error for a particular component.

In issue 1.2 is treated weight least squares method. In this case the errors have unequal variance where the form is known. A test that controls the dispersion of errors in a regression is called "**Breusch-Pagan test**" by Breusch, and Pagan, 1979, which control the hypothesis if the variance of errors is constant versus the alternative the error variance changes with the level of response.

In issue 1.3 is treated bootstrap methods to estimate parameters of linear regression. Finally in issue 1.4 is simulated a linear regression with heteroscedasticity errors. Using R software we have found confidence intervals, estimated coefficients, plots and results for any case. In this issue are shown the commands that are used in R to get the results.

1.1 Least squares linear regression

A vector of observations $y=(y_1,...,y_n)$ having n component is assumed to be a realization of a random variable Y whose components are independently. If we repeat the observation for a value given by the explanatory variable, we will not necessarily have the same value from the explained variable, so we may have another value because they explained variable will be considered a random variable (Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, "Probability and Statistics for engineers and Scientists").

Explanatory variable X will be considered as fixed values, matrix with size $n\times p$.

The linear regression model is:

$$Y=X\beta+e, \tag{1}$$

where: $C(Y/X)=C(e)=\sigma^2 I_n$, $E(e)=0$. $C(e)$ is covariance matrix, I_n is matrix unit $n\times n$,

$E(Y/X)=X\beta$.

To estimate the parameters $\beta_0,\beta_1,...,\beta_p$, we use the least squares method. The sum of the squared errors (SSE) is:

$$SSE(\beta)=\sum_{i=1}^{n}(Y_i-\sum_{j=0}^{p}\beta_j x_{ij})^2=(Y-X\beta)^T(Y-X\beta). \tag{2}$$

The least squares estimators $\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p$, which minimizes $RSS(\beta)$ are:

$$\hat{\beta} = (X^T X)^{-1} X^T y \quad (3)$$

if $(X^T X)^{-1}$ exists. The equation $Y = X\beta$ is called the theoretical regression equation, while $\hat{y} = X\hat{\beta}$ the fitted least squares regression equation.

$$X\hat{\beta} = X(X^T X)^{-1} X^T y = Hy \quad (4)$$

$H = X(X^T X)^{-1} X^T$ is called hat matrix. Residuals: $\hat{e} = y - \hat{y} = (I_n - H)Y$.

Studentized residuals are:

$$r_i = \frac{\hat{e}_i}{\hat{\sigma}\sqrt{1-h_i}}, \quad (5)$$

which if the model is assumed correctly they have variance equal to 1 and $corr(r_i, r_j)$ tend to be small. The quantities $h_i = h_{ii}$ are known as leverages, the main diagonal elements of hat matrix *H*. Studentization of residuals can only correct the natural non-constant variance in residuals when the errors have constant variance. If there is some underlying heteroscedascity in the errors, studentization can not correct for it.

$$E(r_i) = E\left(\frac{e_i}{\sigma\sqrt{(1-h_i)}}\right) = 0, \text{ and } \operatorname{var}(r_i) = \frac{1}{\sigma^2(\sqrt{1-h_i})^2}\operatorname{var}(e_i) = \frac{\sigma^2(1-h_i)}{\sigma^2(1-h_i)} = 1.$$

Residual sum of squares:

$$RSS(\hat{\beta}) = \min_{\beta}[RSS(\beta)] = \hat{e}^T\hat{e} = (Y - X\hat{\beta})^T(Y - X\hat{\beta})y^T = (I_n - H)Y.$$

If $E(e) = 0$ and $C(e) = \sigma^2 I_n$, then

$$E(\hat{\beta}) = \beta, \ \operatorname{var}(\hat{\beta}) = \sigma^2(X^T X)^{-1} \quad (6)$$

$$\hat{\sigma}^2 = \frac{RSS(\hat{\beta})}{n-(p+1)}. \quad (7)$$

The standard error for a particular component is: $se(\hat{\beta}_i) = \sqrt{(X^T X)^{-1}_{ii}}\,\hat{\sigma}$.

If errors are independent and identically normally distributed with mean 0 and variance σ^2, which means $y \sim N(X\beta, \sigma^2 I_n)$, then using the fact that linear combinations of normally distributed values are also normal:

$$\hat{\beta} \sim N(\beta, (X^T X)^{-1}\sigma^2). \quad (8)$$

We can use these results to test hypothesis about parameters and to construct confidence intervals for β. The estimators have normal distribution, if the errors have normal distribution. If the errors do not have normal distribution, then this method cannot be safe for accurate conclusions.

1.2 Weighted least squares method

The basic tool for examining the fit is the residuals, and we have already looked for patterns in residuals and assessed the normality of their distribution. Where the errors are uncorrelated, but have unequal variance where the form is known we can use weighted least squares (WLS). A simple case:

$$E(Y / X = x_i) = \beta^T x_i, \tag{9}$$

$$\text{var}(Y / X = x_i) = \text{var}(e_i) = \frac{\sigma^2}{w_i}, \tag{10}$$

where w_i for i = 1,...,n are known. In this situation the model is:

$$Y = X\beta + e, \ \text{var}(e) = \sigma^2 W^{-1}. \tag{11}$$

The sum of the squared errors in this case is:

$$RSS(\beta) = (Y - X\beta)^T W (Y - X\beta) = \sum w_i (Y_i - x_i^T \beta)^2. \tag{12}$$

Weighted least squares estimations of parameters are:

$$\hat{\beta} = (X^T W X)^{-1} X^T W Y. \tag{13}$$

Weighted least squares is appropriate when the form of the non-constant variance is either known exactly. A test that controls the dispersion of errors in a regression is called "**Breusch-Pagan test**" by Breusch, and Pagan, 1979, which control the hypothesis if the variance of errors is constant versus the alternative the error variance changes with the level of response (fitted values), or with a linear combination of predictors.

Suppose we have a regression model:

$$Y = \beta_0 + \beta_1 X_1 + \cdots \beta_p X_p + e, \quad i = 1, \ldots, n \tag{14}$$

$$\text{var}(e \,|\, \underline{X}) = \sigma^2 h(\underline{X}). \tag{15}$$

If the function h (.) is linear in relation to independent variables, then the last equation can be written in the form: $\text{var}(e \,|\, \underline{X}) = \sigma^2 (\alpha_0 + \alpha_1 X_1 + \cdots \alpha_p X_p)$.

We build auxiliary regression: $\hat{e}^2 = \gamma_0 + \gamma_1 X_1 + \cdots + \gamma_p X_p$. Estimate the regression coefficients with the least square method and check the hypothesis that all coefficients of the least regression equation, in addition to the first coefficient, get the zero value.

$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_p = 0$ *versus* H_a *: there is at least one i where* $\gamma_i \neq 0, i = 1, \ldots, p.$

1.3 Re-sampling errors and re-sampling cases in regression

Re-sampling errors are a bootstrap method to provide more accurate analysis. The application of bootstrap in linear regressions was originally made by Bradley Efron and Robert J Tibshirani (1993). The algorithm for model based re-sampling in simple linear regression (A.C.Davison and D.V.Hinkley 1997, p. 262). Another bootstrap method is re-sampling cases (A.C.Davison and D.V.Hinkley 1997, p. 264). In this method we make no assumption if variance is constant.

1.4 Simulation: True regression, OLS regression, WLS regression and bootstrap methods. Using R software

We start with regression where we know the true linear regression and simulation the errors of regression: $Y = 20 + 10X$,

where Y- define the response variable. X - predictor variable (explanatory variable).

We simulate the errors in regression such that their variances are linearly dependent on variable X. The aim here is to see the difference between results derived from different methods. At first download packages in R program: library(car), library(MASS), library(nlme), library(boot).

The blue and bold words written below are the commands we used in the program R.

x<-seq(0,29,by=1) # explanatory variable

y = 20+10*x + rnorm(30,0,sapply(x,function(x){1+5*x})) # response variable

sim.data<-data.frame(y,x)

OLS regression

fit.ols = lm(y~x, data=sim.data) # OLS regression

summary(fit.ols)

Coefficients (OLS regression):					
	Estimate	Std. Error	t	value	Pr(>|t|)
Intercept β_0	9.348	25.969	0.360	0.722	
β_1	10.024	1.538	6.518	0.00013***	

Residual standard error: 74.56 on 28 degrees of freedom. Multiple R-squared: 0.6028, Adjusted R-squared: 0.5886. F-statistic: 42.48 on 1 and 28 DF, p-value: 0.0001305.

We find the confidence intervals for the coefficients of the OLS regression using: **confint(fit.ols).**

WLS regression

With command **ncvTest** computes a score test of the hypothesis of constant error variance against the alternative that the error variance changes with the level of the response (fitted values), or with a linear combination of predictors. Breusch, T. S. and Pagan, A. R.1979.

ncvTest(fit.ols) # test for heteroscedasticity

Non-constant Variance Score Test. Chisquare = 5.75913, Df = 1, p = 0.01640319. The p-values suggestion to reject the hypothesis of homoscedasticity and we accept heteroscedasticity random errors. We expected this result because such as we designed simulated errors. Now we use weight least square method. Try fitting the regression with weights and see what the difference is.

fit.wls = lm(y~x, weights=1/(1+5*x), data=sim.data)

summary(fit.wls)

Coefficients (WLS regression):

		Estimate	Std. Error	t value	Pr(>|t|)
Intercept	β_0	19.5242	6.5710	2.971	0.00245**
	β_1	9.3218	0.9343	9.998	1.01e-10***

Residual standard error: 7.674 on 28 degrees of freedom. Multiple R-squared: 0.7805. Adjusted R-squared: 0.7726. F-statistic: 99.55 on 1 and 28 DF, p-value: 1.012e-10. We find the confidence intervals for the coefficients of the WLS regression using: **confint (fit.wls).**

We use the following commands to make the graphics of the regressions:

name <-c("True regression","OLS regression", "WLS regression")

colors.reg <- c("red", "blue", "black")

type.line <- c(1, 2, 4)

line.width <- c(3, 2, 4)

plot(x,y, main = "True regression, OLS regression and WLS regression",lwd=1, bty="o")

abline(a=20,b=10,col=colors.reg[1],lty = type.line[1],lwd =line.width [1])

abline(fit.ols$coefficients, col=colors.reg[2], lty = type.line[2], lwd =line.width[2])

abline(fit.wls$coefficients,col=colors.reg[3],lty = type.line[3],lwd =line.width[3])

legend ("topleft", legend = name, col = colors.reg, lty = type.line,lwd = line.width ,bty ="n")

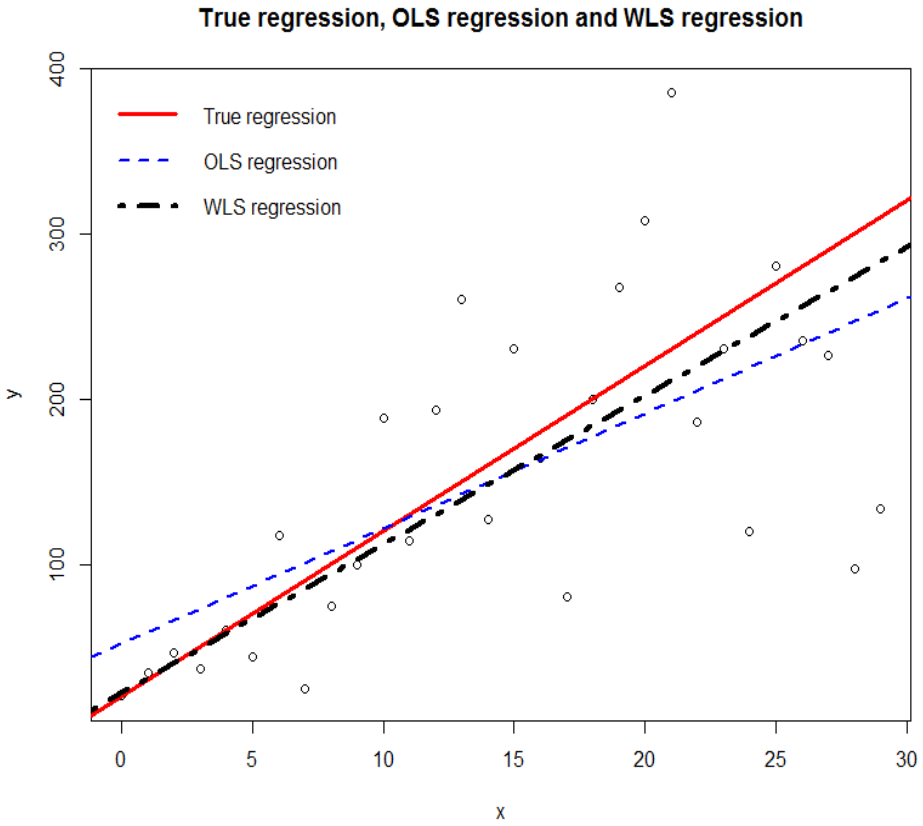

Figure 1 Truth regression, OLS regression and WLS regression for the simulate data.

Bootstrap method: re-sampling errors.

Another method for to estimating regression coefficients and standard deviation is the bootstrap method of re-sampling errors.

sim.lm<-glm(y~x,data=sim.data) # the regression

sim.fit<-function(data) coef(glm(data$y~data$x))

```
sim.diag<-glm.diag.plots(sim.lm,ret=T) # Diagnostics plots
sim.res<-sim.diag$res*sim.diag$sd
sim.res<-sim.res-mean(sim.res)
sim.df<-data.frame(sim.data,res=sim.res,fit=fitted(sim.lm))
sim.model<-function(data,i){
        d<-data
        d$y<-d$fit+d$res[i]
        sim.fit(d)
}
fit.boot<-boot(sim.df,sim.model,R=9999)
```

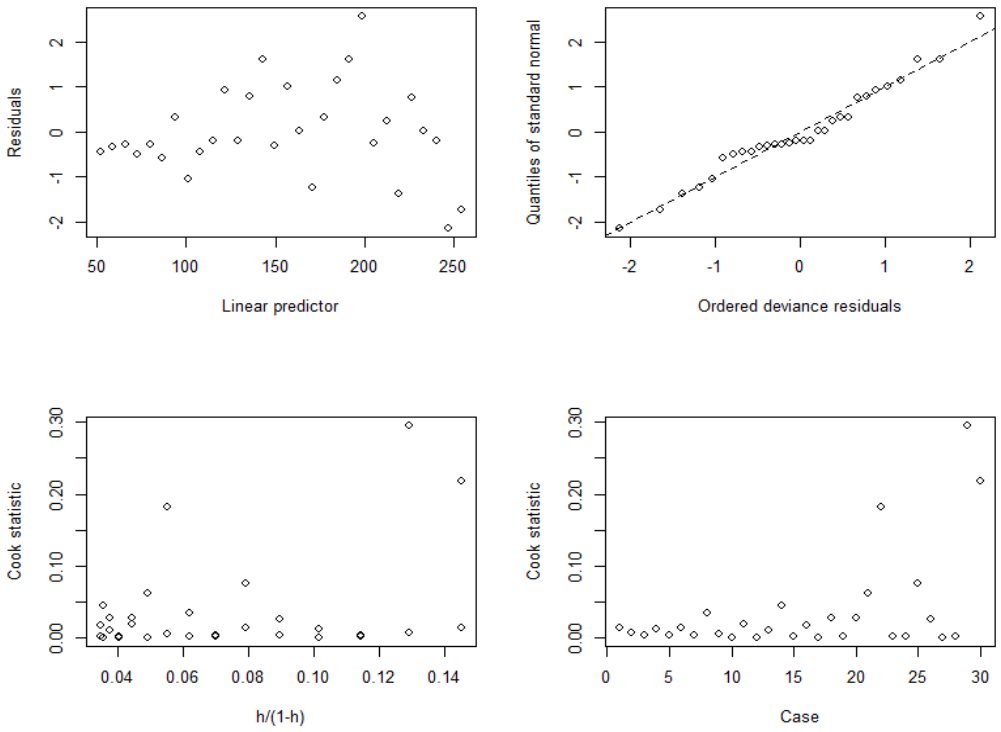

Figure 2 *Diagnostics plots for sim.lm regression.*

The plot on the top left is a plot of the jackknife deviance residuals against the fitted values. The plot on the top right is a normal QQ plot of the standardized deviance

residuals. The dotted line is the expected line if the standardized residuals are normally distributed. The bottom two panels are plots of the Cook statistics. On the left is a plot of the Cook statistics against the standardized leverages. The final plot again shows the Cook statistic this time plotted against case number enabling us to find which observations are influential (Davison, A.C. and Snell, E.J., 1991).

summary (fit.boot)

Bootstrap Statistics (re-sampling errors without weight):

	R	Original	bias	std.error
1	9999	9.038062	0.01855348	26.3335
2	9999	10.045030	0.00809159	1.5606

Bootstrap percentile interval and adjusted bootstrap percentile (BCa) interval.

Calculations based on 9999 bootstrap replicates.

boot.ci(fit.boot, index = 1:min(2,length(fit.boot$t0)),type = c('perc', 'bca'))

Level	Percentile	BCa
95%	(-41.597, 62.66)	(-39.150, 65.582)

boot.ci(fit.boot, index = 2:min(2,length(fit.boot$t0)),type = c('perc','bca'))

Level	Percentile	BCa
95%	(6.98, 13.08)	(7.00, 13.11)

Bootstrap method: re-sampling case.

sim.fit<-function(data)coef(glm(data$y~data$x))

sim.case<-function(data,i)sim.fit(data[i,])

fit.boot.case<-boot(sim.data,sim.case,R=9999)

fit.boot.case

Bootstrap Statistics: (re-sampling cases without weight):

	Original	bias	std. error
t1*	9.3481021	0.3387152	18.40179
t2*	10.023602	-0.0594927	1.64067

Bootstrap method: weight error resamling.

plot(fit.boot.case,index=2,jack=T)

sim.lm.w<-glm(y~x,weight=1/(1+5*x),data=sim.data)

sim.fit.w<-function(data) coef(glm(data$y~data$x))

sim.diag.w<-glm.diag.plots(sim.lm.w,ret=T)

The last command makes plot of jackknife deviance residuals against linear predictor, normal scores plots of standardized deviance residuals, plot of approximate Cook statistics against $\frac{h_i}{1-h_i}$ where h_i are the leverages.

sim.res.w<-sim.res-mean(sim.res.w)

sim.df.w<-data.frame(sim.data,res=sim.res.w,fit=fitted(sim.lm.w))

sim.model<-function(data,i){ d<-data

d$y<-d$fit+d$res[i]

sim.fit(d)}

fit.boot.w<-boot(sim.df.w,sim.model,R=9999); fit.boot.w

Bootstrap Statistics (re-sampling errors with weights):

	Original	bias	std. error
t1*	19.289300	-0.50796443	26.0353
t2*	9.3432310	0.03159331	1.5364

We find bootstrap percentile interval and adjusted bootstrap percentile (BCa) interval only for the second coefficient of the regression fit.boot.w.

Level	Percentile	BCa
95%	(6.318, 12.342)	(6.471, 12.518)

Bootstrap methods: weight errors and re-sampling case

sim.fit.w<-function(data)coef(glm(data$y~data$x,weight=1/(1+5*x)))

sim.case.w<-function(data,i)sim.fit(data[i,])

fit.boot.case.w<-boot(sim.data,sim.case,R=9999); summary(fit.boot.case.w)

Summary table: Estimate errors and standard deviations.

Method	Coefficients	Std. Error
True regression	β_0 =20	-
	β_1 **=10**	-

OLS regression	$\hat{\beta}_0 = 9.348$	25.969
	$\hat{\beta}_1 = \mathbf{10.024}$	**1.538**
WLS regression	$\hat{\beta}_0 = 19.5242$	6.5710
	$\hat{\beta}_1 = \mathbf{9.3218}$	**0.9343**
Bootstrap method resampling errors without weight	$\hat{\beta}_0 = 9.038062$	26.3335
	$\hat{\beta}_1 = \mathbf{10.045030}$	**1.5606**
Bootstrap method resampling cases without weight	$\hat{\beta}_0 = 9.3481021$	18.40179
	$\hat{\beta}_1 = \mathbf{10.023602}$	**1.64067**
Bootstrap method resampling errors with weights	$\hat{\beta}_0 = 19.289300$	26.0353
	$\hat{\beta}_1 = \mathbf{8.957616}$	**1.5364**
Bootstrap method resampling case with weights	$\hat{\beta}_0 = 9.348102$	18.5540
	$\hat{\beta}_1 = \mathbf{10.02364}$	**1.6552**

Conclusions

For linear regression with normal random errors having constant variance, the least squares method of the coefficients estimators and standard deviation are accurate, but when the random errors having non constant variance, the estimators with this method are not accurate. The bootstrap methods improved quantile estimation. The WLS method is the most accurate where we know weights errors. Where we don't know completely the form of the variance of error, we can use bootstrap method resampling errors. We can use packages inside R software to get results.

References

[1] Bradley Efron and Robert J Tibshirani (1993): *An Introduction to the Bootstrap*. Chapman and Hall/CRC.

[2] Breusch, T. S. and Pagan, A. R., (1979) *A simple test for heteroscedasticity and random coefficient variation*. Econometrica 47, 1287–1294.

[3] Cook, R. Dennis; Weisberg, Sanford (1982). *Residuals and Influence in Regression*. New York: Chapman and Hall. ISBN 041224280X.

[4] D. A. Freedman and S. C. Peters: *Bootstrapping a Regression Equation: Some Empirical Results*. Jurnal of the American Statistical Association, Vol.79.385. (1984), pp: 97-106.
[5] Davison, A. C. and Hinkley, D. V. (1997). "Bootstrap Methods and Their Application". Cambridge University Press.
[6] Davison, A.C. and Snell, E.J., (1991). Residuals and diagnostics. In Statistical Theory and Modelling: In Honour of Sir David Cox D.V. Hinkley, N. Reid, and E.J. Snell (editors), 83–106. Chapman and Hall.
[7] P. J. Bickel and D. A. Freedman: *Asymptotic Normality and the bootstrap in stratified sampling*. The Annals of Statistics. 1984, Vol12, No2, pp: 470-482.
[8] R. D. Cook and S. Weisberg: *Diagnostics for heteroscedasticity in regression,* (1983). Oxford University Press, Biometrika, Vol.70, No.1, pp.1-10.
[9] Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye, "Probability and Statistics for engineers and Scientists", 2002, pp. 400-425. Prentice Hall. ISBN 0-13-041529-4.
[10] White, Halbert (1980). *"A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity*". Econometrica. 48 (4): 817–838. doi:10.2307/1912934.

Using Unmanned Aerial Vehicles (UAVs) to Analyze the Urban Environment

Efthimios Bakogiannis

Transportation Engineer & Urban Planner, Ph.D.,

Sustainable Mobility Unit, Department of Geography & Regional Planning,

National Technical University of Athens

Charalampos Kyriakidis

Ph.D. Candidate, Urban & Regional Planner, Transportation Engineer,

Sustainable Mobility Unit, Department of Geography & Regional Planning,

National Technical University of Athens

Vasileios Zafeiris

Surveying Engineer,

Sustainable Mobility Unit, Department of Geography & Regional Planning,

National Technical University of Athens

Abstract

Over the last decades, the evolution of technology has helped us to facilitate various types of works in areas related to land and property management as well as spatial planning. The exploitation of new tools and methods has prompted the international interest in the recording and modeling of geospatial information in more than two dimensions depicted in traditional projects, until then. This has contributed to address a series of issues related to intense urbanization, as well as challenges in identifying complex ownership and building structures. A relatively recent such method is the mapping of buildings and wider spatial units by using Unmanned Aerial Vehicles (UAVs) that has contributed to the production of 3D models by using the appropriate software. This technology finds resonant in recent years in Greece. However, it has not been applied to the mapping of large spatial units such as urban areas. This research paper performs a wide area mapping using UAV. Its purpose is to investigate to what extent the UAVs can do it successfully. For this reason, a brief evaluation is attempted, taking into account the accuracy of the data as well as the cost and time required in relation to traditional techniques. The result justifies the specific technique

that appears to produce good quality metering and quality data while helping to save resources.

Keywords: UAVs, orthomosaic, 3D mapping, urban planning, parking habits, people behavior.

Introduction

Transition from conventional to digital urban planning

From the early 19th century and onwards, numerous models of cities characterised by the concept of futurism, have been presented. They attempted to provide solutions on a series of problems and improve the quality of life of city residents. One characteristic example in Greece was the "electronic urbanism" of the Athenian architect Takis Zenetos, who presented a series of futuristic urban planning proposals (see Yiannoudes, 2016), on an era that was possibly not compatible, with such bold ideas.

In the course of time, urban space is transforming into a digital one. Thus, emerging planning ideas incorporate technological advances. In addition to the proposed plans and planning models, new technologies are gaining ground in the design process, following a series of new tools and technologies utilized by urban planners and practitioners for data collection, analysis and understanding of the specific characteristics of an area as well as the conceptualization and depiction of the proposed interventions.

Sure enough, new methodological tools, like crowdsensing and crowdsourcing techniques come to the forefront of collecting, urban mobility data, environmental data and cadastral information (Basiouka and Potsiou, 2012; Bakogiannis, *et.al.*, 2017; Bakogiannis, *et.al*., 2018; Delitheou, *et.al.,* 2019). In fact, web-based platforms, such as Open Street Map (OSM) (Basiouka, *et.al.,* 2015; Bakogiannis, *et.al*., 2018), have been successfully used to collect urban planning data resulting in the production of Open Source Data, creating a new type of geographic view, which is named 'Neogeography'. This term was proposed by Di-AnnEisnor, during the decade of 2000 (Haklay, *et.al.*, 2008; Stamatopoulou, 2013), in order to describe a process in which people feel free to participate in data gathering processes, apart from traditional consultations (Somarakis and Stratigea, 2014).

Also, the needs for increased information in fields like cadastre and land information management system as well as the need for combined information regarding space and time, resulted in the development of multidimensional models, beyond the two dimensions (Ioannidis, *et.al*., 2000; Dimopoulou, 2015; Doulamis, *et.al.*, 2015 a; Doulamis, *et.al.*, 2015 b). However, even the simplest procedures were implemented through traditional methods, such as the study of urban development through sequential map layout (maps presented different information were used as different layers), today is a process that is simplified through specific applications in a variety

of design software (Stamatopoulou, 2013). In addition to the work done in the field and office, new technologies have also entered the sphere of information dissemination, since the need for citizens' access to information, especially information on urban and societal resources, and their active participation in the design process is the core idea behind the "urban access" design (Nikitas and Rahe, 2013), which in turn forms the basis for promoting a combined urban planning and traffic planning (Christodoulopoulou and Kyriakidis, 2014).

The above are just a few examples of the transition from conventional to digital urban planning, both in the ideological context and in planning process. The increased use of new technologies is clearly related to the simplification of traditional design and planning practices. However, an important parameter is the cost reduction which, in times of economic crisis, the exploitation of alternative practices becomes an important parameter for the elaboration of studies and the implementation of urban interventions (Bakogiannis, *et.al*., 2018).

In that context, this paper focuses on spatial framing processes using UAVs. The specific technique is now widespread for plotting buildings or monuments (Barrile, *et.al*., 2017) since capturing a building or a monument could be implemented by collecting a sufficient number of appropriate photo shoots. Thus, at the same time, measurable and qualitative information is collected (Neitzel and Klonowski, 2011) that can be utilized in various types of analysis. However, the point that differentiates this particular research project is the use of UAVs in a real city environment that consists of an innovation for urban studies in Greece.

In this way, it is expected to collect information on both the physical characteristics of urban space and the dynamic characteristics associated with road traffic. Therefore, it is possible to study the existing traffic –provided that a video is recorded- and proposals for its optimization could be suggested.

Case Study: The wider area of the former railway station in Kozani, Greece

Aim and Objectives

Aim of this paper is to explore the usefulness of UAVs on the capture process of an urban area, with the objective of collecting qualitative and quantifiable information. More specific, the research question could be expressed as: *How UAVs can help urban planners in order to better understand and analyse urban environment (urban form and urban mobility)?* This occurs through the comparative analysis of methods in terms of data quality, implementation timeframe and their operational costs.

Methodology

In order to address the research topics presented above, a decision was made. A pilot study took place on March 2017, as a part of a Sustainable Mobility Unit research project (2017-2018). Study area is located in the city of Kozani, Greece, and includes the Old Railway Station (ORS) area and Olympou Street.

Primary reason for opting in a road on a close proximity to the ORS area was that the use of UAVs in studying mobility parameters should be examined. Moreover, according the regeneration plan of the ORS area, a new cycle path is going to be developed across Olympou Street. In that way, information collected through this research can be used for the implementation of this construction works.

Moving a step further, UAV flight plans were developed. This specific plan was based on the size of the study area, the terrain which had been measured through Google Earth, and finally the technical characteristics of the UAV used. In order to optimize the outcome, the final plan was developed by using an open source software (PiX4dCapture). Flight paths were developed as Figure 1 presents.

Figure 1. Design of flights by using the free software PiX4d Capture. Source: Own elaboration.

Figure 1 presents the six (6) different flights took place. The number of flights is explained by drone's small autonomy. For safety reasons the maximum flight duration was decided to be that of 15 minutes. The flights were fully automated: PiX4dCapture software served as autopilot. On the automated flights, the flight altitude was defined in sixty (60) meters above the ground, so the Ground Sample Distance (GSD)[1] is estimated to be 2.5 cm.

Front overlapping in aerial photographs, that are geolocalized, was selected at 80%, based on the research of Giang, *et.al*. (2017). Finally, It should be noted that, in addition to these photos (1,297 photos), other photos were taken with aerial ramps

[1] GSD is how big each pixel is on the ground.

(239 photos), angled at 45 degrees, so that buildings can be visualized in a 3D model. At the same time, an additional videographer has produced a video of the study area.

In order to collect quantifiable data effectively, Ground Control Points (G.C.P.) were defined. G.C.P. are points with known coordinates which are highlighted before the aerial photography. These points were distributed internally and around the region. In order to be measured, duplex GPS using Base-Rover indicators (for points on the ground) and Total Station (for points in buildings) were used.

Following the collection of data, using the PiX4dMapper software, photos were aligned and placed in the correct shooting position to produce point cloud and then the Digital Elevation Model (DEM). GlobalMapper, another open source software, has also been used for gaining this goal.

Figure 2. The produced point cloud. Source: Sustainable Mobility Unit Archive (2017). And the point cloud presented in Figure 2. Moreover, a series of issues related to traffic and parking management were studied, through the videos produced.

Figure 2. The point cloud developed. Source: Own elaboration.

Overall, in order to assess the quality of the produced results, the new orthomosaic has been compared to the old one as well as to the official urban plans. In that way, the quality of the new orthomosaic was evaluated and the evolution of the spatial

form of this area was also recorded, over time. Moreover, deviations from the official urban plan were checked, demonstrating buildings constructed earlier or in an informal way. This information could be a first means of assessing the legality of structures, saving up human and financial resources.

Method assessment

Following the implementation of the aforementioned processes, a new orthomosaic was developed. For the evaluation of the method, it was decided to investigate the reliability of data, the photo resolution and the interpretation of the urban space as well as the feasibility of work.

In order to evaluate the reliability of the data obtained, it was decided to compare the produced orthomosaic with the pre-existing one and with the layout of the existing urban plan. Following the comparisons, a common conclusion was extracted: there were slight variations between the new and old orthomosaic and the layout coincided to a large extent. As mentioned in the previous section, differentiations were expected due to the fact that the urban form could be changed over time and informal housing could be developed. However, such phenomena have not been observed. Figure 3 shows the match between the layout of the buildings and the new orthomosaic, demonstrating that urban planning regulations were followed to a sufficient extent. This observation is related to the fact that: (a) the area studied was constructed decades ago and (b) the general trend in Greek urban planning system is to recognize and legalize the pre-existing urban morphology.

Figure 3. Comparing the new orthomosaic with the official urban plans. Source: Sustainable Mobility Unit Archive (2017).

Figure 4. Part of the study area presented in the new orthomosaic (2017). Source: Sustainable Mobility Unit Archive (2017).

Figure 5. Part of the study area presented in the old orthomosaic (2007-2009). Source: Sustainable Mobility Unit Archive (2017).

The next parameter to be evaluated was the resolution of the orthomosaic produced. For this purpose, the quality of the new orthomosaic was compared with the old one (the official one) that was developed during the period 2007-2009. The GSD in the

new orthomosaic is estimated to be 2.5 cm (Figure 4) as opposed to the old one that the GSD is estimated to be 30 cm/pixel (Figure 5). Therefore, the clarity of the new map is 12 times better than the clarity of the old one. As a result the information collected is of better quality. Meanwhile, more accurate measurement data have been obtained, validating the accuracy of recorded data. The equipment used was contributed to that, since the resolution could vary depending on the camera used. However, nowadays, good resolution cameras are available in the market at tempting prices. Thus, modern photos are of better quality than the old ones (2007-2009). The results are different when comparing the new orthomosaic with the urban plan produced by using traditional techniques. In the second case, the measurements are of better accuracy and the plan presents much more detail (Table 1). Nonetheless, in cases of large urban areas, like the one studied in this paper or even settlements and towns, the precision provided by this technology is particularly adequate.

	Traditional Mapping methods	Mapping with UAV
Precision	Millimetres to centimetres	centimetres
Pre-processing time (Office)	1 hour	2 hours
Duration at field	3 working days	1 working day
Personnel	2 working groups	1 working group
Measurements processing (office)	2 hours	1 working day
Total processing time	3 working days	4 working days

Every working group is composed by 2 persons, one skilled engineer and one unskilled assistant

Table 2. Comparing traditional techniques with UAVs method.

Figure 6. Mobility and parking behaviors. Source: Sustainable Mobility Unit Archive (2017).

In addition to the quality features and the capabilities of the new orthomosaic, an important parameter for evaluating the method is its cost, measured both in money and in human resources. For that reason, during the survey, a calendar was kept, recording the times required for each individual job and the persons who had to work to complete it. The whole work lasted one working day (14/03/2017) and it was conducted by a two-man workshop. This work is divided into two parts:

(a) Office work: the flight plan developed.

(b) Fieldwork: the definition of the G.C.P. and the flights took place.

After that, it was decided to compare the cost of work performed in relation to the cost of the same work in case it was conducted by traditional mapping methods. Given the fact that this work was not implemented by both methods in order to obtain a precise comparison, assumptions were made as to the number of people required as well as to the duration of the work. According to our assumptions, the number of persons required is twice as high (2 topographers and 2 workers as support staff) that will have to work for 4 days. On the basis of the above data, it was proved that implementing the work by using UAVs is more efficient in economic terms than conducting the same project by using traditional techniques. Volkmann (2017), who carried out a comparative analysis between traditional methods and UAVs focusing on economic terms, reaches a similar conclusion.

Given the fact that the precise calculation of the cost is a procedure that is also affected by a series of other criteria (*i.e.* the area that has to be captured), beside the working hours, it was decided not to present specific budgets but rather a relation of expected cost of the two methodologies, as resulted by Table 1.

Finally, it should be mentioned that through this project it was also possible to create a data base concerning street behavior. More specific, traffic and parking habits were recorded, as presented in Figure 6. Moreover, pedestrians' behavior was also studied by analyzing the videos. In a different case, this could be possible only by observation, id est researchers who would observe people's behavior in specific street sections. Such traditional techniques are economic inefficient due to the fact that the number of researchers required is higher.

Conclusions

Technological evolution has significantly impacted the design of research and the spatial planning process. A range of new technological tools are now widely used in data collection as well as in the depiction and evaluation of planning proposals. UAVs are such tools that can make a significant contribution to mapping cities.

In the context of this study, UAVs used to map an urban area that is going to be regenerated; the Old Railway Station of Kozani and a road located in a close proximity to the station. The results derived helped us in comparing this method with traditional methods. More specific:

- The comparison between the orthomosaics performed showed that the resolution of the new orthhomosaic was quite better than the previous one. As a result, the accuracy of the qualitative and quantitative data will be high enough.
- Using UAVs seems to be a cost-effective procedure comparing to traditional methods, since the demands on time and people were considerably reduced.
- Mobility patterns and street behavior was also recorded.
- Apart from 2D topographical diagrams, 3D mapping can also be developed. 3D capture provides better understanding of the study area.

The above findings resulting from this application of a UAV could be a tangible proof of their usefulness in surveying in large urban environments, where the needs are more in relation to private ownership.

Finally, taking into account the possibility of video recording that UAVs can provide, it is believed that their use in urban planning can be extended to fields like traffic management and study of human behavior, contributing to a better understanding of the parameters influence the function of urban spaces.

References

[1] Bakogiannis, E., Kyriakidis, C., Siti, M., Iliadis, F. and Potsiou, C. (2017). Toward a methodology for noise mapping using VGI. In: *VGI: Emerging applications in public science and citizen participation-FIG Commission 3 Annual Meeting*, Lisbon, Portugal, 27-30 November 2017.

[2] Bakogiannis, E., Kyriakidis, C., Siti, M., Kougioumtzidis, N. and Potsiou, C. (2017). The use of VGI in noise mapping. In: *GeoPreVi 2017*, Bucharest, Romania, 14-16 September 2017.

[3] Barrile, V., Gelsomino, V. and Bilotta, G. (2017). UAV and computer vision in 3D modeling of cultural heritage in southern Italy. In: *IOP Conference Series: Materials science and engineering*, New Delhi, India. Available at: https://iopscience.iop.org/article/10.1088/1757-899X/225/1/012196 (Accessed January 2018).

[4] Basiouka, S. and Potsiou, C. (2012). VGI in cadastre: A Greek experiment to investigate the potential of crowdsourcing techniques in Cadastral Mapping. *Survey Review*, 44(325), 153-161.

[5] Basiouka, S., Potsiou, C. and Bakogiannis, E. (2015).Open Street Map for Cadastral Purposes: An application using CGI for official processes in urban areas. *Survey Review*, 47, 344 (2015), 333-341.

[6] Christodoulopoulou, G. and Kyriakidis, C. (2017). *Using a combined urban and traffic planning for promoting sustainable mobility in the Municipality of*

Piraeus. Available at: http://www.citybranding.gr/2014/12/blog-post_9.html (Accessed September 2017).

[7] Delitheou, V., Bakogiannis, E. and Kyriakidis, C. (2019). Urban planning: Integrating smart applications to promote community engagement. Heliyon, 5(5), e01672.

[8] Dimopoulou, E. (2015). *nD Cadastre: Planning-Development-Standards-Modeling* [e-book]. Kallipos, 2015. Availabe at: https://repository.kallipos.gr/bitstream/ 11419/3661/1/nd_cadastre.pdf (Accessed January 2018).

[9] Doulamis, A., Soile, S., Doulamis, N., Chrisouli, C., Grammalidis, N., Dimitropoulos, K., Manesis, C., Potsiou, C. and Ioannidis, C. (2015a). Selective 4D modeling framework for spatial-temporla land information management. In: 3rd *International Conference on Remote Sensing and Geoinformation of the Enviornment,* Paphos, Cyprus, 16-19 March 2015.

[10] Doulamis, A., Doulamis, N., Ioannidis, C., Chrysouli, C., Grammalidis, N., Dimitropoulos, K., Potsiou, C., Stathopoulou, E. and Ioannides, M. (2015b). 5D modelling: An efficient approach for creating spatiotemporal predictive 3D maps of large-scale cultural resources. In *25th International CIPA Symposium*, Taipei, Taiwan, 31 August-5 September 2015.

[11] Haklay, M., Singleton, A. and Parker, C. (2008). Web mapping 2.0: The neogegraphy of the GeoWeb. *Geography Compass*, 2(6), 2.011-2.039.

[12] Ioannidis, C., Potsiou, C., Soile, S. and Badekas, J. (2000). Detailed 3D representation of aerchaelogical sites. *Proceedings of the International Archives of Photogrammetry and Remote Sensing (ARPS)*, Vol. XXXIII, Part B5. Amsterdam 2000, 642-649.

[13] Neitzel, F. and Klonowski, J. (2011). Mobile 3D mapping with a low-cost UAV system. In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Zurich, Switzerland, 2011, 37,1: C22.

[14] Nikitas, A. and Rahe, U. (2013). Designing the metropolitan future of Shanghai: A local interaction platform looking to incorporate urban access design considerations in planning for a faier, denser and greener Mega-City. In: *10th European Academy of Design Conference «Crafting the Future»,* Gothenburg, Sweeden,17-19 April 2013.

[15] Stamatopoulou, C. (2013). *3D representation for the exploitation of touristic routes in the Municipality of Lamia* – M.Sc. Thesis University of Thessaly. Available at: http://ir.lib.uth.gr/bitstream/handle/11615/41528/11416.pdf?sequence=2&isAllowed=y (Accessed January 2018).

[16] Somarakis, G. and Stratigea, A. (2014). Public involvement in taking legislative action as to the spatial development of the tourist sector in Greece – The «OpenGov» platform experience, *Future Internet*, 6 (2014), 735-759.

[17] Sustainable Mobility Unit (2017-2018). *Research for the Municipality of Kozani for preparing an action plan towards Sustainable Mobility: Framework for a strategy to raise people awareness and enhance local community towards environmental and traffic problems*. Scientific director: Vlastos Thanos.

[18] Volkmann, W. (2017). *Small unmanned aerial system mapping versus conventional methods.* CTA Working Paper 17/07. Wageningen: CTA. Available at: https://publications.cta.int/media/publications/downloads/1987_PDF.pdf (Accessed January 2018).

[19] Yiannoudes, S. (2016). *Architecture and adaption. From cybernetics to tangible computing*. New York, NY; London,UK: Routledge.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

The Design of a Tomato Powder Processing Plant with a Capacity of 75,000 Tonnes Per Annum in Oyo State, Nigeria, West Africa

Osaretin N.I. Ebuehi

Department of Chemical and Petroleum Engineering, Faculty of Engineering, University of Lagos, Lagos, Nigeria

Abstract

This project designed a plant that will process 250 tonnes of raw tomatoes per day. The duration of this production was spread across 24 hours and operated for 300 days in a year, which gave a yield of 17.5 tons of tomato powder per day. From these specifications, the best process route was synthesized. The process route selected involved choosing and sorting the tomatoes, washing, blanching, blending, spray drying (as the drying process), packaging and labeling. Through a series of investigations, the most feasible way to dry tomato powder on an industrial scale proved to be the spray-drying method because the process was very rapid, required low labor cost, and was relatively simple to operate and sustain. The particle size was also easy to control with this technology, making it easy to correlate with product standards. The equipment to be used was then identified which included the machine vision, a spray washer, a water Blancher, an industrial blender, a spray dryer, and a powder packaging machine. Material and energy balance were evaluated around the whole plant. A process flow diagram and basic piping and instrumentation diagram were also computed using engineering software such as Microsoft Visio. From the previous work done, the selected process unit, the spray dryer, was designed with a detailed piping and instrumentation diagram made around it. The site of the plant was resolved, adjacent oriental foods along the Lagos-Ibadan Expressway. It was necessary to site the plant in an area where raw materials can easily be delivered to. The area selected was directly linked to a major highway, the Lagos-Ibadan Expressway, thereby making transport quite affordable. The site layout alongside the plant layout was constructed. The layouts clearly showed the standard flow process.

Keywords: tomatoes, processing plant, storage of tomatoes, tomato powder production, spray drying.

Introduction

Tomato is a red, edible fruit that has high nutritional value and is versatile in the food industry. However, it spoils easily. It was estimated that about 65% of the tomato produced in Nigeria spoils due to lack of storage (Hayes, Smith, & Morris, 2010).

A feasibility study was carried out on the project and it was concluded that it is technically feasible to construct the plant because the equipment is available, process routes are feasible and process conditions are attainable. The best way to produce the tomato powder on an industrial scale was using the spray drying method. This method was chosen because it was very rapid, required low labor cost was relatively easy to carry out and particle size of the product can easily be controlled.

The site selected for our plant is in the Oluyole local government area in Oyo State, Nigeria. It is going to be adjacent to Oriental Foods Industries Limited. It has an area of 4,000km^2 and a population of 202,725 as of 2006. It shares boundaries with four native governments, Ibadan South West, Ibadan South East, Ona Ara, and Ido ingenious government all within the Ibadan city. The site was selected due to its proximity to the Lagos-Ibadan expressway and the availability of labor and land.

The aims and objectives of the Chemical Engineering design project were to address the challenges that arose as a result of perishable qualities of tomatoes by choosing a process route for a plant for the production of tomato powder processing plant, carrying out a feasibility study to serve as an indicator of the technical and economic viability of the process, design a process plant to produce 75,000 tonnes per annum in Oyo State, Nigeria. The key challenges that we faced by the plant in the processing of tomato powder included tomato price fluctuation and maintenance of product quality. These challenges and others were dealt with by proper project management and maintenance practices.

The varieties of tomatoes available in the market including Dandino, Danyeka, UTC, Roma, Ibadan local, etc. They had different tastes and coloration, and for quality and attractive production, varieties can as well be blended such as Dandino (having the sweetest taste) and UTC (having the brightest deep Red color) (Abdulmalik et al., 2014). The best species of tomatoes that are grown or cultivated in Nigeria are the Roma tomatoes. These acid tomatoes are known for their chewy flesh and relatively low water content, about 93% moisture content, which makes them a good choice for making tomato powder (Jayathunge et al., 2012). Hence, for this design, Roma tomatoes will be our choice in producing tomato powder

The raw tomatoes were obtained from Kano, Taraba, Gombe, Bauchi, Kaduna, Sokoto, Zamfara, Katsina, and Jigawa states. This was administered to keep up with the high capacity needed to be processed per day.

Fig. 1. Roma tomatoes cultivated in Nigeria.

There are a variety of ways to produce tomato powder from raw tomatoes, the main process to be considered to be the dehydration process by which moisture was removed from the tomatoes. The universal routes taken in the production of tomato powder include choosing and sorting the tomato, washing, and blanching. From the blanching process, there are different routes that plants can take to get tomato powder. For some processing plants, after the blanching process, the tomatoes are pulped to remove the solid particles within, then dried to obtain tomato powder (Sylvanus & Victor, 2012). For other processing plants, the tomatoes are sliced after blanching, then dried, after which the dry tomato slices are milled to get tomato powder (Arjenaki, Moghaddam, & Motlagh, 2013).

The process routes taken by 7 senior year Chemical Engineering students in my research group included choosing and sorting the tomatoes, washing, blanching, blending, drying, and then packaging the tomato powder. Blending was picked over pulping so that all the solid particles will contribute to the yield of the tomato powder.

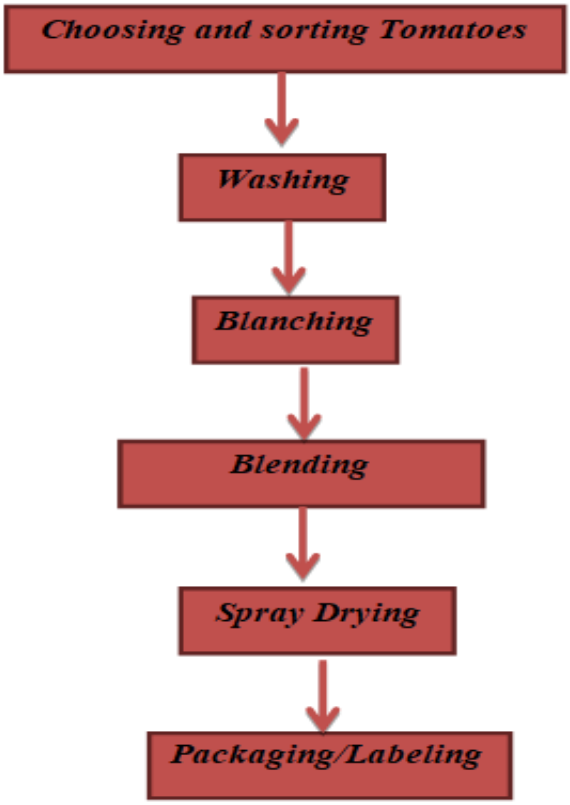

Fig. 2. Selected Process route

Sorting refers to the removal of rotten, damaged, and cracked tomato fruits from healthy, bright, and clean fruits. The raw materials have to be processed as soon as possible (within 48 hours after harvesting) to avoid deterioration. The efficiency and effectiveness of sorting govern the quality standard of the packing lines and the product (Jarimopas & Jaisin, 2008). Consequently, it is imperative to have a swift, coherent, genuine, and vigorous approach for sorting. The machine vision will be used for the sorting process. It is the best technique for quality evaluation and sorting and its use will contribute to the automation of sorting and reduce labor costs and the number of employees required. Among the other advantages of machine vision are non-destructive, accuracy, and consistency. The machinery involved a conveyor, power drive with the inverter, light source, and CCD camera. The software consists of separate algorithms for shape, size, maturity, and defect detection.

After harvesting, the tomato contains mold. These molds are a risk for the tomatoes and need to be removed during the processing. The fruit is emptied into the washer where it is washed and an intensive cleaning result is achieved. Lye solution may also

be used in washing tomatoes to remove drosophila eggs and larvae. It was reported that soaking tomatoes in a 0.5% lye solution for 3 min at 130°F (54°C) was effective for this purpose (Maul et al., 2000).

Blanching is simply scalding of vegetables or fruits in hot water or steam to reduce the enzyme action which causes a change in color, texture, and loss of flavor. Blanching cleanses the surface of dirt and organisms, brightens the color, and helps retard the loss of vitamins. A water blancher comprises a pan with inlet and outlet points. At the bottom of the pan are several outlets that are connected to the circulation system. The circulation system distributes the water back to the pan through injectors. The water is boiled by injected steam or by a heat exchanger. According to research and development centers in Anuradhapura, Sri Lanka, hot water blanching of ripe tomatoes at 60°C for one minute is the most effective pre-treatment in the preservation of the red color of the dehydrated product. As a result, this will be the process condition.

Blending is the process where the natural liquid, vitamins, and minerals are extracted from raw fruits. This process strips away any solid matter from the fruit and what is left is liquid only. Blending is important for food manufacturers so that products are the same every time—consistency is key. If a recipe is not blended properly, the taste, texture, color, and appearance of the final product can all be affected. Spray drying is a method of producing a dry powder from a liquid or slurry by rapidly drying with a hot gas (Seltzer & Settelmeyer, 1949). This is the preferred method of drying foods or pharmaceuticals. Air is the heated drying medium (Baker & Mckenzie, 2014). Criteria like the amount of juice, feed flow rate, and the inlet, outlet air temperatures have a vital role in the output of fruit powders. Fruit powder is hygroscopic and requires drying agents (Siddick & Ganesh, 2016). Although the spray drying method has high capital and installation cost, lower thermal efficiency, and products are susceptible to heat degradation, it has the following advantages over other types of drying such as low labor cost, relatively simple operation and maintenance, reduces transport weight of foods, compliance with the product standards, continuous and easy to control the process, applicable to both heat-sensitive and heat resistant materials, satisfies aseptic/hygienic drying conditions (Mujumdar, 2010).

As a result of the above, spray drying has been chosen as the drying method. Conventional spray dryer processes consist of the following three stages such as atomization of feed into the droplet, spray-air contact and drying droplets, product recovery, and final air treatment (Muzaffar, Nayik, & Kumar, 2015).

Fig. 3. Industrial spray dryer

The tomato powder is then packaged, labeled, and boxed to be distributed for distribution. Aluminum foil will be used to package the tomato powder (Sobowale, Olatidoye, Odunmbaku, & Raji, 2012).

Literature Review

[Hayes, Smith, & Morris, (2010)] examined the quality, composition, standards, and requirements for tomato production. They emphasized that the characteristics of importance in tomato production were color, cohesion, and consistency. They saw a lack of uniformity of quality measures. They stated that, while color may be assessed, no color criteria for tomato production are established.

[Abdulmalik et al., (2014)] investigated the degradation of tomatoes after harvesting by farmers to reduce the enormous wastages that occur. They stated that the farmers are compelled to sell at a low price and do not profit from their produce. They proposed canning as a method of storing tomato paste and spray drying as a more efficient method of producing tomato powder. They also discussed the benefits and drawbacks of hot air convective dehydration with milling and refractive window drying.

[Jayathunge et al., (2012)] developed a suitable drying technique for the manufacturing of dehydrated tomato powder and its utilization in food production. Based on the product color and water activity, they calculated the best conditions for

producing dehydrated tomato powder. Depending on his findings, it was recommended that appropriate packing materials be used based on requirements such as moisture content, water activity, and rehydration ratio.

[Arjenaki, Moghaddam, & Motlagh, (2013)] explored the necessity to develop an effective machine vision system for tomato sorting. They investigated characteristics such as form, size, maturity, and flaws. They sampled around 210 tomato species and evaluated the pictures of each sample, shape, size, and system precisions were approximately 84.4 %, 90.9 %, 94.5 %, and 90%. They obtained these findings with a single line at a rate of 2517tomatoes/hr.

[Maul et al., (2000)] studied the flavor and aroma of ripe tomatoes kept at several temperatures, including 5°C, 10°C, 12.5°C, and 20°C. It was observed that tomato held below 20°C was judged poorer in mature fragrance, sweetness, and flavor by sensory reviewers and greater in sourness (people trained to evaluate the taste, flavor, and texture of food products). They also observed that tomatoes at 5°C had the least developed aroma and flavor.

[Siddick & Ganesh, (2016)] emphasized the importance of spray drying as a post-harvest method for producing tomato powder with a long shelf life. They investigated the application of spray drying technology for tomato cultivars, concentrating on factors such as feed flow rate, input and output temperature, and the addition of barley dextrin using a spray drier in India.

[Adegbola et al., (2012)] investigated the processing of tomatoes into various products, starting with tomato paste. They addressed the massive waste involved in tomato production owing to the nature of the crop, which renders it very prone to degradation after harvest if not stored properly. They outlined the procedures to be followed in the manufacturing of processed tomatoes, as well as the benefits and challenges of processed tomato and information on processes to be followed for the production of tomato produce They also investigated the issues connected with the packaging of finished tomato products, as well as marketing techniques to raise product awareness. They proposed that the establishment of processing plants by investors would produce money and jobs while also assisting in the reduction of waste in Nigeria.

Methodology

The material balance for each process in the plant was drawn up. After then, the energy balance was also carried out. The process flow diagram for the plant was then drawn up using Microsoft Visio. Microsoft Visio was picked because the software has a user-friendly interface. It also has general, flexible drawing tools to accommodate the needs of the user.

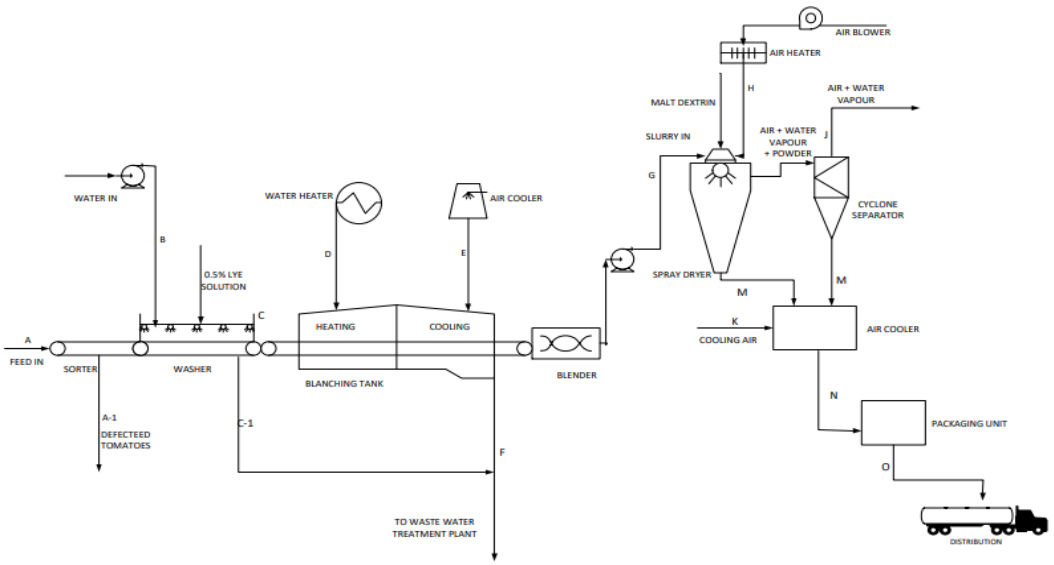

Fig. 4. Process flow diagram of tomato powder processing plant.

The engineering design of the spray dryer involves chemical engineering design as well as mechanical engineering design. The chemical engineering design involves considering the sections that make up the spray dryer and how each section operates. The mechanical engineering design, on the other hand, involves the design of the spray dryer as it is a pressure vessel, specifying the basic design information. The spray dryer is made up of 5 major components namely the drying chamber, hot air supply system, feed supply system, atomizing device, and powder recovery cyclone.

The spray drying chambers are typically field fabricated. They are built on the ground in two sections, the chamber, and the cone, and then a set of structural steel by cranes.

Assuming the vessel has a residence time of 1 minute (60secs). From the Process Flow Diagram, the airflow rate is 3000 m3/hr = 29.4 cfs.

Dryer Volume, Vd = 29.4 * 60 = 1,764 Cu ft

Assuming the straight sides of the vessel is 4 times the diameter and that the vessel has a 600 cone,

1764 = 4D [22D2/7(4)] + 0.866(22) D3/7(12) = 3.3684D3

Therefore, D3 = 523.69

The diameter of the vessel = 8ft

Many factors have to be considered when selecting engineering materials, but for a chemical process plant, the overriding considerations are usually high temperature, strength, and the ability to resist corrosion. Stainless steel is classified as the best choice for the construction of spray dryer for the tomato powder production process due to various advantages it has properties over other construction materials. Properties such as high corrosion resistance, good elasticity modulus, ease of fabrication, available in standard sizes and favorable relative costing rate. To impart corrosion resistance, the chromium content must be above 12%, and the higher the chromium content, the more resistant the alloy to corrosion in oxidizing conditions. Nickel is added to enhance corrosion defiance in non-oxidizing environments.

The hot air will be passed in a co-current manner which implies the same direction as the sprayed liquid atomizer. With the co-current flow, particles spend less time in the system and the particle separator. This flow generally allows the system to operate more efficiently. In this sequence, the atomized blobs penetrating the dryer are in touch with the hot inlet air, but the temperature is kept below due to an increased proportion of evaporation taking place and much at the wet-bulb temperature. It also consists of four main parts namely; supply fan, air filters, air heaters, air dispensers.

The feed supply system has a flowrate of 9,896Kg/hr, a temperature of 35℃, and a pressure of 1.0132Bar. The rotary atomizer considered the most flexible atomizing device is suitable for a wide range of products. The rotary atomizer is suitable for an extensive range of products. Rotary atomizers produce a thin particle size dispersion and circulation of powder. They are preferred because they handle high feed rates without clogging; formation of uniform size particles, low-pressure operation; high efficiency. They can operate at low pressure than that required in hydraulic and pneumatic nozzle atomizers.

Fig. 5. Rotary atomizer

The type of equipment to be used is a **cyclone with a tangential entry.** These types of cyclones have a distinctive and easily recognized form and are widely used in food processing industries. In the agricultural processing industry, 2D2D cyclone designs are the most commonly used abatement devices for particulate matter control.

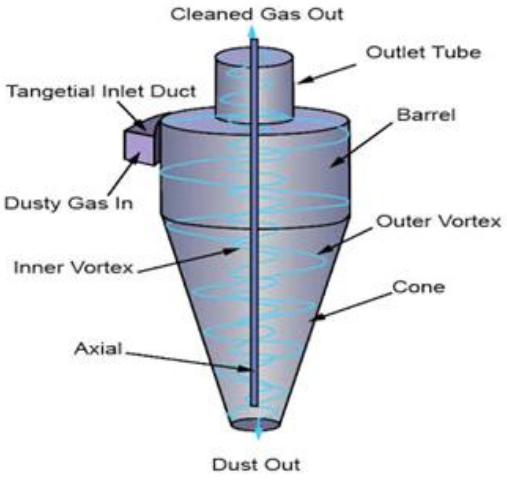

Fig. 6. Cyclone with a tangential entry

The piping and instrumentation diagram (P&ID) for the whole plant was done first, followed by a detailed one for the spray dryer. The P&ID was drawn up using Microsoft Visio.

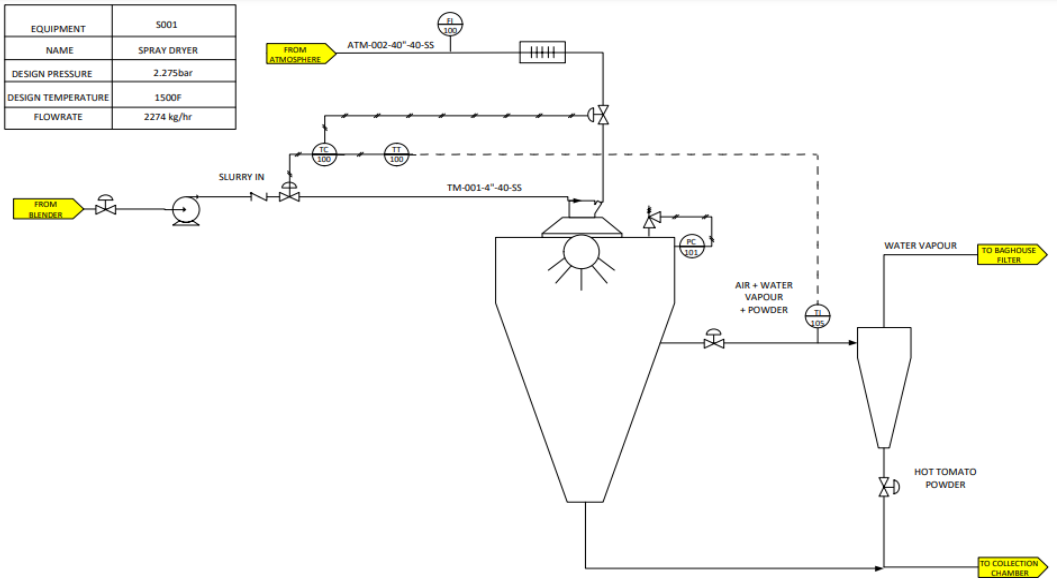

Fig. 7. Piping and instrumentation diagram of spray dryer

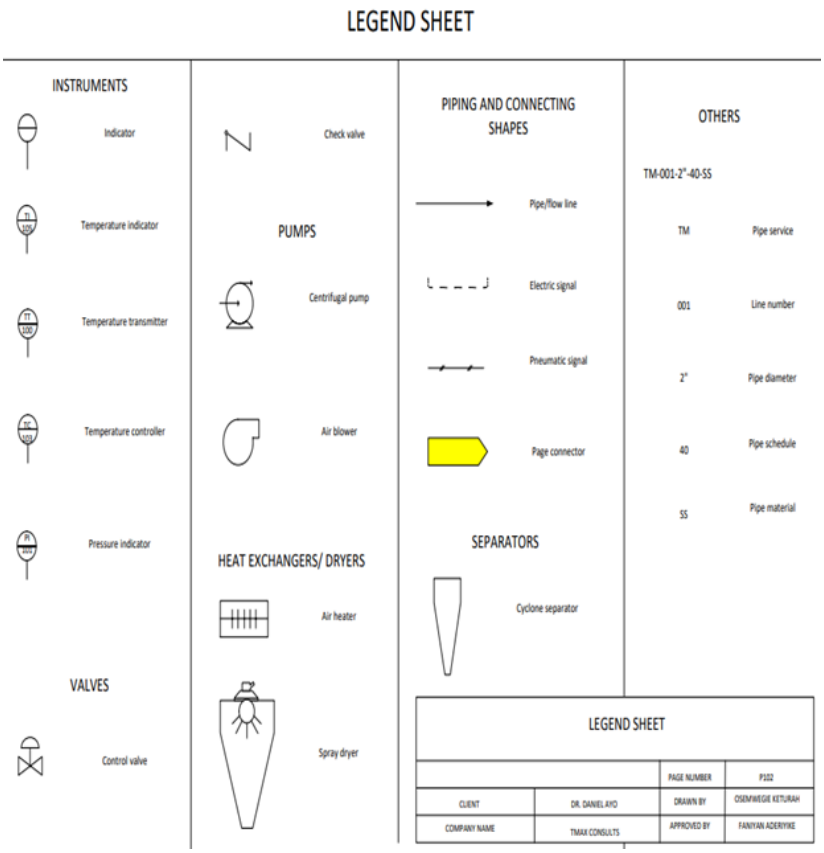

Fig. 8. Legend sheet of elements in P&ID of the spray dryer

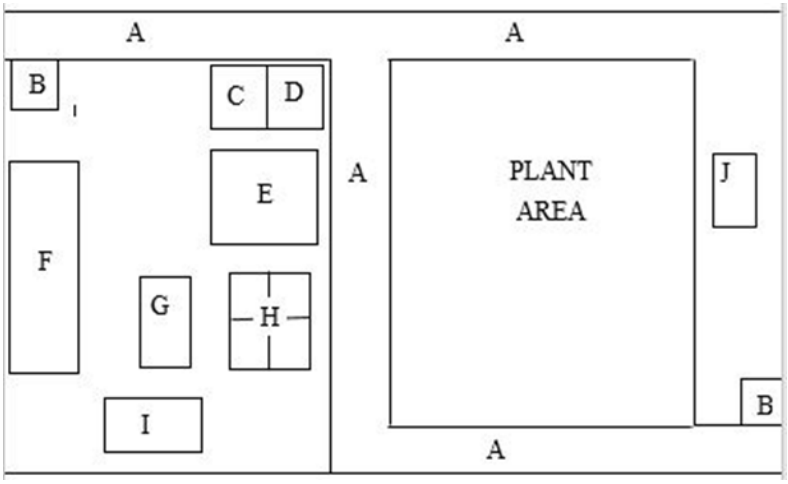

Fig. 9. Site layout

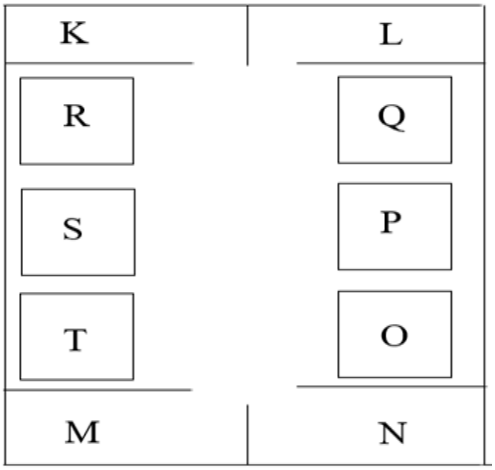

Fig. 10. Plant layout

LEGEND

A – ROADS	K – CONTROL ROOM
B – SECURITY	L – QUALITY CONTROL LAB
C – MAINTENANCE WORKSHOP	M – RECEPTION OF TOMATOES
D – BOILER	N – WEIGHING OF TOMATOES
E – STOREHOUSE	O – CHOOSING & SORTING
F – ADMINISTRATIVE OFFICE	P – WASHING
G – CANTEEN	Q – BLANCHING
H – MALE/FEMALE CHANGING ROOM & TOILET	R – BLENDING
I – CLINIC	S – SPRAY DRYING
J – TRUCK CAR PARK	T – PACKAGING & LABELLING

Fig. 11. Legend sheet showing the representation of each figure in site & plant layout

Roads: If a site is flat, the topography may not influence the location and layout of the building, but on a sloping site, the topography is likely to be a significant factor. The configuration of the road linking where the plant is located (Oluyole LG) opposite Oriental food industrial limited is a flat surface with no slope, which enables easy road access from the main gate to the expressway.

Security: The security building is positioned at the entrance of the gate facing an east flat surface side to have reduced sunlight by day and to have easy access to what is visible to the surrounding. It is also important to monitor the goings of workers and products.

Maintenance workshop: The distance between where this department is located and where the plant area building is situated is for easy repair, correction, and maintenance of faulty or damaged parts in the spray dryer and other utilities in the work area.

Storehouse: This is a building where the final product is stored. It is located close to the plant area because of the easy movement of products from the point of production to the warehouse.

Administrative office: A rule of thumb applies to this which is "minimize the use of highly visible large retaining walls." If they are over a meter, they should be stepped and landscaped.

Canteen: In a site layout, this building is usually close to the administrative block and the clinic. A consideration in site layout is that the cafeteria should not be far from the clinic.

Discussion

The site selected for our plant is in the Oluyole local government area in Oyo State, Nigeria. It is going to be adjacent to Oriental Foods Industries Limited. It has an area of 4,000km^2 and a population of 202,725 as of 2006. It shares boundaries with four native governments, Ibadan South West, Ibadan South East, Ona Ara, and Ido ingenious government all within the Ibadan city. The Urban segment of the native government constitutes the Lagos/Ibadan Express Road, Old Lagos Road, and New Garage, where giant companies were situated. Companies like; British America Tobacco (BAT), ROM Oil, Black-Hors plastic company, Jubaili Agro-Limited, KAMAR industries. The populaces in Oluyole local government are farmers with arable soil for crops like cocoa, cashews, orange, mango, etc.

The site was selected due to its proximity to the Lagos-Ibadan expressway and the availability of labor and land. It is approximately 150km from Lagos by the most direct route and 659km from Abuja, the Federal Capital Territory (FCT). It has a well-developed road network.

The economic construction and efficient operation of a processing unit will depend on how well the plant and equipment specified on the process flowsheet are laid out and the following are the factors that affected the site selection of the tomato powder production plant in the Oluyole local government in Oyo State:

Availability of raw materials: Unrestricted and regular supply of raw materials from Kaduna or Kano State. Since tomatoes are weight loss and cannot be preserved for a long time, adequate transport facilities are ensured to supply the raw materials.

Transport facilities: The transport of raw materials and products to and from the plant is also a major factor. The site being located in Ibadan makes road transportation easier, and road transportation will be majorly used for conveying raw

materials and finished products. Air transport is considered for the movement of personnel and equipment.

Availability of labor: Labor will be needed for the construction of the plant and its operation. For a location like Ibadan, there is an adequate pool of unskilled labor available locally and labor suitable for training to operate the plant.

Utilities (services): The plant is located near a source of water of suitable quality. The process water can be drawn from a well. There is also a cheap source of power supply and a relatively affordable source of fuel.

Nearness to Market: Market changes influence the enterprise of a manufacturing unit. The site, located in Ibadan is near markets like Ogun and Lagos, which is the center of commerce. Tomato powder is also highly needed in Lagos because of its high shelf life compared to raw tomatoes and tomato paste.

Environmental impact and effluent disposal: All industrial processes produce waste products, and full consideration must be given to the cost of disposal. The disposal of toxic and harmful effluents will be covered by local regulations, and the appropriate authorities must be consulted during the initial site survey to determine the standards that must be met. The environmental laws guiding Oluyole, Ibadan is the same as the ones guiding Nigeria.

Soil, climate, and topography: The climate conditions of Ibadan are stable to set up an industrial plant and to help workers to be efficient. Ibadan is not subject to adverse climatic conditions such as earthquakes; hence it is favorable to set up a site. The Oluyole area in Ibadan has a good topography (its plain land, it's not hilly) which makes it suitable for industries. Oluyole, Ibadan area has sufficient suitable land available for the proposed plant and future expansion. The land is flat, well-drained, and has suitable load-bearing characteristics.

Government, policies, and regulations: Industrial Development and Regulation Act of 1951 laid down certain rules, regulations, and formalities to be complied with before setting up an industrial unit. Prior permission and license are necessary under this act before setting up a new industrial unit. All these will be taken into considerations and necessary procedures will be followed before setting up the plant at Oluyole, Ibadan.

Local community considerations: The intended plant must be satisfactory to the residents of the Oluwole local government. Safety measures are mandatory in the positioning of the plant, so it doesn't serve as a threat to the lives and properties of communities within the vicinity. Plant proximity brings about productivity and resources to the community. The traditional institution of Oluyole Local Government comprises the Baale which is under the control of Olubadan of Ibadan Land. There is a strong attachment between the traditional rulers and other elders in their community. As a result, community conflict is very unlikely after the plant begins

operations. The corporate social responsibility of the industry to the community will be duly performed.

Nearness to adequate banking and credit facilities: For the productive management of the business, banks should be located in the community to offer services to the people, which creates economic growth. Nearness to banks and financial institutions is an important consideration for the location of an industrial unit.

Maintenance: Equipment that requires dismantling for maintenance, such as compressors and large pumps around the spray dryer unit, should be placed undercover.

Safety: Blast walls may be needed to isolate potentially hazardous equipment and confine the effects of an explosion. At least two escape routes for operators must be provided for each level in process buildings (Esguerra, Rolle, & Rapusas, 2013).

Plant expansion: Equipment should be located so that it can be conveniently tied in with any future expansion of the process.

RECOMMENDATION

Tomatoes are a highly nutritious food ingredient used by all Nigerian tribes in the preparation of numerous foods. Its strong demand has made it a very profitable business for individuals, and it is this business that ensures a good return on investment.

Given the inadequacy of local tomato powder processing facilities, I recommend resources be made available for the fabrication and construction of tomato processing facilities. Executing this will help barely noticed tomato products supplied locally to attract foreign currency income and promote domestic food sufficiency.

There needs to be a greater emphasis on tomato processing in Nigeria due to the increase in the quality of life in cities and the rapid urbanization of rural areas, which contributes to high export potential and a positive return on investment.

The lack of storage and processing facilities and under-developed marketing channels resulted in 50% of tomato production is destroyed. Converting tomato to powder will reduce the high percentage of tomato that generally rots away, and investors can benefit from trade liberalization in West Africa (ECOWAS) to commercialize their products.

Conclusion

The production of tomato powder will respond to the challenge of tomato deterioration in Oyo State and elsewhere. This measure was intended to reduce the deterioration of tomatoes in Nigeria and to reduce the import of processed tomato products.

Through a series of investigations, the most feasible way to dry Tomato powder on an industrial scale proved to be the spray-drying method because the process was very rapid, required low labor cost, and was relatively simple to operate and sustain.

Based on the results obtained during the design of the plant, it was recommended that investments be made for the establishment of this plant. More investment should be made in growing raw tomatoes, which will help reduce the cost of raw tomato material (Adegbola et al., 2012).

The feasibility study carried out indicated that it is technically feasible to construct the plant because the equipment is available, the process route is feasible, and the process conditions are attainable. It was shown from the safety analysis that the safety operability concerns of the plant can be taken care of by proper engineering control, environmental control, and personal protection.

In addition, the economic analysis shows that the plant is profitable.

Acknowledgment

The author is grateful to the academic and non-teaching staff of the Department of Chemical and Petroleum Engineering, Faculty of Engineering, University of Lagos, Lagos, for the technical assistance received in the execution of this project. The moral and financial support received from Prof. O.A.T. Ebuehi, and the expertise and knowledge received from Dr. (Mrs.) O.M. Ebuehi, Osaruyi Ebuehi, Eseosa Ebuehi, Ayomide Adedeji, Dr. Daniel Ayo, Faniyan Aderiyike, Osemwegie Keturah, Shobowale Ahmed, Ezeka Francis, Olawale Oyindamola, Nwankwo Dave, and Adeyemi Boluwatife in the Department of Chemical and Petroleum Engineering.

References

[1] B. Jarimopas & N. Jaisin. (2008). "An experimental machine vision system for sorting sweet tamarind," *J. Food Eng.*, vol. 89, no. 3, pp. 291–297. DOI: 10.1016/j.jfoodeng.2008.05.007.

[2] C. G. J. Baker & K. A. Mckenzie. (2014). "Drying Technology : An Energy Consumption of Industrial Spray Dryers," no. October, pp. 37–41. DOI: 10.1081/DRT-200047665.

[3] E. B. Esguerra, R. Rolle, & R. S. Rapusas. (2013). "Post-harvest management of tomato for quality and safety assurance," *Handbook of Food and Agriculture Organization of the United Nations,* pp. 8-15.

[4] E. Seltzer & J. T. Settelmeyer. (1949). "Spray Drying of Foods," *Adv. Food Res.*, vol. 2, no. C, pp. 399–520. DOI: 10.1016/S0065-2628(08)60048-8.

[5] F. Maul, S.A. Sargent, C.A. Sims, E.A. & Baldwin. (2000). "Tomato Flavor and Aroma Quality," *Food Sci. Sens. Nutr. Qual. Food*, vol. 65, no. 7, pp. 1228–1237.

[6] J. A. Adegbola, E. A. Adu, U. D. Anugwom, D. T. Ishola, & A. A. Bodunde. (2012). "Investment opportunities in tomato processing in Kano, Northern Nigeria.," *Glob. Adv. Res. J. Agric. Sci. (ISSN 2315)*, vol. 1, no. 10, pp. 288–297.

[7] K. Jayathunge, R. Kapilarathne, B. Thilakarathne, K. Palipane, & P.Prasanna. (2012). "Development of a methodology for the production of dehydrated tomato powder and study the acceptability of the product," *J. Agric. Technol.*, vol. 8, no. 2, pp. 765–773.

[8] K. Muzaffar. (2015). "Stickiness Problem Associated with Spray Drying of Sugar and Acid Rich Foods: A Mini-Review," *J. Nutr. Food Sci.*, vol. s12, pp. 11–13. DOI: 10.4172/2155-9600.s12-003.

[9] O. Abdulmalik, I., M. C. Amonye, A. O, & P. O. Umeanuka. (2014). "Appropriate Technology for Tomato Powder Production," *Int. J. Eng. Invent.*, vol. 3, no. 8, pp. 29–34.

[10] O. O. Arjenaki, P. A. Moghaddam, & A. M. Motlagh. (2013). "Online tomato sorting based on shape, maturity, size, and surface defects using machine vision," *Turkish J. Agric. For.*, vol. 37, no. 1, pp. 62–68. DOI: 10.3906/tar-1201-10.

[11] S. A. Siddick & S. Ganesh. (2016). "Spray drying technology for producing fruit powders from tomatoes and tamarillo," *Acta Hortic.*, vol. 1120, pp. 343–348.DOI: 10.17660/ActaHortic.2016.1120.52.

[12] S. Mujumdar. (2020). "Handbook of Industrial Drying," *Handb. Ind. Dry.* DOI: 10.1201/9780429289774.

[13] S. S. Sobowale, O. P. Olatidoye, L. A. Odunmbaku, & O. H. Raji. (2012). "A Comparative Study on Physicochemical and Rheological Properties of Imported Tomato Paste in Nigeria," *Sustain. Agric. Res.*, vol. 1, no. 2, p. 51. DOI: 10.5539/sar.v1n2p51.

[14] U. F. Sylvanus & I. E. Victor. (2012). "Impact of tomato processing and wastages in Kano, Nigeria," "International Journal Of Social Sciences & Interdisciplinary Research," *Int. J. Soc. Sci.*, vol. 1, no. 3, pp. 297–310.

[15] W. A. Hayes, P. G. Smith, & E. J. Morris. (2010). "Critical Reviews in Food Science and Nutrition The Production and Quality of Tomato Concentrates," *Crit. Rev. Food Sci. Nutr.*, vol. 38, no. 7, pp. 537–564.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Detection of Anomalies in the Computer Network Behaviour

Danijela Protić

MSc, Centre of Applied Mathematics and Electronics, Belgrade, Serbia

Miomir Stanković

PhD, Mathematical Institute, Serbian Academy of Science and Arts, Belgrade, Serbia

Abstract

The goal of anomaly-based intrusion detection is to build a system which monitors computer network behaviour and generates alerts if either a known attack or an anomaly is detected. Anomaly-based intrusion detection system detects intrusions based on a reference model which identifies normal behaviour of the computer network and flags an anomaly. Basic challenges in anomaly-based detection are difficulties to identify a 'normal' network behaviour and complexity of the dataset needed to train the intrusion detection system. Supervised machine learning can be used to train the binary classifiers in order to recognize the notion of normality. In this paper we present an algorithm for feature selection and instances normalization which reduces the Kyoto 2006+ dataset in order to increase accuracy and decrease time for training, testing and validating intrusion detection systems based on five models: k-Nearest Neighbour (k-NN), weighted k-NN (wk-NN), Support Vector Machine (SVM), Decision Tree, and Feedforward Neural Network (FNN).

Keywords: intrusion detection, Kyoto 2006+, k-NN, wk-NN, SVM, decision tree, FNN

Introduction

Intrusion detection systems (IDSs) protect computer networks from malicious activities which compromise network security and affect the confidentiality, integrity and availability of the data. IDSs can be grouped into the signature-based, anomaly-based, and hybrid (Ganpathy et al., 2015, 44-50).The basic idea of signature-based detection is to represent an attack in the form of pattern in such a way that any known attack and its variation can be detected. The main disadvantage of this approach is difficulty for detecting unknown attacks. Anomaly-based IDS detects changes in the network behaviour. The goal of anomaly-based detection is to build a statistical model

that describes the normal behaviour of the computer network and then looks for activities which differ from the created model. It detects both intrusions and/or misuse, and classifies them as either 'normal' or 'anomaly'. The biggest challenge in anomaly-based detection is to identify what is considered normal network behaviour. Machine learning (ML) models can be trained as binary classifiers in order to recognize the notion of normality. In a supervised ML, the data have to be collected over a period of time to create a model of normal behaviour of users, hosts, and networks.

A number of records needed for training the complex computer networks can be large, which makes evaluation of the IDS computationally expensive since the processing time and memory usage rise with a size of the dataset. In anomaly-based detection some of the recorded data can be discarded to decrease time needed for training and increase accuracy of IDSs. This paper presents one pre-processing technique applied to reduce the size of the Kyoto 2006+ dataset. The proposed algorithm cuts off all categorical features, features which are intended for further analyses of the evaluated models, and features containing instances which cannot be normalized into the range [-1, 1], excluding the feature 'Label' that flags either normal network behaviour or an anomaly (Ramasamy & Rani, 2018, 1060-1067). After the pre-processing, of 24 features describing each instance nine features left. In this paper we present results on the accuracy and computation time testing for five supervised learning models: k-Nearest Neighbour (k-NN), weighted k-NN (wk-NN), Support Vector Machine (SVM), Decision Tree and Feedforward Neural Network (FNN).

1. The Kyoto 2006+ Dataset

The Kyoto 2006+ dataset is built on the three years (November 2006 to August 2009) of the real network traffic data, collected on five different computer networks inside and outside the Kyoto University. The data set is designed to provide evaluation of the network-based intrusion detection systems (NIDS). It consists of 14 statistical features derived from the KDD Cup '99 dataset (1999) and 10 additional features which can be used for evaluation and further analyses of NIDS (Protic, 2018, 580-595). The Kyoto 2006+ dataset is captured using honeypots, darknet sensors, e-mail servers, web crawlers and other intrusion detection systems (Sing, 2014, 31-35). During the observation period, more than 50 million sessions of normal traffic, over 43 million sessions of known attacks and almost 426 thousand sessions of unknown attack were recorded (Song et al, 2011). Of the 41 features derived from the KDD Cup '99 dataset, authors discarded redundant data and content features which are not suitable for NIDS (See Table 1).

Table 1 First 14 features from the KDD Cup '99 dataset

No	Feature	Description
1	Duration	The length of the connection (seconds).
2	Service	The connection's server type (dns, ssh, other).

3	Source bytes	The number of data bytes sent by the source IP address.
4	Destination bytes	The number of data bytes sent by the destination IP address.
5	Count	The number of connections whose source IP address and destination IP address are the same to those of the current connection in the past two seconds.
6	Same_srv_rate	% of connections to the same service in the Count feature.
7	Serror_rate	% of connections that have 'SYN' errors in Count feature.
8	Srv_serror_rate	% of connections that have 'SYN' errors in Srv_count (% of connections whose service type is the same to that of the current connections in the past two seconds) feature.
9	Dst_host_count	Among the past 100 connections whose destination IP address is the same to that of the current connection, the number of connections whose source IP address is also the same to that of the current connection.
10	Dst_host_srv_count	Among the past 100 connections whose destination IP address is the same to that of the current connection, the number of connections whose service type is also the same to that of the current connection.
11	Dst_host_same_src_port_rate	% of connections whose source port is the same to that of the current connection in Dst_host_count feature.
12	Dst_host_serror_rate	% of connections that have 'SYN' errors in Dst_host_count feature.
13	Dst_host_srv_serror_rate	% of connections that have 'SYN' errors in Dst_host_srv_count feature.
14	Flag	The state of the connection at the time of connection was written (tcp, udp).

(Source: Song et al., 2011)

In this way, the 14 features left consisted of one categorical feature 'Flag' and 13 continuous features. Moreover, authors extracted 10 additional features: 'Label' which indicated normal traffic or attacks, four features describing source and destination addresses and port numbers, two features describing start time and duration of the session, and three features for IDS, malware and Ashula detection (See Table 2).

Table 2. Additional features

No	Feature	Description
1	IDS_detection	Reflects if IDS triggered an alert for the connection; '0' means any alerts were not triggered and an arabic numeral means the different kind of alerts. Parenthesis indicates the number of the same alert.
2	Malware_detection	Indicates if malware, also known as malicious software, was observed at the connection; '0' means no malware was observed, and string indicates the corresponding malware observed at the connection. Parenthesis indicates the number of the same malware.
3	Ashula_detection	Means if shellcodes and exploit codes were used in the connection; '0' means no shellcode nor exploit code were observed, and an arabic numeral means the different kinds of the shellcodes or exploit codes. Parenthesis indicates the number of the same shellcode or exploit code
4	Label	Indicates whether the session was attack or not; '1' means normal. '-1' means known attack was observed in the session, and '-2' means unknown attack was observed in the session.
5	Source_IP_Address	Means source IP address used in the session. The original IP address on IPv4 was sanitized to one of the Unique Local IPv6 Unicast Addresses. Also, the same private IP addresses are only valid in the same month; if two private IP addresses are the same within the same month, it means their IP addresses on IPv4 were also the same, otherwise are different.
6	Source_Port_Number	Indicates the source port number used in the session.
7	Destination_IP_Address	It was also sanitized.
8	Destination_Port_Number	Indicates the destination port number used in the session.
9	Start_Time	Indicates when the session was started.
10	Duration	Indicates how long the session was being established.

(Source: Song et al., 2011)

2. Feature Selection

One of the major issues associated with the Kyoto 2006+ dataset is its size. Features selection reduces the data dimensionality by determining whether a feature is relevant or not for evaluation of the NIDS model. Using effective features in designing classifiers not only reduce the dataset but also improve performances of the classifier (Jayakumar, Revathi & Karpagam, 2015, 728-734).

In this paper we present two-step pre-processing algorithm for feature selection given as follows:

Step 1: Discard all categorical features and all features which are intended for further analyses, excluding feature 'Label'.

Step 2: Cut features containing instances which cannot be normalized into the range [-1, 1].

Of the 24 features of the Kyoto 2006+ dataset, 17 features are discarded after the first algorithm step. Nine features (5-13) left after the pre-processing is done. These features are normalized to fall into the range [-1, 1] by applying the hyperbolic tangent function given with Eq. (1):

$$\tanh(n) = \frac{2}{1+e^{-2n}} - 1 \qquad (1)$$

In this way the values of instances are scaled so that the effect of one feature cannot dominate the others. Moreover, normalized instances speed up FNN. Network training is more efficient when this pre-processing is performed on inputs. If the inputs are greater than 3 ($e^{-3} \approx 0.05$) sigmoid functions, which are used in the hidden network layer, become essentially saturated. If this happens at the beginning of the training process the gradients will be very small and the network training will be slow.

3. Machine Learning Models

Supervised ML algorithms use the known dataset to evaluate a model that generates prediction of unknown data. Assume that all data points belong either to the class 'normal' or class 'anomaly' (binary classification). Than each training data point x_i, from a vector of d-dimensional feature space, can be labelled by y_i as follows (see Eq. (2)):

$$y_i = \begin{cases} y_n, & x_i \in class_{normal} \\ y_a, & x_i \in class_{anomaly} \end{cases} \qquad (2)$$

The training dataset is denoted as follows $\{(x_i, y_i),\ i = 1, \cdots, k\}$.

This paper presents five models used for binary classification: k-NN, wk-NN (Hechenbichler & Schliep, 2004), SVM (Burgess, 1998, 283-298), Decision Tree (Sebastiani, 2002, 13) and FNN (Protic & Milosavljevic, 2006, 643-646).

3.1 k-Nearest Neighbour

k-NN is nonparametric method where new observation is placed into the class of observation from the learning set (Hechenbichler & Schliep, 2004). In this paper we present results on k-NN based prediction of unknown instances which finds the largest similarity of instances based on the Euclidean distance measure (Eq. (3)):

$$d(x_i, y_i) = \sqrt{\sum_{s=1}^{p} (x_{is} - y_{is})^2} \qquad (3)$$

3.2 Weighted k-Nearest Neighbour

The main idea of wk-NN is to extend the k-NN method in so that the observations within the learning set, which are particularly close to the new observation, should get a higher weight in the decision than such neighbours which are more distant one from that observation (Hechenbichler & Schliep, 2004). To reach this aim the distances have to be transformed into the similarity measures (Eq. (4)), which can be used as weights (Eq. (5)):

$$dist = \sqrt{\sum_{i=1}^{p} (x_i - y_i)} \qquad (4)$$

$$w = \frac{1}{dist^2} \qquad (5)$$

3.3 Support Vector Machine

The goal of SVM algorithm is to find a hyperplane that distinctly classifies the data points into various classes (Burgess, 1998, 283-298). To separate instances into the classes 'normal' or 'anomaly' the algorithm finds a linear hyperplane that has the maximum distance ρ=2/||w|| between data points of both classes. For the training set

$\mathbf{x}(x_i, y_i), x_i \in R^d, y_i \in \{-1,1\}$, the discriminant function takes the form (Eq. (6)):

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b = \sum_i w_i x_i + b \qquad (6)$$

where $\boldsymbol{w}$ is normal vector to the hyperplane that can be determined as follows (Eq. (7)):

$$\begin{aligned} & x_i \mathbf{w} + b \geq +1, && y_i = +1 \\ & x_i \mathbf{w} + b \leq -1 && y_i = -1 \\ & y_i (x_i \mathbf{w} + b) - 1 \geq 0, && \forall i,\ 1 \leq i \leq n \end{aligned} \qquad (7)$$

The idea is to find min||w|| which maximizes the distance ρ.

3.4 Decision Tree

Decision Tree prediction is based on the principle of recursive partitioning by monitoring decisions from the root to the last node (Sebastiani, 2002, 13). It is one of the graph-like algorithms which use branching methods to illustrate every possible outcome of the decisions, where nodes represent features, links represent decision rules and leafs represent the outcomes. In the experiments we applied the Iterative Dichotomy 3 algorithm (ID3) (Colin, 1996, 107-110) which calculates entropy and information gain to build a tree. Entropy is a measure which controls how the tree decides to split the data. If the target feature can take on k different values then entropy of S relative to this k-wise classification can be calculated as follows (Eq. (8)):

$$Entropy(S) = -\sum_{i=1}^{k} p_i \log_2(p_i) \quad (8)$$

where p_i is the proportion S belonging to class i. Information gain represents expected reduction in entropy based on the decrease in entropy after the dataset is split on the feature. The feature with highest information gain will split first. Information gain can be calculated with the formula (Eq. (9)):

$$Gain(S,A) = Entropy(S) - \sum_{v \in avlues(A)} \frac{|S_v|}{|S|} \cdot Entropy(S_v) \quad (9)$$

Gain(S,A) of a feature A relative to a collection of examples S provides information about the target function value, given the value of some other feature A (when A splits the set S into the subsets S_v).

3.5 Feedforward Neural Network

FNN is based on the back-propagation algorithm. The nonlinear transfer function of the FNN is given with Eq. (10):

$$y_i(\mathbf{w},\mathbf{W}) = F_i\left(\sum_{j=1}^{q} W_{ij} f_j\left(\sum_{l=1}^{m} w_{jl} x_l + w_{j0}\right) + W_{i0}\right) \quad (10)$$

where x_l are inputs, y_i are outputs, **w** and **W** are weight matrices, f_j and F_i denote transfer functions of hidden and output layers, *m* represents the number of inputs, *q* represents the number of outputs, and w_{j0} and W_{i0} denote biases (Protic & Milosavljevic, 2006, 643-646). The objective of FNN presented in this paper is to minimize output error in accordance with the Levenberg-Marquardt algorithm (Protic, 2015, 11-28). It should be pointed out that the FNN has to be large enough to improve network generalization and provide an adequate fit.

4. Results

In our experiments, performances of the models are measured based on accuracy (ACC), which represents the ratio of number of instances correctly classified to the total number of instances given with Eq. (11) (Ambedkar & Babu, 2015, 25-29):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (11)$$

where true positive (TP), true negative (TN), false positive (FP) and false negative (FN) denote detected network behaviour as follows:

TP – 'anomaly' is detected as 'anomaly',

TN – 'normal' is detected as 'normal',

FP – 'normal' is detected as 'anomaly' and

FN – 'anomaly' is detected as 'normal'.

In our previous work we have presented results based on normalized and not-normalized dataset of instances and four ML models, namely k-NN, wk-NN, SVM and Decision Tree (Protic & Stankovic, 2018, 43-48). Here we present results of the experiments conducted to the normalized instances and five models: k-NN, wk-NN, SVM, Decision Tree and FNN. Accuracy of the models and the corresponding computation time (sum of training, testing and validation time) are given in Table 3.

Table 3 Accuracy and computation time

No	Size of the dataset	Accuracy [%] Comp. time [s]	FNN	k-NN	wk-NN	SVM	Decision Tree
1	158572	Accuracy [%]	98.8	98.3	98.4	98.1	97.2
		Comp. time [s]	26	275.72	277.32	449.35	3.8452
2	129651	Accuracy [%]	97.67	91.8	91.8	98.3	97.3
		Comp. time [s]	20	175.84	173.32	254.3	3.3104
3	128740	Accuracy [%]	98.32	98.2	98.1	97.8	97.2
		Comp. time [s]	8	193.82	194.81	280.82	3.3033
4	136625	Accuracy [%]	99.21	99.3	99.4	99.1	98.3
		Comp. time [s]	20	194.83	194.23	217.32	8.3169
5	90129	Accuracy [%]	98.99	99.0	99.1	99.0	98.4
		Comp. time [s]	11	101.28	101.753	86.283	2.2308

6	93999	Accuracy [%]	98.12	96.5	96.5	98.0	97.5
		Comp. time [s]	7	109.25	108.77	111.83	2.2613
7	81807	Accuracy [%]	98.3	98.8	98.8	97.9	98.9
		Comp. time [s]	10	91.25	91.26	227.28	2..2615
8	57278	Accuracy [%]	99.14	99.36	99.3	99.2	99.3
		Comp. time [s]	2	42.704	43.235	33.754	1.743
9	58317	Accuracy [%]	98.97	99.1	99.2	99.1	98.9
		Comp. time [s]	3	31.714	31.738	34.234	1.7482
10	57278	Accuracy [%]	99.2	99.4	99.5	99.2	99.4
		Comp. time [s]	2	43.734	43.272	30.239	1.2901

The experiments are conducted on 10-days records from the Kyoto 2006+ dataset (991.395 instances in total). All models are trained so that out of the total number of randomly selected instances 70% are used for training, 15% for testing and 15% for validation of the models. Experiments are performed using Intel(R) Core(TM) i7-2620M CPU 2.70GHz processor with 16GB RAM Installed Memory.

Results show the highest accuracy (99.5%) of wk-NN model. The results also point to high accuracy for k-NN model (99.36%). Computation time for evaluating Decision Tree model is significantly shorter comparing to the computation time for other model's evaluation, except for the FNN. Number of parameters in the network structure with nine inputs, one hidden-layer and one output is large enough to provide an adequate fit. The highest accuracy of FNN (99.21%) is achieved when the network is trained with the largest subset (136.625 instances in total). As expected, the time period needed for network learning is longer than for the networks trained with the smaller datasets (20s). Time period of training, testing and validating the FNN is significantly shorter than k-NN, wk-NN and SVM, and fall into the range [2s, 26s]. The results also show that the SVM model has a lower accuracy and longer computation time comparing to the other models.

Conclusion

Anomaly-based intrusion detection systems recognize deviations from normal computer network behaviour. A main challenge in anomaly-based detection is to determine the normal network behaviour and flag the anomaly. Machine learning models can be trained to classify the data into categories 'normal' or 'anomaly'. In supervised machine learning the data have to be recorded over a period of time to create a model of normal behaviour. This process can take significant time and may be computationally expensive for complex computer networks. To decrease the period of learning and increase the accuracy of the model the number of features can be significantly reduced. In this paper, we present one pre-processing technique based on the feature selection. A subset containing nine features and normalized instances is used for evaluation of IDSs based on k-NN, wk-NN, SVM, Decision Tree and FNN models. The results show that the highest accuracy gives the wk-NN model, while the shortest computation time has Decision Tree model. Overall, FNN shows higher accuracy and shorter computation time, compared to the other models. In our further work we will present the results of experiments conducted on hybrid model based on wk-NN and FNN which detects variation in the decision on detected anomalies.

References

[1] Ambedkar, Ch. & Babu V. K. (2015). Detection of Probe Attacks Using Machine Learning Techniques. *International Journal of Research Studies in Computer Science and Engineering*. *2*(3), 25-29.

[2] Burgess, M. (1998). Computer Immunology. *12th USENIX Conference on System Administration*, Boston, MA, USA, 06-11 December 1998, 283-298.

[3] [Dataset] The UCI KDD Archive Information and Computer Science University of California, Irvine. Last modified: October 28, 1999. Retrieved from http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[4] Ganpathy, S., Vijayakumar, P., Yogesh, P. & Kannan, A. (2015). An Intelligent CRF Based Feature Selection for Effective Intrusion Detection. *The International Arab Journal of Information Technology*, published online March 8, 44-50.

[5] Hechenbichler, K. & Schliep, K. (2004). Weighted *k*-Nearest-Neighbor Techniques and Ordinal Classification. *Sonderforschungsbereich* 386, Paper 399 (2004). Retrieved from https://epub.ub.uni-muenchen.de/1769/1/paper_399.pdf

[6] Jayakumar, K., Revathi, T. & Karpagam, S. (2015). Intrusion Detection using Artificial Neural Networks with Best Set of Features. *The International Arab Journal of Information Technology*. *12*(6A), 728-734.

[7] Protic, D. (2015). Feedforward neural networks: The Levenberg-Marquardt optimization and the optimal brain surgeon pruning. *Military Technical Courier*. *3*(63), 11-28.

[8] Protic, D. (2018). Review of KDD Cup '99, NSL-KDD and Kyoto 2006+ datasets. *Vojnotehnički glasnik/Military Technical Courier*, *66*(3), 580-596. DOI: 10.5937/vojteh-16670.

[9] Protic, D. & Milosavljevic, M. (2006). NNARX model of speech signal generating system: test error subject to modeling mode selection. *25th International conference on microelectronics*, 643-646.

[10] Protic, D. & Stankovic, M. (2018). Anomaly-Based Intrusion Detection: Feature Selection and Normalization Influence to the Machine Learning Models Accuracy. *European Journal of Formal Science and Engineering*. *1*(3), 43-48. ISSN 2601-8675 (Print), ISSN 2601-8683 (Online).

[11] Ramasamy, R. & Rani, S. (2018). Modified Binary Bat Algorithm for Feature Selection in Unsupervised Learning. *The International Arab Journal of Information Technology. 15*(6), 1060-1067.

[12] Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Comput.Surv. 34*(1),1-47.

[13] Singh S., Khan Z. & Saxena K. (2014). Intrusion Detection Based on Artificial Intelligence Technique. *International Journal of Computer Science Trends and Technology. 2*(4), pp. 31-35.

[14] Song, J., Takakura, H., Okabe, Y., Eto, M., Inoue, D. & Nakao, K. (2011). Statistical Analysis of Honeypot Data and Building of Kyoto 2006+ Dataset for NIDS Evaluation. In: *Proc.1st Work-shop on Building Anal. Datasets and Gathering Experience Returns for Security*. Salzburg. April 10-13, 29-36.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Design and Application of Optimally-Tuned Variable Parameters PID Controller for Nonlinear Engineering Systems

Magdy A. S. Aboelela

Faculty of Engineering, Cairo University,
Electrical Power Dept., Giza, Egypt

Abstract

The goal of this article is to investigate the implementation of the Cuckoo Search Algorithm (CSA) as an optimization technique to determine the parameters of variable parameters PID (VP-PID) controller. The VP-PID has three parameters that have to be optimally evaluated. A case of three physical imbedded nonlinearities in a single area electric power system has been selected to test the suitability of the proposed technique. The integral-square error (ISE) criterion has been considered as a part of the objective function together with the percentage overshoot and settling time. Matlab/Simulink software has been used in the simulation process. The simulation results show that the proposed VP-PID controller furnishes a better performance than the conventional PID controller.

Keywords: Proportional-Integral-Derivative controllers (PID); Variable Parameters PID controller (VP-PID); Cuckoo Search Algorithm; Nonlinear physical Systems; Matlab/Simulink

Introduction

In recent years, the problem of maintaining the power and frequency of a power system free from oscillations has become rapidly crucial need because of irregular load variations and imbedded system nonlinearities [1]. The unexpected load variations result in many undesired behaviours such as the mismatch of generated power and load demand for consumption. This can be achieved by the load frequency control (LFC) methods. Nowadays, plenty research work is going on to make the systems intelligent so the systems can successfully serve the benefits of all customers [2].

Several optimization approaches have been recorded in literature that can be applied to tune the conventional PID controller. This includes but not limited to:

Particle Swarm Optimization (PSO),

Genetic Algorithm (GA),

Bacterial Foraging Optimization (BFO) [2].

Furthermore, the LFC and AGC of a single area power system been investigated by a variety of techniques such as gravitational search algorithm [3], the modified particle swarm optimization [4], the artificial neural network [5], optimal control design [6], fuzzy controllers [23], proportional-integral-observer techniques [7], as well as LQR and Legendre wavelet function [8]. Also, the AGC control of single area power system with distributed generation has been investigated in [9]. All these researches deal only with no nonlinearities in the control loops.

On the other hand, the application of VP-PID controller has been studied in different fields by many research workers. The contribution of the variable parameters class of controllers is the cope the desired characteristic of dynamic systems. Based on the desired behavior of the system output, we can assume variable parameter controller. In general, most of the systems will start their response with peak overshoot and sustained oscillations. This can be reduced by implementing a conventional PID controller [10]. The tuned parameters of this class of controllers may be reduced after sometime when the response tends to have steady state. This is the philosophy behind the use of VP-PID controller [11-12].

In this paper, a new evolutionary optimization technique has been used for tuning the variable parameters PID controller. This algorithm is inspired by lifestyle of a bird family called cuckoo. Specific breeding and egg laying of cuckoos bird is the main basis of this optimization algorithm [13-14].

The Classical PID Controller

The PID controller is considered to be an important component in industrial control systems because of its capability of reducing the steady state error and enhancing the dynamic response and other static characteristics of systems. The PID controller is defined, mathematically, by the next equation [15-17]:

$$u(t) = K_{\mathrm{p}}\mathrm{e(t)} + K_{\mathrm{i}}\int_0^{\mathrm{t}}\mathrm{e}(\tau)\mathrm{d}\tau + K_{\mathrm{d}}\frac{\mathrm{d\,e(t)}}{\mathrm{dt}} \qquad (1)$$

Where e(t) is the system error, K_{i} is the integral constant, K_{p} is the proportional constant, K_{d} is the derivative constant and $u(t)$ is the controller output.

On the other hand, the VP-PID controller can be described by changing controller gains as follows

Case I : The proportional gain parameter (*Kp*) must be high firstly and then decrease to the normal value in order to accelerate system response. Under these conditions, the basic formula of *Kp* is:

$$K_p = C_1 + C_2 * e_{f(i)} \qquad (2)$$

Where C_1 and C_2 are considered constants achieving the following bounderies

$$C_{1_{min}} \leq C_1 \leq C_{1_{max}} \quad (3)$$

$$C_{2_{min}} \leq C_2 \leq C_{2_{max}} \quad (4)$$

and $e_{f(i)}$ is the final value of the system error at optimal tuning step *i*.

Case II : The derivative gain parameter (K_d) is increased to prevent the oscillation and overlapping. But, this increase causes the slowdown of the system response. Thus, change of K_d must be from the highest value to the lowest. K_d can be formulized as

$$K_d = C_3 + C_4 * e_{f(i)} \quad (5)$$

Where C_3 and C_4 are considered constants achieving the following bounderies

$$C_{3_{min}} \leq C_3 \leq C_{3_{max}} \quad (6)$$

$$C_{4_{min}} \leq C_4 \leq C_{4_{max}} \quad (7)$$

Case III : Integral gain parameter (K_i) is provided to initialize the steady state error but big values of this parameter causes oscillation and higher overshoot. Thus, change of K_i must be from the highest value to the lowest. K_i can be formulized as

$$K_i = C_6 - C_5 * e_{f(i)} \quad (8)$$

Where C_5 and C_6 are considered constants achieving the following boundaries

$$C_{5_{min}} \leq C_5 \leq C_{5_{max}} \quad (9)$$

$$C_{6_{min}} \leq C_6 \leq C_{6_{max}} \quad (10)$$

Cuckoo Search Algorithm (CSA)

The cuckoo search algorithm (CSA) is inspired by the cuckoos breeding behavior [13]. Cuckoos select a random nest for laying their eggs; and a cuckoo can place only one egg at the time. Only the eggs with the highest quality (solutions) are carried over to the next generation. The available number of host nests must be fixed. The host can detect an alien egg with a probability $P_a \in [0, 1]$. In this case, the host bird can either throw the egg out or give up the nest so as to construct a fully new nest in a new location.

In the case of generating new solutions x^{t+1} for a cuckoo i, an L*é*vy flight method is applied as follows [14].

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus \text{L}\acute{e}\text{vy}\ (\lambda) \quad (11)$$

Where α is the step size (α≥ 0) and it must be related to the scales of the studied problem. In most cases, it may be selected to be as α = O($\frac{L}{10}$) where L refers to the studied problem characteristic scale. λ is a variable such that $1 < \lambda \leq 3$.

In fact, the Lévy flight in essence presents a random walk whose random step length is drawn from an Lévy distribution. It has an infinite variance with an infinite mean.

Lévy $\backsim$ u=$t^{-\lambda}$, $(1 < \lambda \leq 3)$ (12)

A number of the new solutions should be generated by Lévy around the best solution obtained so far. This will speed up the local search.

The objective function J_i at iteration *i* is given as

$$J_i = \frac{1}{ISE_i} \quad (13)$$

so that

$$ISE_i = \int_0^\infty {e_i}^2(t)\, dt + \text{Settling Time}_i + \text{P.O.}_i \quad (14)$$

Where P.O.$_i$ and Settling Time$_i$ represent the percentage overshoot and settling time at iteration *i* of optimization procedure [15].

Application

The controllers described in this research work can be applied on a single area power system. The system consists of the governor; the turbine and the generator; and the load. The block diagrams implementing the conventional PID and the VP-PID controllers is delineated in Fig. 1.

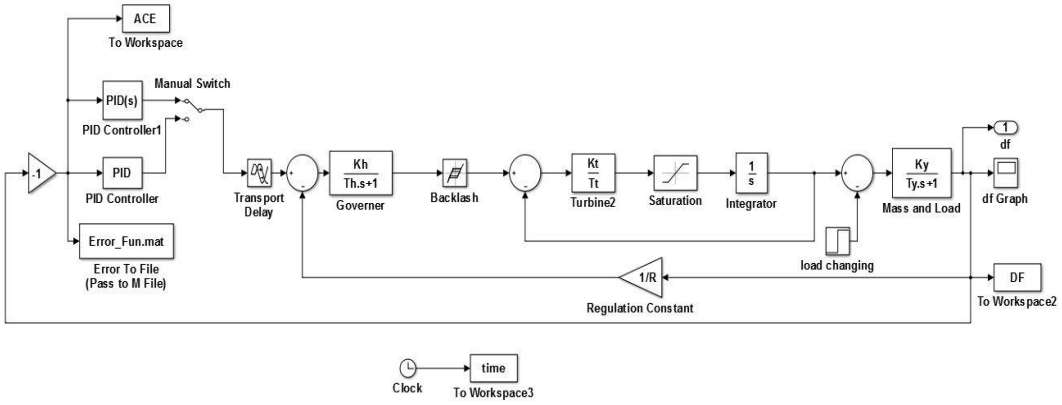

Fig. 1: Diagram of the LFC under investigation using Classical and VP-PID controllers

The parameters for this system as well as the characteristics of the turbine saturation (GRC), the GDB and the time delay as non-linear elements are listed in [1]. The system is subjected to a sudden load change of 0.05 p.u. The parameters of the Cuckoo Search Algorithm (CSA) are delineated in Table 1.

Table 1. The parameters of the Cuckoo Search Algorithm (CSA)

Parameter	Value
Number of individuals	10
Number of generations	100
The probability (P_a)	0.25

The flow chart of the algorithm is illustrated in Fig. 2.

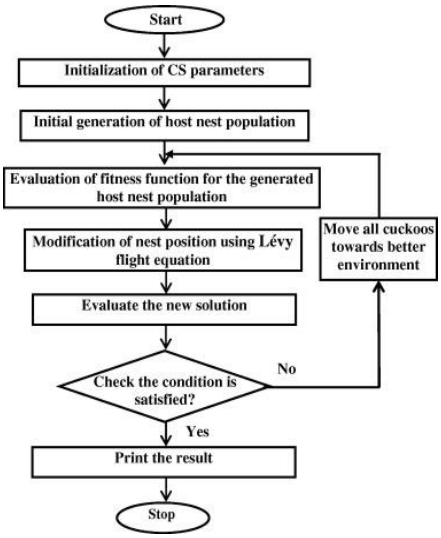

Fig. 2: The flow chart of the Cuckoo Search (CS) Algorithm [13]

Results and Discussion

The frequency control single area power system with GRC, GDB, and time delay nonlinearities incorporating the classical and VP-PID controllers has been studied in this research. The Cuckoo Search Algorithm has been utilized to tune the proposed controllers. The tuning process of the two controllers has been achieved using the Matlab/Simulink software. The results of the tuning procedure using the ISE criterion are summarized in Table 2.

Table 2. Results of the two proposed controllers

Controller	Parameters					
PID Controller	K_p	K_d	K_i	-	-	-
	2.4491	1.8840	0	-	-	-
ISE	12.2269					
	C_1	C_2	C_3	C_4	C_5	C_6
VP-PID controller	1.7509	3.7527	5.1860	1.1559	0	0
ISE	9.4966					

As stated in the literature, a proportional controller (K_p) will reduce the rise time and will decrease, but never eliminate, the steady-state error. An integral control (K_i) will eliminate the steady-state error, but it may make the transient response worse. A derivative control (K_d) will increase the stability of the system, reducing the overshoot, and improving the transient response [15-17].

As mentioned earlier, the first alternative is to implement the conventional PID controller as noted in Fig. 3. The parameters of the first controller (K_p, K_d, and K_i) are tuned using the Cuckoo Search Algorithm but those of the second controller, PID(s), are obtained a special optimization technique assigned to this block in Matlab.

The frequency deviation curve of the PID controller as shown in Fig. 3 is somewhat acceptable. There are two main notes. The first is the irregular effect at steady state. Second is the high value of ISE. The estimated optimal parameters of this controller are shown in Table 2. The tuning procedure using the CSA gives zero value for the integral gain K_i.

On the other hand, the frequency deviation given by the VP-PID controller is shown in Fig. 4. The estimated parameters variation of the VP-PID using CSA are delineated in Fig. 5. This controller will yield a smaller value for the ISE compared with the conventional PID controller as shown in Table 2. This, of course, can be achieved with little high computation time.

The frequency deviations obtained by the classical and VP-PID controllers are illustrated in Fig. 3 and 4. These curves illustrate that we cannot arrive, exactly, to a zero-state error. This is due to the imbedded system nonlinearities and the absence of integral controller. One more thing, the increase of the time delay will result in unstable system.

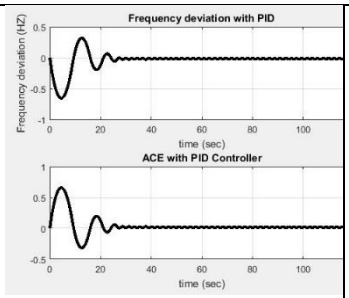

Fig. 3: Results of the PID controller (1.5 seconds time delay)

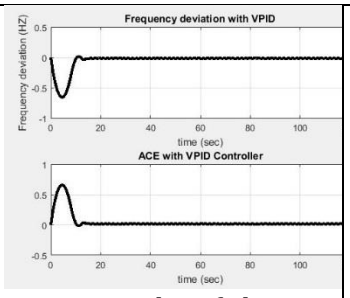

Fig. 4: Results of the VP-PID controller (1.5 seconds time delay)

(VPID=VP-PID)

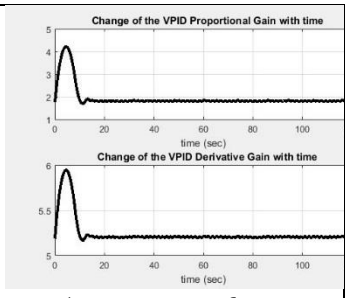

Fig. 5: Variation of VP-PID controller gains (1.5 seconds time delay)

(VPID=VP-PID)

Conclusion

In this paper, two classes of controllers have been applied to a single area LFC. The system under study has three sources of nonlinearities, the GRC, GBD, and time delay. The first controller is the classical PID controller while the second is the VP-PID controller. The parameters of these two controllers have been optimally evaluated using the Cuckoo Search Algorithm. Results show that the second controller behaves better than the first one. Furthermore, the effect of dealing with the three embedded nonlinearities, the GRC, GBD, and time delay, have been investigated but more detailed research is recommended.

References

[1] Aboelela, M. A. S., "Implementation of Artificial Intelligence Based Optimally Tuned Controllers to a Class of Embedded Nonlinear System", *7th International Conference on Modern Circuits and Systems Technologies (MOCAST)*, Thessaloniki, Greece, 7-9 May, 2018.

[2] Ahmed Husseina, S. Sadeem Saliha , and Y. Ghazi Ghasma, “Implementation of Proportional-Integral-Observer Techniques for Load Frequency Control of Power System”, Procedia Computer Science 109C, pp. 754–762, 2017.

[3] B. C. Kuo and F. Golnaraghi, “Automatic Control Systems, 9th Edition, John Wiley and Sons, NY, 2010.

[4] B. Ogbonnaa, S.N. Ndubisib " Neural network based load frequency control for restructuring power industry" Nigerian Journal of Technology (NIJOTECH), Vol. 31, No. 1, pp. 40-47, 2012.

[5] E. Rashedi, H. Nezamabadi-pour, S. Saryazdi, "GSA: a gravitational search algorithm," Information Sciences, vol. 179, 2009, pp. 2232-2248.

[6] Hong Zhang and Bo Hu, “The Application of Nonlinear PID Controller in Generator Excitation System”, Energy Procedia, Vol. 17, pp. 202 –207, 2012.

[7] K. Ogata, "Modern Control Engineering", 5th Edition, Pearson, 2010.
[8] K. Vijayakumar and T. Manigandan, "Nonlinear PID Controller Parameter Optimization using Enhanced Genetic Algorithm for Nonlinear Control System", CEAI, Vol.18, No.2 pp. 3-10, 2016.
[9] L. H. Hassan, Haider A.F. Mohamed, M. Moghavvem, and S.S. Yang "Load Frequency Control of Power Systems with Sugeno Fuzzy Gain Scheduling PID Controller" ICROS-SICE International Joint Conference, 2009, Fukuoka, Japan.
[10] M. M. Ismail, M. A. Mustafa Hassan "Load frequency Control Adaptation using Artificial Intelligent Technique for one and two different areas power system" International Journal of Control, Automation and Systems, Vol. 1, No. 1, Jan. 2012.
[11] O. Aydogdu and M.Korkmaz, "A Simple Approach to Design of Variable Parameter Nonlinear PID Controller", *IPCSIT vol.20, IACSIT Press, Singapore, 2011.*
[12] R. Boruah and B. Talukdar, " Modelling of Load Frequency Control Using PI Controller and Stability Determination with Linear Quadratic Regulator", *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE), Volume 9, Issue 4 Ver. VI (Jul – Aug. 2014), PP 30-36*
[13] R. C. Dorf, and R. H. Bishop, "Modern Control Systems, 13th Edition, Pearson, 2017.
[14] Santosh Kumar and S. Mishra, "AGC for distributed generation", In Proceedings of the Sustainable Energy Technologies, ICSET 2008. IEEE International Conference, 24-27 Nov. 2008. Singapore.
[15] T. S. Gorripotu, S. Gopi, H. Samalla, A. V. L. Prasanna, and B. Samira, "Applications of Computational Intelligence Techniques for Automatic Generation Control Problem-A Short Review from 2010 to 2018", In book: Computational Intelligence in Pattern Recognition, A.K. Das, J. Nayak, B. Naik, S.K. Pati, D. Pelusi (Eds.), Springer, Singapore, January 2020.
[16] X. S. Yang and Deb, S., "Cuckoo search via Lévy flights". In *Nature & Biologically Inspired Computing, World Congress,* IEEE., 210-214, 2009.
[17] Y. Abdelaziz and Ali, E. S., "Cuckoo Search algorithm based load frequency controller design for nonlinear interconnected power system".
" International Journal of Electrical Power & Energy Systems, 73: 632-643, 2015.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

A Study of Gas Diffusion Characteristics on Nano-Structured Ceramic Membranes

Priscilla Ogunlude

Robert Gordon University, Centre for Process Integration and Membrane Technology School of Engineering

Ofasa Abunumah

Robert Gordon University, Centre for Process Integration and Membrane Technology School of Engineering

Edward Gobina

Robert Gordon University, Centre for Process Integration and Membrane Technology School of Engineering

Abstract

The use of membranes for gas upgrading has increasingly become of interest as it has shown great potential for efficient and affective gas purification and a pathway to green energy. The emission of greenhouse gases to the atmosphere has detrimental effects on the economy in terms of global warming which has led to many natural disasters, heat waves, food shortage, loss of life and property. To combat this, studies of capturing and utilizing greenhouse gases are ongoing. In this paper, the study of biogas components (methane and carbon dioxide) diffusion through membranes are studied to employ its use as a solution for the challenge. The study involved the use of membranes of different pore sizes (15, 200 and 6000nm) to ascertain the flow characteristics and regime of the gases under different operating conditions. Single gas permeation tests were conducted, and the results show the flow of gases is dependent on factors including molecular weight, kinematic diameter and viscosity of the gas components. It was observed that pressure has a greater influence on the gas flow through membranes compared to temperature with the effect of pore size having the greatest impact. The flux of methane through the membrane is greater than that of carbon dioxide in regular pore geometry and depicts a greater potential for upgrading of biogas.

Keywords: Biogas, upgrading, emissions, nano-structured, membrane, carbon capture

Introduction

Greenhouse gas emissions (GHGs) and its effects have been a matter of global concern over the past decade(1). With growing energy demands to support developing economies, there has been a challenge of harnessing and utilizing renewable energy to meet these demands. Despite the effect of global warming and the problems associated with it, the use of fossil fuels is still increasing. This problem has negatively impacted the climate because the carbon dioxide evolved from burning fossils are increasing the GHGs. A possible cause of this problem is the technical difficulty of capturing and utilizing the GHGs. Thus, this study investigates a method of channeling biogas for use as a renewable energy source using membrane technology which could remedy the situation(2). Biogas is a mixture of mainly carbon dioxide (CO2) and methane (CH4) which is evolved from food waste, animal waste, sewage and municipal waste that have decomposed under certain conditions of temperature and pressure(3). Biogas holds a promising future to solving these challenges as it utilizes these gases as a renewable fuel to meet economic demand and prevent emissions that cause global warming.

In this article, nano-structured membranes would be utilized for the biogas upgrading process. Upgraded biogas can be used as an alternative energy source to fulfil global energy demands. Ceramic nanostructured membranes are chosen due to their versatility (for use in separation and reaction processes), high stability, long life-span, compact size, low energy demand, high efficiency and high resistance to aggressive chemical environment/harsh environmental conditions of temperature and pressure which prevail in industry(4).

Methodology

The membranes used for this experiment were inserted into an insulated reactor chamber. Permeation tests for the sample gases were carried out to study the behaviour of the methane/carbon dioxide and the selectivity of the membranes for each gas. The membranes were investigated under different conditions of temperature and pressure which were stabilized during each test run, this was confirmed using a pressure gauge and temperature indicator which was connected to various points of the reactor chamber. Tests were performed at various operating conditions. The outlet gas flow estimates how much methane versus carbon dioxide is evolved in order to figure out how well the membrane will separate the gases in an industrial application.

Figure 1 Experimental set-up showing all equipment including; pressure gauge(1), membrane module covered with heating tape(2), gas regulator(3), gas cylinder(4), heat regulator(5), volumetric meter(6) and temperature indicator(7)

Results and Discussion

Both CH4 and CO2 were passed through the membrane, but the flux of CH4 is significantly higher therefore this means that CO2 collides with the pore walls more than CH4 and thus loses momentum and absolute velocity to the wall; this can be attributed to the kinetic diameter of CO2 which is larger than that of CH4 and also shows molecular weight dependency which is representative of Knudsen mechanism.

The tables below show separation factors for the gases using different membranes and at varied temperature ranges. The separation factor was calculated based on the ratio of flow of

each gas which was obtained from experiments whilst the theoretical (Knudsen) separation factor is defined by (5):

Where MWB AND MWA are the molecular weights of B and A respectively.

$$\propto_{A/B} = \sqrt{\frac{MW_B}{MW_A}}$$

15nm membrane with inlet Pressure of 1 bar	
Ideal Knudsen Value	1.66
Experimental Knudsen Value at 298K	1.59
Experimental Knudsen Value at 323K	1.59
Experimental Knudsen Value at 343K	1.60
Experimental Knudsen Value at 373K	1.61

Conditions of 373K and 1 bar	
Ideal Knudsen Value	1.66
Experimental Knudsen Value using 15nm membrane	1.61
Experimental Knudsen Value using 200nm membrane	1.56
Experimental Knudsen Value using 6000nm membrane	1.55

This shows that pore size has a greater influence on the gas flow through membranes compared to temperature within this range. The flux of methane through the membrane is greater than that of carbon dioxide in regular pore geometry and depicts a great potential for separation/upgrading of biogas which is a mixture of both. Results also show that flux is dependent on the gas molecular weight, kinematic viscosity and viscosity as the heavier, larger and more viscous gas CO2 did not pass through the membrane as quickly as the lighter, smaller and less viscous CH4 gas.

It can also be seen that the separation factor increases with decreasing pore size as the 6000nm membrane show low selectivity compared to the 15nm membranes, this is attributed to the smaller pores that restrict the flow of carbon dioxide. These results show that by tuning the pore size of the membranes below 15nm, a greater separation efficiency can be achieved.

Conclusion

The experiments show that pore size has a greater influence on the gas flow through membranes compared to temperature with no significant changes in the graphs from 20 – 100oC. The flux of methane through the membrane is greater than that of carbon dioxide in regular pore geometry due to factors including molecular weight, kinematic diameter and viscosity of the gas components. This depicts a greater potential for separation/upgrading of biogas.

References

[1] Al Seadi T. Biogas handbook. Odense, Denmark: Syddansk Universitet; 2008.

[2] da Silva Biron D, dos Santos V, Zeni M. Ceramic Membranes Applied in Separation Processes [Internet]. Cham: Springer International Publishing; 2018 [cited 2019 May 23]. (Topics in Mining, Metallurgy and Materials Engineering). Available from: http://link.springer.com/10.1007/978-3-319-58604-5

[3] Damle A. An introduction to the utilization of membrane technology in the production of clean and renewable power. In: Membranes for Clean and

Renewable Power Applications [Internet]. Amsterdam: Elsevier; 2014 [cited 2019 May 23]. p. 3–43. Available from: https://linkinghub.elsevier.com/retrieve/pii/B978085709545950001X

[4] Domenico De Meis. Gas-transport-through-porous-membranes. Rome, Italy: Frascati Research Center; 2017.

[5] Lashof DA, Ahuja DR. Relative contributions of greenhouse gas emissions to global warming. Nature. 1990;344(6266):529.

ISSN 2601-8683 (Print)
ISSN 2601-8675 (Online)
European Journal of
Formal Sciences and Engineering
July - December 2022
Volume 5, Issue 2

Comparing Different Equipment and Applications in Pavement Data Collection as Part of Road Management System

Arjol Lule

PhD Cand, Department of Applied Geology, Environment and Geo-Informatics,

Faculty of Geology and Mining, Polytechnic University of Tirana

Shkelqim Daja

Prof. As. Department of Applied Geology, Environment and Geo-Informatics,

Faculty of Geology and Mining, Polytechnic University of Tirana

Abstract

National roads are the main arteries in road transport infrastructure. Therefore, all agencies or authorities responsible of road infrastructure, pay attention to road management systems. Albania is experiencing an increase in road infrastructure investments and maintenance of this road network. There have been some attempts to establish national and secondary road management systems. These systems attempt to achieve different objectives, such as the provision of an adequate level of service, the preservation of the road infrastructure, etc. A good **R**oad **A**sset **M**anagement **S**ystem (RAMS), helps to carry out all the actions of inventory, storage and maintenance of road assets as well as, supports the decision-making process. At present, there are several data collection devices and applications that carry out the job efficiently. The purpose of this paper is to present the analysis of the use and comparison of some equipment and Cell Phone Based Systems (MiniROMDAS, PaveProf-V2 and RoadLab_Pro) used for the road pavement data collection, necessary in the calculation of the International Roughness Index (IRI), along the national road network in Albania. The comparison is made, by analyzing the data and results obtained along a 20 km long road segment in Albania, using the various above-mentioned technologies. Also, an overview of the currently available technologies providing information that could assist managers in establishing an appropriate data collection program is given.

Keywords: MiniROMDAS, PaveProf-V2, Road Lab_ Pro, IRI.

Introduction

The national road network in Albania is about 4,000 km length and maintained by Albania Road Authority. The secondary and local road network is about 9500 km length and maintained by 61 municipalities of Albania. In Albania, too, there have been some attempts to establish both national and secondary road management systems. These systems combine several objectives, such as provision of an adequate level of service, preservation of the facility etc. At present, there are several data collection equipment and systems that carry out the job efficiently. ROMDAS, PaveProf-V2 and RoadLab_Pro are some of those systems, which are widely used to collect pavement data with different cost and quality. Different systems are used to evaluate pavement quality. All methods of evaluation are based on IRI (The **I**nternational **R**oughness **I**ndex). The International Roughness Index (IRI) is the roughness index most commonly obtained from measured longitudinal road profiles [1]. This study aims to assess the pavement road conditions in the Maminas – Shen Pjeter road, using different tools and system of data collection and to make a comparison between the results obtained.

Methodology

The pavement road conditions are assessed by the use of International Roughness Index (IRI) via different systems, such as ROMDAS, PaveProf-V2 and RoadLab_Pro. These systems are used in the pavement data collection in several important projects in Albania. The results are compared with the purpose of providing some recommendations regarding the tools to be used as function of the established evaluation requirements.

Raw Data Collection and Data Processing

Data collection in this study was carried out in Maminas – Shen Pjeter national road in Albania region with 21km length. Roughness data was recorded with 100 meters interval using ROMDAS, Pave Prof-V2 and RoadLab_Pro systems. Raw data collections are processed by each software to get the final product of this survey. In order to have the same data collection conditions, an almost constant velocity of 40 km/h was maintained.

MiniROMDAS

ROMDAS is a cost effective and modular system designed to collect road and pavement data using any vehicle. The MiniROMDAS, was used in road pavement data collection on about 385 km of the Albanian national road network in the frame of the Project "**O**utput and **P**erformance based **R**oad **M**aintenance **C**ontracts (OPRMC)", funded by Government of Albania and the World Bank. A schematic presentation of data collection process using the Mini ROMDAS is shown in the Figure 1.

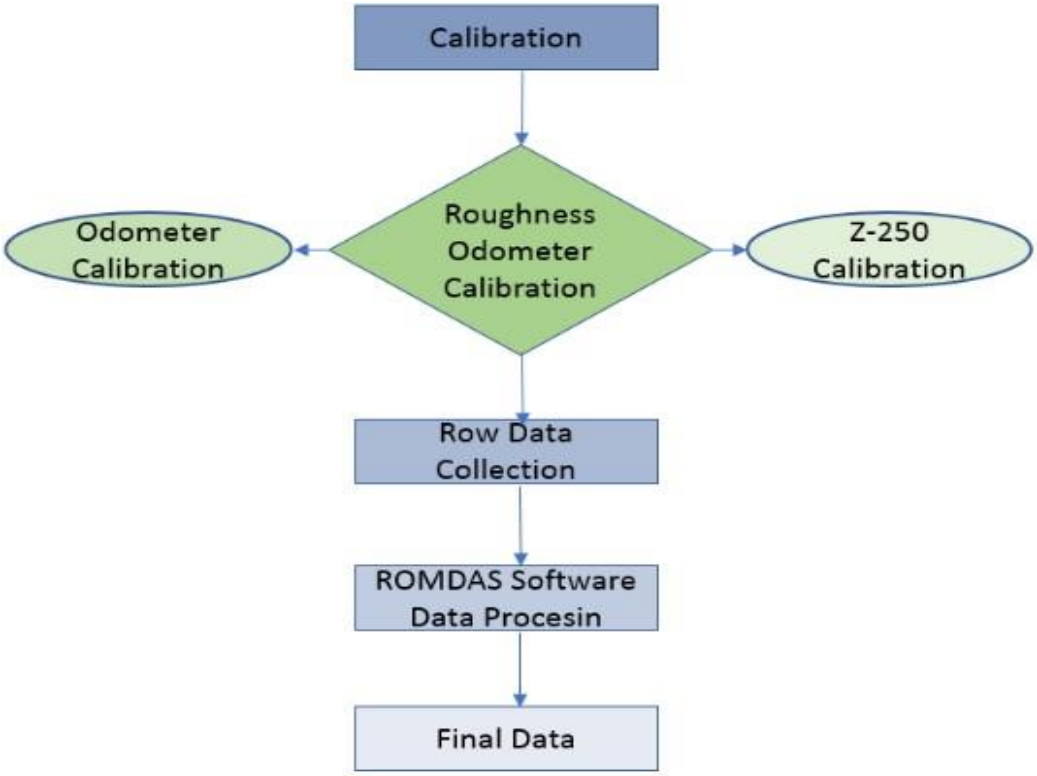

Figure 1.Framework of Data Collection Process

Before performing the data collection, the system needs to be calibrated. The MiniROMDAS calibration system is divided into three parts, as follows: 1) Z-250 reference profiler calibration, 2) odometer calibration and 3) roughness meter calibration.

Z-250 reference profiler was calibrated at the beginning in office, while odometer and roughness meter calibration were carried out in the certified service of the deliverer. Table 1 lists the data entered, necessary for providing the Z-250 calibration equation shown in Figure 2. Therefore, the slope adjustment factor (1,039) found from the analysis is entered into the profiler tab of data logger.

Z-250 Start Elevation: -0.09					
Shim Placements		Elevation (mm)			
Foot A	Foot B	Shim	Mean Elevation	Test 1	
				Display	Corrected
2	-	2	1.92	1.83	1.83
10	6	4	3.73	3.64	3.64
6	-	6	5.92	5.83	5.83
8	-	8	8.06	7.97	7.97
10	-	10	10.10	10.01	10.01
15	-	15	14.88	14.79	14.79
-	2	-2	-2.07	-2.16	-2.16
6	10	-4	-3.82	-3.91	-3.91
-	6	-6	-5.47	-5.56	-5.56
-	8	-8	-7.22	-7.31	-7.31
-	10	-10	-9.41	-9.50	-9.50
-	15	-15	-13.88	-13.97	-13.97

Table 1.Data Entry Component of Z-250 Calibration

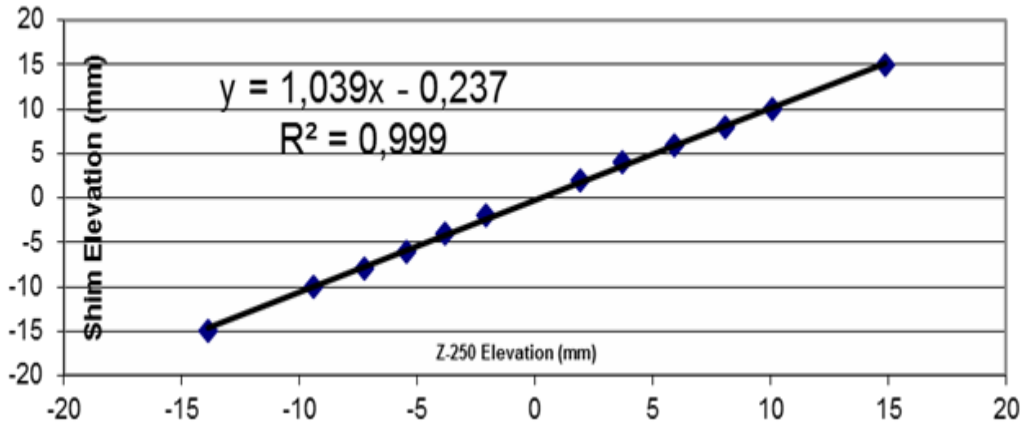

Figure 2.Regression Analysis for the Determination of Slope Adjustment Factor

Installation and calibration of the odometer are also necessary for the roughness meter calibration. The odometer, used for the accurate determination of the length and the speed, was calibrated in a 200-meter section using four runs as shown in the Table 2.

Table 2. Lists the data entry component of Odometer Calibration

Odometer Calibration Factor								Error Tolerance:			0.10%	
Run Number				Sample	Mean	Sdev	S.Error	S. Error	Beta	Beta	Pass/Fail	Pass/Fail
1	2	3	4					(%)	90%	95%	90%	95%
930	930	931	930	4	930.250	0.50	0.25	0.03	0.59	0.80	Pass	Pass

The Bumper Integrator BI, was installed in a vehicle having a Solid Rear Axle, measuring the so-called the 'Half-Car' roughness as shown in figure 3. [2]

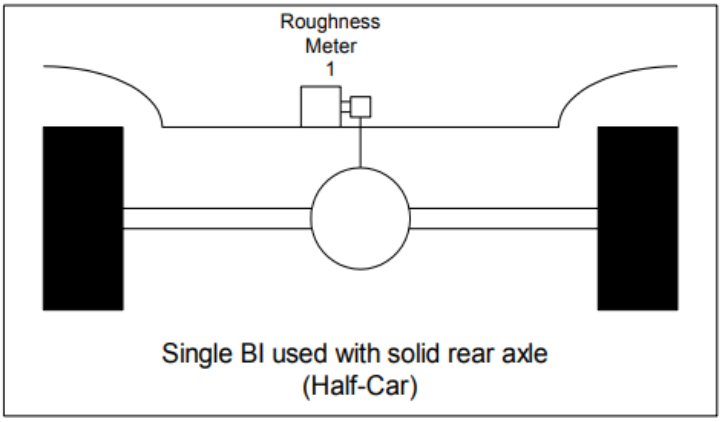

Figure 3 Single BI used with Solid Rear Axle (Source: ROMDAS user guide)

The roughness meter calibration in ROMDAS system was done using the calibrated Z-250 reference profiler. Seven different sections of 300 m length were used to measure the reference profiles. The output of BI is generated in terms of count per km, representing the cumulative number of bumps in one kilometer, as shown in table 3.

Table 3 Roughness of the seven Different Sections Roughness

Calibration Site	Site IRI (m/	Site Length (m)	Calibration Speed	ROMDAS Raw BI un Number 1	2	3	Number of	Run Mean Raw BI Count	Raw BI coun	Sdev	S.Error	S. Error	Beta 90%	Beta 95%	Pass/Fail 90%	Pass/Fail 95%
1	4.09	300	40	1273	1355	1136	3	1255	4182	111	64	1.5	186.5	274.88	Pass	Pass
2	4.5	300	40	1173	1170	1179	3	1174	3913	5	3	0.1	7.73	11.38	Pass	Pass
3	3.75	300	40	980	939	962	3	960	3201	21	12	0.4	34.65	51.06	Pass	Pass
4	3.09	300	40	866	882	889	3	879	2930	12	7	0.2	19.88	29.29	Pass	Pass
5	1.92	300	40	607	591	618	3	605	2018	14	8	0.4	22.89	33.73	Pass	Pass
6	2.25	300	40	708	699	709	3	705	2351	6	3	0.1	9.29	13.68	Pass	Pass
7	5.11	300	40	1491	1485	1490	3	1489	4962	3	2	0.0	5.42	7.99	Pass	Pass
1	4.09	300	60	1282	1342	1342	3	1322	4407	35	20	0.5	58.40	86.06	Pass	Pass

The ROMDAS software was used for the calculation of the Roughness Calibration Equation Coefficients, shown in figure 4. The R-squared value is above 0.9, showing a good repeatability and therefore good reliability of the BI results. The slope and the intercept respectively the coefficients a1 and a2 of the Calibration Equation are used in the calculation of IRI of the road segment under study.

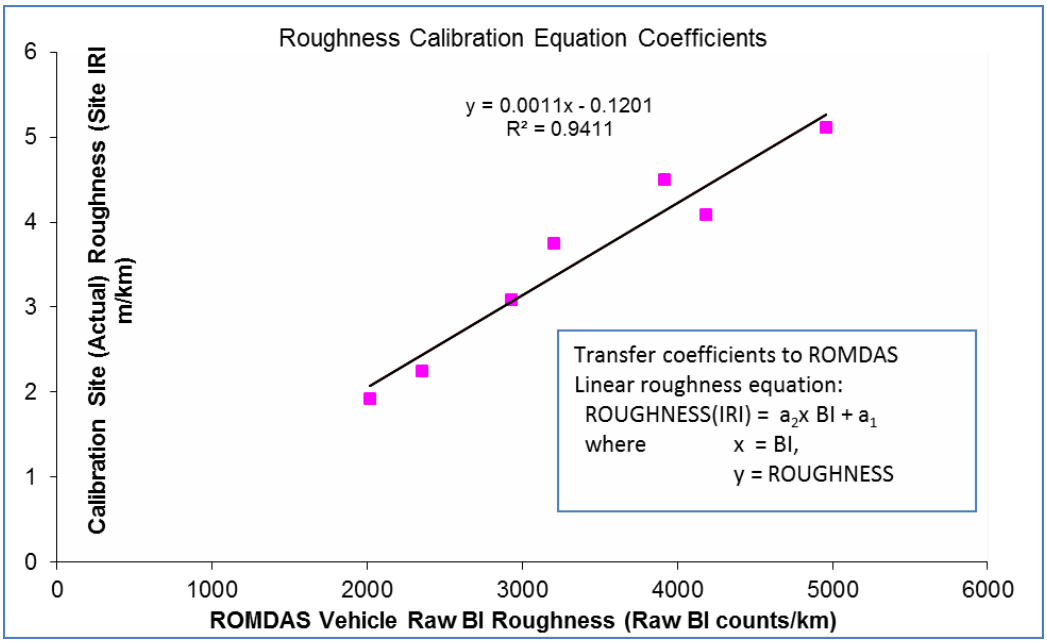

Figure 4 Calibration Equation Coefficients Worksheet

The miniRODMAS assembly is given in pictures of the figure 5.

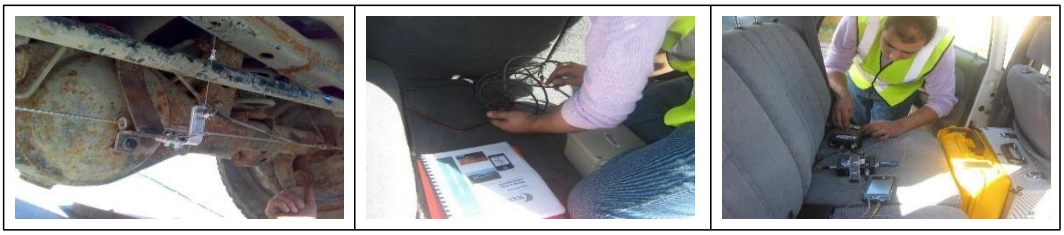

Figure 5 Assemble and calibrate the ROMDAS equipment

PaveProf-V2

*Pave*Prof-V2 is a modular system that uses laser sensors to measure pavement profiles for applications such as highways and runways [3]. A single laser and accelerometer system were used to collect road data on the Albanian road network in the frame of "**O**utput and **P**erformance based **R**oad **C**ontracts (OPRC)" Project. We had the opportunity by the contractor to use this system for collecting pavement data for Maminas-Shen Pjeter Road.

RoadLab_Pro

RoadLab_Pro, is designed as a data collection tool for engineer by the World Bank in collaboration with Beldor Center, SoftTeco and Progress Analytics LLC [4]. With accelerometers on Smartphone's, this app evaluates road conditions, map road networks, detects major road bumps, and reports road safety hazards [4].

The Road lab Pro is used on data collection in two important Projects in Albania, the "Technical Assistance for monitoring, communication and visibility of Transport Sector with Focus on Roads" and "Regional and Local Roads Connectivity". The RoadLab_Pro app was downloaded on the Smartphone, mounted in a vertical position. The data collected was emailed and the IRI values for every 100 m length are obtained. Accompanied by the coordinates of each point the app allows the mapping of the IRI values and consequently the road conditions.

Results and Discussion

The average every 100 m IRI values obtained using three different data collection systems (PaveProf-V2, ROMDAS and RoadLab_Pro), for Maminas – ShenPjeter road, are presented in the figure 6 and figure 7.

Pave of inas Pjeter Pr Ma-					BI_R _Ma _Sh r OMDmin enP					Road ro pj Ma Lab_ Sh et min					
Dista nce/	Dista nce/	Sp ee	IRI L1	GPS	CHA INA	LRP_ FRO	LRP _TO	SPE ED	C_RO UGH_	latit ude	longi tude	len gt	sp ee	roug hne	con ditio

0+000	0+100	44.05	3.365	41.378575N:19.606	100	0	100	38.7	3.08	41.5338	19.516491	110.1	39.6	2.8	Good
0+100	0+200	44.16	1.263	41.379123N:19.604	200	100	200	42.1	1.32	41.5330	19.517023	112.1	49.6	1.38	Very Good
0+200	0+300	44.32	1.487	41.379512N:19.603	300	200	300	42	1.41	41.5320	19.517225	117.4	39.9	1.34	Very Good
0+300	0+400	39.2	1.822	41.379898N:19.602	400	300	400	41.5	1.59	41.5311	19.517878	100.1	40.0	1.35	Very Good
0+400	0+500	38.6	1.646	41.380618N:19.601	500	400	500	41.2	1.57	41.5303	19.518488	107.5	40.5	1.5	Very Good
0+500	0+600	37.7	1.721	41.381428N:19.601	600	500	600	41.6	1.47	41.5295	19.517907	104.1	40.6	1.21	Very Good
0+600	0+700	39.3	2.068	41.382363N:19.601	700	600	700	40.9	1.75	41.5285	19.517945	106.6	44.0	1.44	Very Good
0+700	0+800	37.05	1.353	41.383198N:19.600	800	700	800	42.5	1.31	41.5277	19.519913	111.3	44.1	1.26	Very Good
0+800	0+900	35.12	1.635	41.384042N:19.600	900	800	900	42.6	1.56	41.5267	19.520368	112.5	44.3	1.48	Very Good
0+900	1+000	35.27	2.589	41.384898N:19.599	1000	900	1000	41.9	2.19	41.5257	19.520796	108.5	46.8	3.97	Fair
1+000	1+100	42.76	1.071	41.385798N:19.599	1100	1000	1100	41.6	1.31	41.5242	19.521903	107.6	40.5	1.55	Very Good
1+100	1+200	43.91	1.343	41.386907N:19.599	1200	1100	1200	41.1	1.40	41.5236	19.52294	111.8	40.6	1.46	Very Good
1+200	1+300	43.94	2.694	41.387602N:19.599	1300	1200	1300	39.5	2.12	41.5232	19.524151	112	44.0	1.55	Very Good

Figure 6.Processed Roughness Data

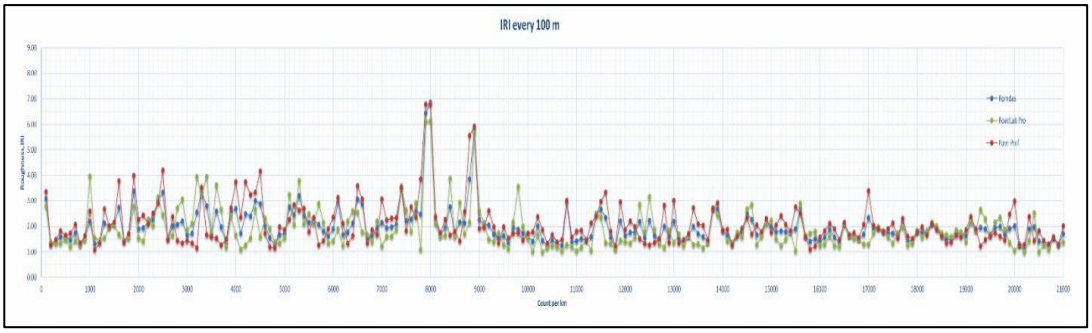

Figure 7.Chart of IRI every 100m for each system

The quality of the road pavement, for all systems used, was evaluated by the use of IRI (International Roughness Index), using the same scale of road conditions classification. The road pavement conditions are qualified as Very Good if the IRI are less than 2 m/km; Good if the IRI values are between 2 m/km and 4 m/km; Fair if the IRI values are between 4 m/km and 6 m/km; Poor if the IRI values are between 6 m/km and 8 m/km and Very Poor if the IRI values is larger than 8 m/km.

In the table 4 are summarized the road conditions based on IRI, related to the data collection system used in the survey.

Table 4. Road pavement classes based on IRI values in Maminas-ShenPjeter road based on different systems of data collection

Road pavement quality	Number of IRI values relative to road			Road class percentage		
	PaveProf	ROMDAS	RoadLab_Pr	PavePro	ROMDAS	RoadLab_Pr
Very good	125	134	144	59.52	63.81	68.57
Good (2<=IRI<4)	79	72	63	37.62	34.29	30.00
Fair (4<=IRI<6)	4	2	1	1.90	0.95	0.48
Poor (6<=IRI<8)	2	2	2	0.95	0.95	0.95
Very poor	0	0	0	0.00	0.00	0.00
Total	210	210	210	100	100	100

Based on the above table, regardless of the data collection tools or system used, the Maminas-Shenpjeter road pavement conditions are generally good. Taking into account the IRI scale chosen, about 60% of the road segment under study are classified as very good conditions, 35% as good and the rest as fair condition.

Differences between data collection tools are observed for IRI values less than 2 m/km, corresponding to very good road pavement conditions. In this case the PaveProf-V2 system is the most rigorous. The differences decrease with the increase of the IRI values, becoming equal for IRI values greater than 6 m/km, corresponding to poor road conditions.

In the following figure 8 is enlarged a segment which reflects the same degree of road condition classification for all three data collection systems used. The figure shows that for IRI values larger than 4 the results are similar.

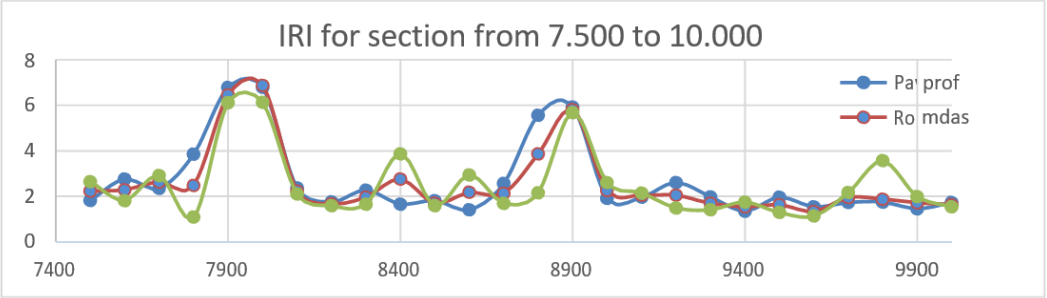

Figure 8.Chart of IRI from km 7+500 to km 10+000 for each system

The results of the study are also mapped showing the geographic location of each road condition class in GIS environment as shown in the figure 9, below.

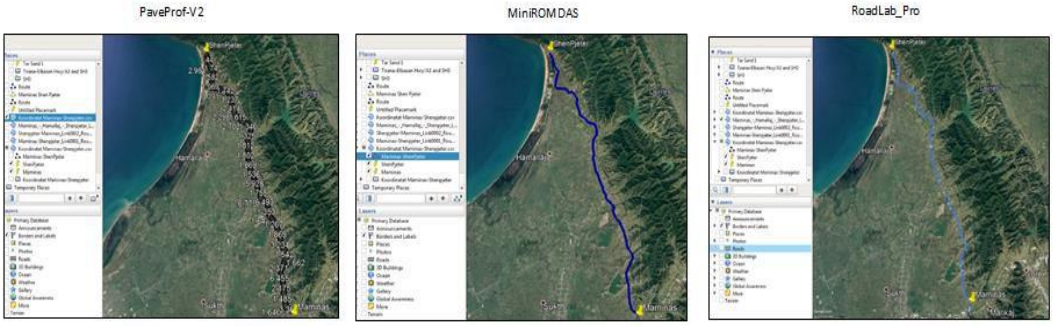

Figure 9. Map of road Maminas – Shen Pjeter for each system

Conclusion

The road pavement conditions in Maminas-Shenpjeter road are generally good, with small exceptions representing about 1.5-3 % of the total length, depending on the data collection system or tool used.

For good and very good conditions the use of PaveProf gives better results, while for IRI values higher than 4 the results are similar. In these terms the data collection system selection to be used depends on the road conditions to be assessed. In case of poor road conditions, a most convenient economic system can be used depending also in the project requirements. The RoadLab_Pro, representing the most economically convenient system, can be used in the assessment of roads in poor pavement conditions in Albania.

References

[1] Christopher R. Bennett, Alondra Chamorro, Chen Chen, Hernan de Solminihac, Gerardo W. Flintsch. Data Collection Technologies for Road Management, East Asia Pacific Transport Unit. The World Bank Washington, D.C, 2007

[2] Christopher R. Bennett, Alondra Chamorro, Chen Chen, Hernan de Solminihac, Gerardo W. Flintsch. Data Collection Technologies for Road Management, East Asia Pacific Transport Unit. The World Bank Washington, D.C, 2007, page 35

[3] Hunter, P. K., and Porter, H. A. ROMDAS for Windows User's Guide. Data Collection Limited, Motueka 7161, New Zealand, 2005.

[4] Sayers, M.W., Gillespie, T. D., and Paterson, W.D. Guidelines for the Conduct and Calibration of Road Roughness Measurements, World Bank Technical Paper No. 46, The World Bank, Washington DC, 1986.

[5] Sayers, M.W.; Karamihas, S.M. *(1998). "Little Book of Profiling" (PDF). University of Michigan Transportation Research Institute. Retrieved 2010-03-07.*

[6] http://himachalservices.nic.in/hpridc/ROMDAS/User Manual.pdf