Jordan Journal of Mechanical and Industrial Engineering (JJMIE)

JJMIE is a high-quality scientific journal devoted to fields of Mechanical and Industrial Engineering. It is published by Hashemite University in cooperation with the Jordanian Scientific Research and Innovation Support Fund, Ministry of Higher Education and Scientific Research.

SUBMISSION ADDRESS:

Prof. Moh'd Sami Ashhab, Editor-in-Chief
Jordan Journal of Mechanical & Industrial Engineering,
Hashemite University,
PO Box 330127, Zarqa, 13133 , Jordan
E-mail: jjmie@hu.edu.jo

Hashemite Kingdom of Jordan

Hashemite University

Jordan Journal of Mechanical and Industrial Engineering

JJMIE

An International Peer-Reviewed Scientific Journal

Financed by Scientific Research Support Fund

http://jjmie.hu.edu.jo/

ISSN 1995-6665

Jordan Journal of Mechanical and Industrial Engineering (JJMIE)

JJMIE is a high-quality scientific journal devoted to fields of Mechanical and Industrial Engineering. It is published by Hashemite University in cooperation with the Jordanian Scientific Research and Innovation Support Fund, Ministry of Higher Education and Scientific Research.

Introduction: The Editorial Board is very committed to build the Journal as one of the leading international journals in mechanical and industrial engineering sciences in the next few years. With the support of the Ministry of Higher Education and Scientific Research and Jordanian Universities, it is expected that a heavy resource to be channeled into the Journal to establish its international reputation. The Journal's reputation will be enhanced from arrangements with several organizers of international conferences in publishing selected best papers of the conference proceedings.

Aims and Scope: *Jordan Journal of Mechanical and Industrial Engineering* (JJMIE) is a refereed international journal to be of interest and use to all those concerned with research in various fields of, or closely related to, mechanical and industrial engineering disciplines. *Jordan Journal of Mechanical and Industrial Engineering* aims to provide a highly readable and valuable addition to the literature which will serve as an indispensable reference tool for years to come. The coverage of the journal includes all new theoretical and experimental findings in the fields of mechanical and industrial engineering or any closely related fields (Materials, Manufacturing, Management, Design, Thermal, Fluid, Energy, Control, Mechatronics, and Biomedical). The journal also encourages the submission of critical review articles covering advances in recent research of such fields as well as technical notes.

Guide for Authors

Manuscript Submission:
High-quality submissions to this new journal are welcome now and manuscripts may be either submitted online or email.

Online: For online and email submission upload one copy of the full paper including graphics and all figures at the online, submission site, accessed via http://jjmie.hu.edu.jo. The manuscript must be written in MS Word Format. All correspondence including notification of the Editor's decision and requests for revision, takes place by e-mail and via the Author's homepage, removing the need for a hard-copy paper trail

Submission address and contact :
Prof. Moh'd Sami Ashhab
Editor-in-Chief
Jordan Journal of Mechanical & Industrial Engineering, Hashemite University
PO Box 330127, Zarqa, 13115, Jordan
E-mail: jjmie@hu.edu.jo

Types of contributions*:* Original research papers and Technical reports

Corresponding author: Clearly indicate who is responsible for correspondence at all stages of refereeing and publication, including post-publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Full postal addresses must be given for all co-authors.

Original material: Submission of an article implies that the work described has not been published previously (except in the form of itstan abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, tha publication is approved by all authors and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher. Authors found to be deliberately contravening the submission guidelines on originality and exclusivity shall not be considered for future publication in this journal.

Withdrawing: If the author chooses to withdraw his article after it has been assessed, he shall reimburse JJMIE with the cost of reviewing the paper.

Manuscript Preparation:

General: Editors reserve the right to adjust style to certain standards of uniformity. Original manuscripts are discarded after .publication unless the Publisher is asked to return original material after use. Please use MS Word for the text of your manuscript

Structure: Follow this order when typing manuscripts: Title, Authors, Authors title, Affiliations, Abstract, Keywords, Introduction, Main text, Conclusions, Acknowledgements, Appendix, References, Figure Captions, Figures and then Tables. Please supply figures imported into the text AND also separately as original graphics files. Collate acknowledgements in a separate section at the end of the article and do not include them on the title page, as a footnote to the title or otherwise.

Text Layout: Use 1.5 line spacing and wide (3 cm) margins. Ensure that each new paragraph is clearly indicated. Present tables and figure legends on separate pages at the end of the manuscript. If possible, consult a recent issue of the journal to become familiar with layout and conventions. All footnotes (except for table and corresponding author footnotes) should be identified with superscript Arabic numbers. To conserve space, authors are requested to mark the less important parts of the paper (such as records of experimental results) for printing in smaller type. For long papers (more than 4000 words) sections which could be deleted without destroying either the sense or the continuity of the paper should be indicated as a guide for the editor. Nomenclature should conform to that most frequently used in the scientific field concerned. Number all pages consecutively; use 12 or 10 pt font size and standard fonts.

Corresponding author: Clearly indicate who is responsible for correspondence at all stages of refereeing and publication, including post-publication. The corresponding author should be identified with an asterisk and footnote. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Full postal addresses must be given for all co-authors. Please consult a recent journal paper for style if possible.

Abstract: A self-contained abstract outlining in a single paragraph the aims, scope and conclusions of the paper must be supplied.

Keywords: Immediately after the abstract, provide a maximum of six keywords (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible.

Symbols: All Greek letters and unusual symbols should be identified by name in the margin, the first time they are used.

Units: Follow internationally accepted rules and conventions: use the international system of units (SI). If other quantities are mentioned, give their equivalent quantities in SI

Maths: Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

References: All publications cited in the text should be presented in a list of references following the text of the manuscript.

Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.

List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.

Examples:

Reference to a journal publication:

[1] M.S. Mohsen, B.A. Akash, "Evaluation of domestic solar water heating system in Jordan using analytic hierarchy process". Energy Conversion & Management, Vol. 38 (1997) No. 9, 1815-1822.

Reference to a book:

[2] Strunk Jr W, White EB. The elements of style. 3rd ed. New York: Macmillan; 1979.

Reference to a conference proceeding:

[3] B. Akash, S. Odeh, S. Nijmeh, "Modeling of solar-assisted double-tube evaporator heat pump system under local climate conditions". 5th Jordanian International Mechanical Engineering Conference, Amman, Jordan, 2004.

Reference to a chapter in an edited book:

[4] Mettam GR, Adams LB. How to prepare an electronic version of your article. In: Jones BS, Smith RZ, editors. Introduction to the electronic age, New York: E-Publishing Inc; 1999, p. 281-304

Free Online Color : If, together with your accepted article, you submit usable color and black/white figures then the journal will ensure that these figures will appear in color on the journal website electronic version.

Tables: Tables should be numbered consecutively and given suitable captions and each table should begin on a new page. No vertical rules should be used. Tables should not unnecessarily duplicate results presented elsewhere in the manuscript (for example, in graphs). Footnotes to tables should be typed below the table and should be referred to by superscript lowercase letters.

Notification: Authors will be notified of the acceptance of their paper by the editor. The Publisher will also send a notification of receipt of the paper in production.

Copyright: All authors must sign the Transfer of Copyright agreement before the article can be published. This transfer agreement enables Jordan Journal of Mechanical and Industrial Engineering to protect the copyrighted material for the authors, but does not relinquish the authors' proprietary rights. The copyright transfer covers the exclusive rights to reproduce and distribute the article, including reprints, photographic reproductions, microfilm or any other reproductions of similar nature and translations.

Proof Reading: One set of page proofs in MS Word format will be sent by e-mail to the corresponding author, to be checked for typesetting/editing. The corrections should be returned within **48 hours**. No changes in, or additions to, the accepted (and subsequently edited) manuscript will be allowed at this stage. Proofreading is solely the author's responsibility. Any queries should be answered in full. Please correct factual errors only, or errors introduced by typesetting. Please note that once your paper has been proofed we publish the identical paper online as in print.

Author Benefits:

No page charges: Publication in this journal is free of charge.

Free offprints: One journal issues of which the article appears will be supplied free of charge to the corresponding author and additional offprint for each co-author. Corresponding authors will be given the choice to buy extra offprints before printing of the article.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665

Jordan Journal of Mechanical and Industrial Engineering

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 141 - 148

Jordan Journal of Mechanical and Industrial Engineering

The Influence of Internal Carotid Artery Off-plane Angle on the Hemodynamics of Blood Flow in the Carotid Artery

Thakir AlMomani[* a], Suleiman Bani Hani[b], Othman Smadi[a], Salah Azghoul[a], Laila Kiswani[c], Majdoleen Assarsour[c]

[a]Bioemdical Engineering department, Faculty of engineering, The Hashemite University, Zarqa, Jordan
[b]Mechatronics Engineering department, Faculty of engineering, The Hashemite University, Zarqa,Jordan
[c]The Hashemite University, Zarqa, Jordan

Received 21 May. 2019

Abstract

The geometry of the carotid artery (IC) is thought to have a major role in causing blood flow related diseases in this artery. With age, different geometrical changes are observed to occur. These changes include: diameter, bifurcation angle, and internal carotid off-plane angle. The current study focuses on the changes that happen to the off-plane angle of the internal carotid artery to investigate the hemodynamic parameters that are thought to lead to the formation of plaques and atherosclerosis in this artery. Physiological values found in literature are used to compare vorticity and velocity fields for elderly and young people. Steady, three dimensional CFD computations were performed. Inlet velocity of 0.3 mm/sec, with no slip boundary conditions and rigid walls are used. The results showed that as the age or off-plane increases, the flow abnormalities and vorticity values increase. At the apex, the velocity values were the maximum near the inner wall, while at the mid and in the second half of the sinus, the velocity values were minimal and show back flow with higher values of vorticities. These results of the current study indicated that as the off-plane angle increases, as in elderly people, diseases of blood flow, such as atherosclerosis and unfavorable plaques, have higher probability to occur in such complicated artery.

Keywords: Carotid artery, aging, internal carotid off-plane angle, vorticity, velocity magnitude, atherosclerosis, plaques;

1. Introduction

Cardiovascular system is the system that is responsible for blood supply to different living tissues through blood circulation. Blood flow through this system is associated with many diseases that might cause insufficient blood supply, or lead to more complicated cases, such as strokes, heart attacks, and atherosclerosis. (1) Both experimental and numerical studies have shown that the initiation and the growth of cardiovascular diseases is not random. Regions with secondary flows, higher vorticities, and abnormal shear stress values are thought to promote more cell damage, deconstruction, and growth in the vessel lumen. (1)

One of the geometrical, critical, regions in the cardiovascular system is the carotid artery that provides brain and face with blood. The presence of the branching with the, normally, local dilation in the internal carotid artery (sinus) makes its geometry as one of the most complicated and favorable geometries that are thought to cause flow abnormalities that associated with atherogenic and atherosclerosis in this part of the arterial tree. (1)

1.1. *Carotid Artery Related Diseases:*

The preliminary stage of atherosclerosis and plaque formation is the presence of lipids in the vascular wall caused by sub endothelial retention of lipoproteins. High plasma concentrations of LDL-c (low-density lipoprotein cholesterol) and low plasma concentrations of HDL-c (high-density lipoprotein cholesterol) are each independent risk factors for the development of such diseases. (2) LDL has been shown to be the major carrier of cholesterol and the main atherogenic lipoprotein for atherosclerosis. In addition to concentrations of LDL and HDL in systemic circulation, some physical and fluid mechanical factors play an important role in the development of atherosclerotic plaques through the activation of blood constituents and endothelial cells. (2) The surface concentration of LDL is locally elevated in regions of fluctuated flow and fluctuated shear stress, which causes more flow abnormalities and provides a favorable condition for the formation of atherosclerosis and aggregated plaques. (2) Studies on flow dynamics in the carotid artery have demonstrated that laminar flow is

* Corresponding author e-mail: thakir2000@hu.edu.jo

disturbed at carotid bifurcation regions. The greatest plaque accumulation typically occurs on the outer wall of the proximal segment of the sinus of the internal carotid artery, in the region of the lowest wall shear stress. [1,2] The intimal thickness is the least on the flow divider side at the junction of the internal and external carotid arteries where wall stress is the highest. [3]

Experimentally, it was reported that atherosclerosis is the beginning of most of the diseases in brain that related to the blood flow, as more studies have demonstrated that occlusive thrombus triggered by plaque rupture is one of the major determinants of ischemic stroke in patients affected by carotid atherosclerotic disease. [4] The majority of ischemic strokes seem to result from embolization from an atherosclerotic plaque or acute occlusion of the carotid artery and propagation of the thrombus distally rather than static occlusion. The explanation for the low rate of total occlusions in carotid plaques is most likely related to high flow rates that limit thrombotic occlusions, unless there is severe luminal narrowing caused by repeated plaque ruptures. [5]

Due to the travelling parts of thrombus plaque from the carotid artery to the small arteries of the brain, which affects directly the hippocampus which is the part of the brain that is vital for learning and memory, particularly converting your short-term memory into long-term. [6] This part of the brain, also, helps human comprehend spatial relationships and navigate the world around him. Consequently, hippocampus is one of the first parts of the brain that suffers damage from plaques and tangles that light lead to Alzheimer's or paralysis diseases. Furthermore, clinical studies of the association of atherosclerosis risk factors and dementia have shown positive associations between dementia and carotid atherosclerosis. [6]

1.2. Flow Abnormalities in the Carotid Artery:

The unique geometry of the carotid artery has been correlated to the higher chances of formation of, mainly, atherosclerosis and the un-preferable plaques in this artery which might lead to strokes. As can be seen in figure 1, carotid artery combines branching, curvature, and blubbing (or dilation). All these geometrical complications cause to change the uniform, laminar flow upstream the apex into a turbulent - like, non-uniform flow downstream the internal and the external carotid arteries. [7] Such abnormalities are associated with abnormal shear stress values and higher vorticity values which will lead to cell deconstruction and proliferation. [8]

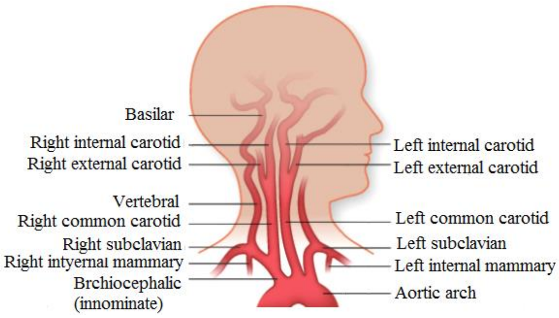


Figure 1. The geometry of the carotid artery

Vital cells as platelets, red blood cells, and endothelial cells are highly sensitive to the variation and abnormal fluctuations in the shear stress values. For example, Gnasso at el. (1996) have indicated the normal shear stress in human circulation is between 3.1 to 29.5dynes/cm^2. [9] In their study, Harish Shankaran at el (2002) have demonstrated that platelet activation occurs at 83 dynes/cm^2 value of shear stress. [10] Moreover, endothelial cells are used to be exposed to values of shear stress of 29.5 to 3.1 dynes/cm^2, higher or lower values of shear stress are, experimentally, observed to injure these cells, cause aggregation, growth, and proliferation new cells in the lumen of the blood vessel.[11, 12] More over, in the case of the negative pressure, recent experimental studies have indicated that these cells are showing higher trend to growth toward the internal of the artery lumen, causing the formation of un-preferable plaques in this area. [12]

On the other hand, experimental measurements of flow parameters, mainly shear stress and vorticities, in a scale of human carotid artery model have been conducted using different techniques such as: laser Doppler anemometry and MRI techniques. [13-18] The results of such studies are considered as a spark for many researchers to carry new in-vivo and in vitro experimental works or computational studies. In many cases, scientists tend to match experimental results with those that can be obtained from computational fluid dynamics simulations (CFD). [19,20] These recent studies aimed at defining the regions of high and low stress regions or the regions of diseases initiation.

As a result of these efforts, two theories have been adapted. The first one, which appeared early in 1960s, has mentioned that the main cell damage, deconstruction, and growth in the carotid artery resulted from the low shear stress values. [1] Based on their in vivo observations, and depending on different imaging techniques, they clearly found that more cells deposition and plaque formation is occurred on the external wall of the internal carotid artery. [1] Such regions are usually are associated with low shear stress values. Furthermore, they clearly stated that such low shear stress values are combined with both higher vorticities and negative pressure values. Such conditions are responsible to increase the residence and exposure time of the cells in such abnormal flow regime areas. [1]

Contradicting to the low shear stress theory, newer researchers performed more intensive in vitro experiments using flow regimes and geometries, similar to those that could be found in the physiological carotid artery. [1] In their results, they have mentioned that the cells are more sensitive to high shear stress values than lower shear stress values. More cell damage and deconstruction were observed to occur in the apex of the internal carotid artery, where higher blood streams and shear stress values are observed. Based on that, they conclude that diseases in the carotid artery are initially appeared in the high shear stress values then it develops there. [1]

In the last two decades, these two theories were studied intensively. More concern and attention were given to the low shear stress theory. Many researchers have reported that accumulative effect of flow abnormalities on both blood constituents and endothelial cells is more critical than the instantaneous effect of abnormal values. Cells need to stay for a certain and longer time to get activated and start to from clots and plaques. The presence of vorticities causes them to be exposed more to the abnormal flow values, and hence abnormal growth in such areas. [21]

One of the major geometrical issues that has been reported to affect the flow abnormalities in the carotid artery is the bifurcation angle between the internal and the external carotid arteries and the plane off angle. In their work, Thomas et al. (2005) and Jeon et al. (2017) accomplished a statistical study between the old and young people to investigate the differences between the geometry of the carotid artery between young and old people using MRI imaging technique. [4, 21] In their findings, they show that young people have, almost, uniform shape with little or no curvature in their carotid artery. While old people, they reported that carotid artery has a non-uniform with curvature rates occur in the internal carotid artery. [21] In numbers, they mentioned that the bifurcation angle for young people has an average of 48.5^o people while it has an average value of 63.6 o for elderly people. On the other hand, they reported that the plane off angle for young people is almost 0^o while it's around 16^o for elderly ones. These variation in shape between the young's and elders' carotid artery may be related to the reason why the cerebral vascular accident occurrence in the elder has higher probability than young. [4, 21]

1.3. Objectives

The current work utilized CFD simulations to investigate the influence of off-plane angle variation on the flow parameters of the fluid, mainly velocity and vorticity fields. Off-plane angles of 0^o, 5^o, 10^o, and 15^o will be used. More focus will be given to the flow patterns in the sinus of the internal artery. As mentioned earlier this normal dilatation is thought to be the most critical place where more brain diseases start at. Computational models are built based on the geometries found in literature. [4, 21] FLUENT-ANSYS software is used to create the mesh, identify boundary and initial conditions, and solve flow equations.

2. Methodology

2.1. Geometry and Dimensions

Dimension and geometries reported in Thomas Et al. 2005 is used to construct the current geometry. The dimensions are shown below in table 1. As mentioned earlier the off-plane angles are 0^o, 5^o, 10^o, and 15^o as shown in figure (2). [4, 22]

Table 1. Dimension of simulated carotid artery

Parameter	Diameter (mm)	Length (mm)
CCA Diameter	8	47
ICA Diameter	5.5536	60
ECA Diameter	4.6224	60
major Sinus	22	
minor Sinus	8	

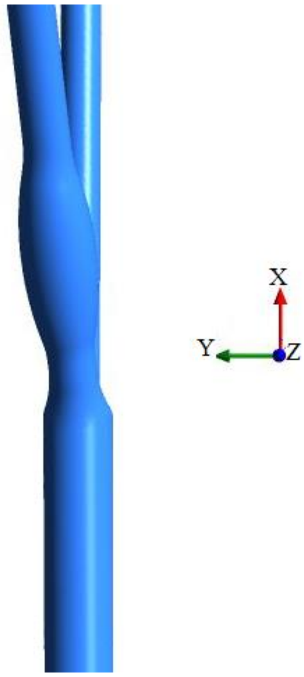


Figure 2. Model of the used carotid artery that shows the off-plane angle.

2.2. Governing equations, boundary, and initial conditions:

The governing equations include the incompressible equations of mass and momentum conservation given by:

$$\frac{\partial u_x}{\partial x}+\frac{\partial u_y}{\partial y}=0 \tag{1}$$

$$\rho\left(\frac{\partial u_x}{\partial t}+u_x\frac{\partial u_x}{\partial x}+u_y\frac{\partial u_y}{\partial y}+u_z\frac{\partial u_z}{\partial z}\right)= -\frac{\partial P}{\partial x}+\mu\left(\frac{\partial^2 u_x}{\partial x^2}+\frac{\partial^2 u_y}{\partial y^2}+\frac{\partial^2 u_z}{\partial z^2}\right) \tag{2}$$

Here, u_x, u_y,and u_z are the velocity along, x, y, and z-directions. ρ is the blood density, μ is the blood viscosity.

For the boundary conditions, the vessel walls are prescribed as rigid walls with no slip boundary condition. At the inlet, the velocity specified to be 0.3mm/s. At the outlet, the standard outflow condition is applied, i.e., the velocities are linearly extrapolated and corrected to be consistent with global mass conservation and a Neumann condition is applied on the pressure. The blood is treated as non-Newtonian homogenous fluid with density of 1060 kg/m^3 and a viscosity of 0.0035 kg/m.s. [1] The iterative technique is used to solve the above equations with initial velocity vector of zero magnitude.

3. Computational method:

Flow in carotid artery is simulated using finite volume-based software package ANSYS-FLUENT. The mesh is constructed using pro-E software. Different off-plane angles were used to simulate the flow in elderly and young carotid arteries. Figure 3 shows the velocity magnitude plotted versus radial location at the inlet of the sinus region for different mesh sizes for an off-plane angle of 0°. Grid independency was performed for many cell numbers starting with 353209 and ending with 583139, the magnitude of the velocity magnitude is presented versus the radial location of the centre plane of the carotid sinus. As can be observed on this figure grid independency occurred when the cell number was 436261. Consequently, this mesh size is used for the rest of the simulations.

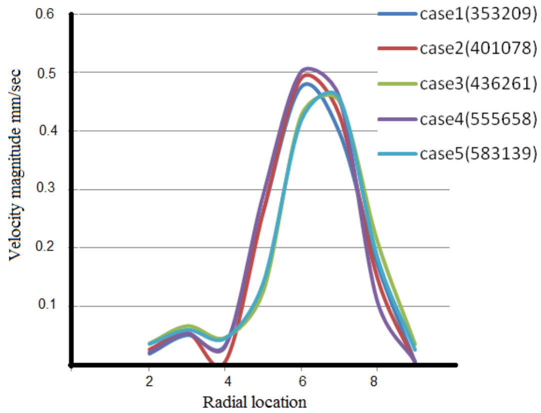


Figure 3. Mesh independency; velocity magnitude versus radial position in the mid of the sinus plane.

4. Computational Results:

Four different IC off-plane angles have been simulated and analyzed to study the influence of aging with the hemodynamics of blood and carotid bifurcation. Physiological flow conditions are applied for all cases. The results are mainly include the velocity and vorticity contours for different planes; three of them present the sinus region, and one plane is the medial plane that perpendicular on the IC such that this plane divided the IC carotid artery into two symmetric volumes, so that the internal and external walls of the IC artery are captured in this plane (figure 4). Consequently, the flow changes in the inner and the outer walls of the sinus cane be clearly demonstrated.

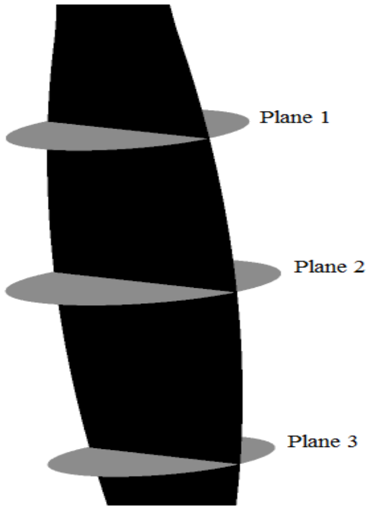


Figure 4. Presented planes to show simulated results.

4.1. Velocity contours

Figures (5-7) shows the contours of the velocity magnitude, for the studied off-plane angles for locations number 1, 2, and 3 (as numbered and shown on figure 4). For all cases, at the outer IC wall, velocity magnitude became lower and lower, while at the internal wall of the IC the velocity magnitude became higher and higher. Also, in all cases velocity gradient, and hence shear stress, is changed as off-plane angle increases.

Figure 8 shows the velocity contours for the frontal plane of the carotid artery. Also, in this figure, values of velocity are dependent on the off-plane angle. The minimal velocity values occurred on the outer wall region and decrease with increasing the off-plane angle, while maximal values occurred near the inner walls and increase with increasing the off-plane angle. In addition to the inner and the outer walls, this figure shows the maximum velocity value and magnitude occurs at the division between the internal and the external arteries, i.e., the apex of the carotid arteries. As can be seen in the figure, the blood stream became higher and greater downstream the apex or the bifurcation angle as the off-plane angle increases.

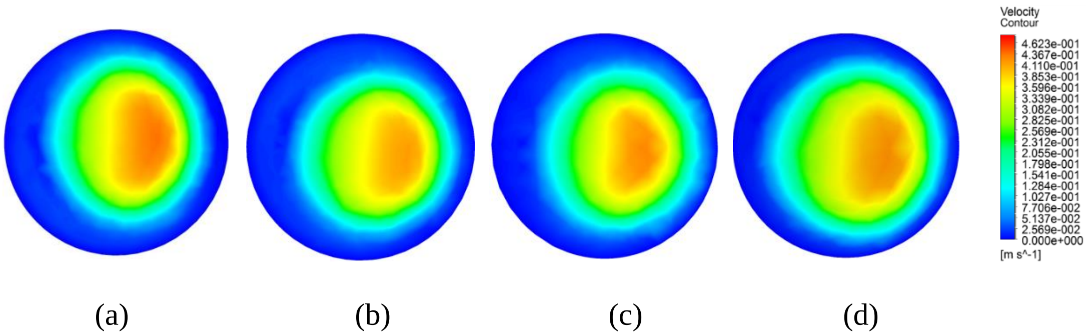
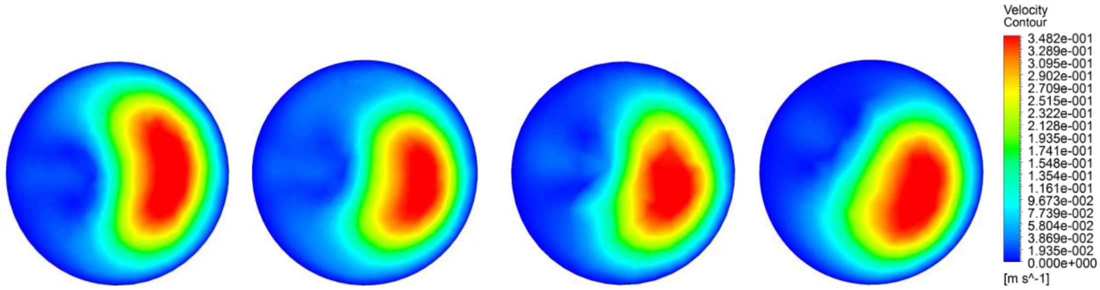
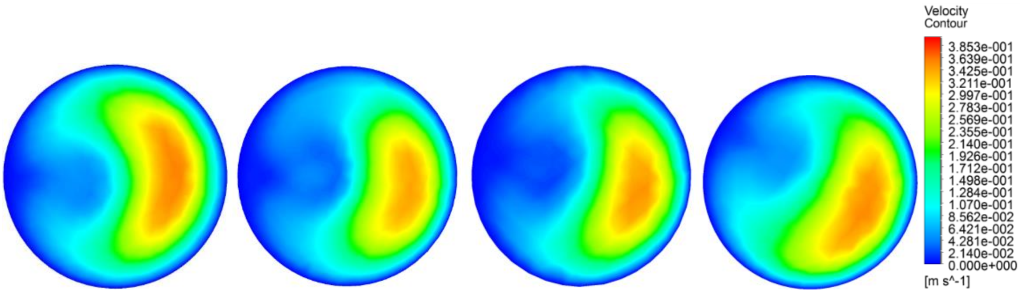


Figure 5. Velocity contours (plane 3: start of the sinus); off-plane angle of: a) 0° b) 5° c) 10° d) 15°



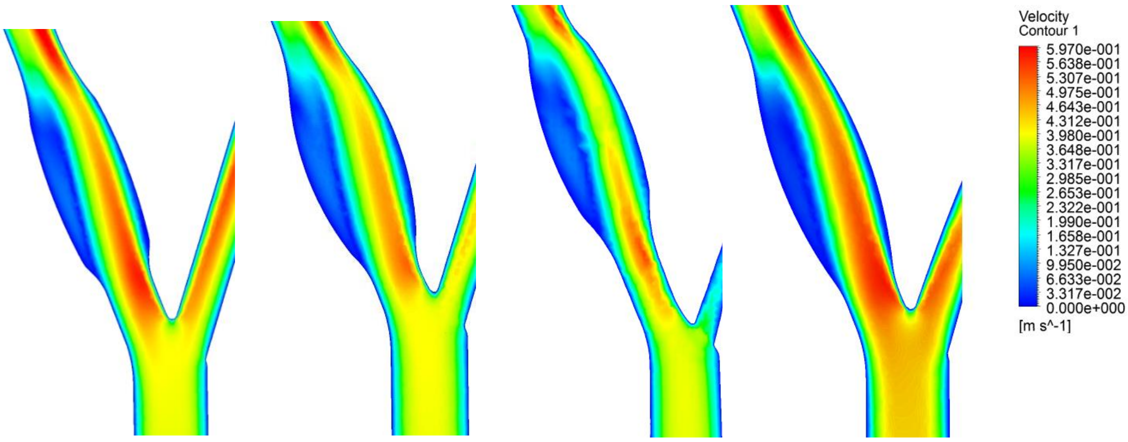
(a) (b) (c) (d)

Figure 6. Velocity contours (plane 2: Mid of the sinus); off-plane angle of: a) 0^o b) 5^o c) 10^o d) 15^o



(a) (b) (c) (d)

Figure 7. Velocity contours (plane 1: end of the sinus); off-plane angle of: a) 0^o b) 5^o c) 10^o d) 15^o



(a) (b) (c) (d)

Figure 8. Velocity contours in the medial plane of the sinus with off-plane angle of: a) 0^o b) 5^o c) 10^o d) 15^o

4.2. Vorticity contours:

To study flow abnormalities, vorticity contours are plotted for different angles, different planes (figures 9-12).in all cases, it's clearly observed that aging is causing strong changes in vorticity values. In all sinus planes, the vorticity became maximal on the internal wall and minimal on the outer walls. These values are clearly changed with the off-plane angle. The greatest changing can be observed in figure 11 that represents the second half of the sinus. To get more comprehensive picture, the vorticity contours are plotted for different angles on the frontal plane of the carotid artery (figure 12). The results of this figure show the trend of increasing the vorticity values near the internal wall, and the decreasing of these values near the outer wall. According to literature the accumulative effect of the elevated vorticity values is thought to have a major influence in the disease initiation in the sinus of the carotid artery. Lower vorticity values are indication of the longer residence time for blood constituents, like platelets and LDL, that consider the primary cells that are responsible for plaque formation. Also, the long residence of the LDL can be correlated to the injuries that could happen in the endothelial cells.

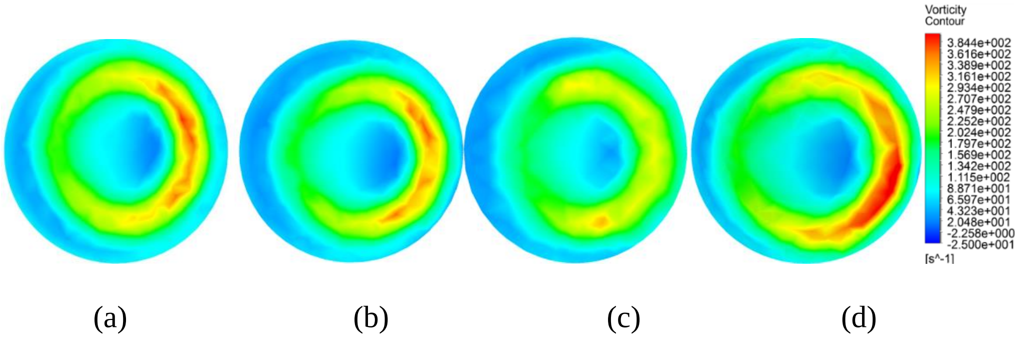


Figure 9. Vorticity contours (plane 3: start of the sinus); off-plane angle of : a) 0^o b) 5^o c) 10^o d) 15^o

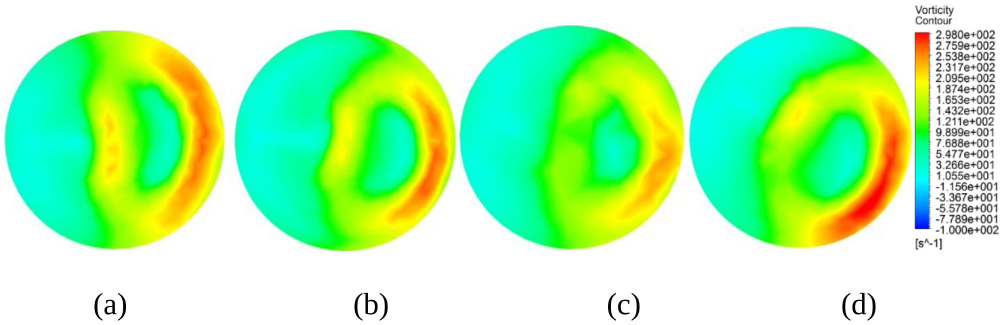


Figure 10. Vorticity contours (plane 2: start of the sinus); off-plane angle of: a) 0^o b) 5^o c) 10^o d) 15^o

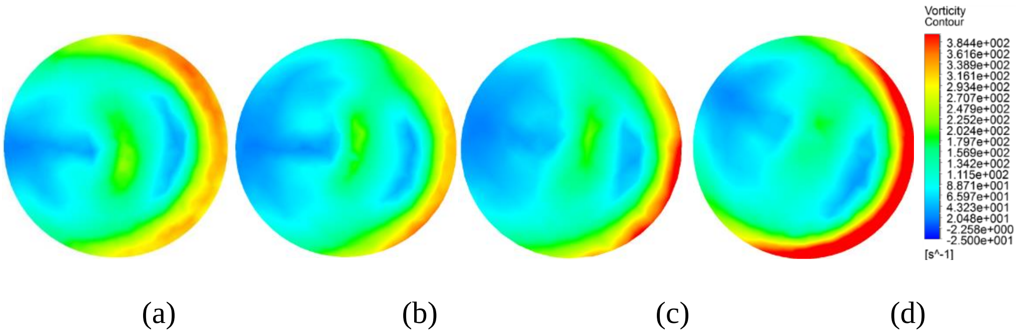


Figure 11. Vorticity contours (plane 1: start of the sinus); off-plane angle of : a) 0^o b) 5^o c) 10^o d) 15^o

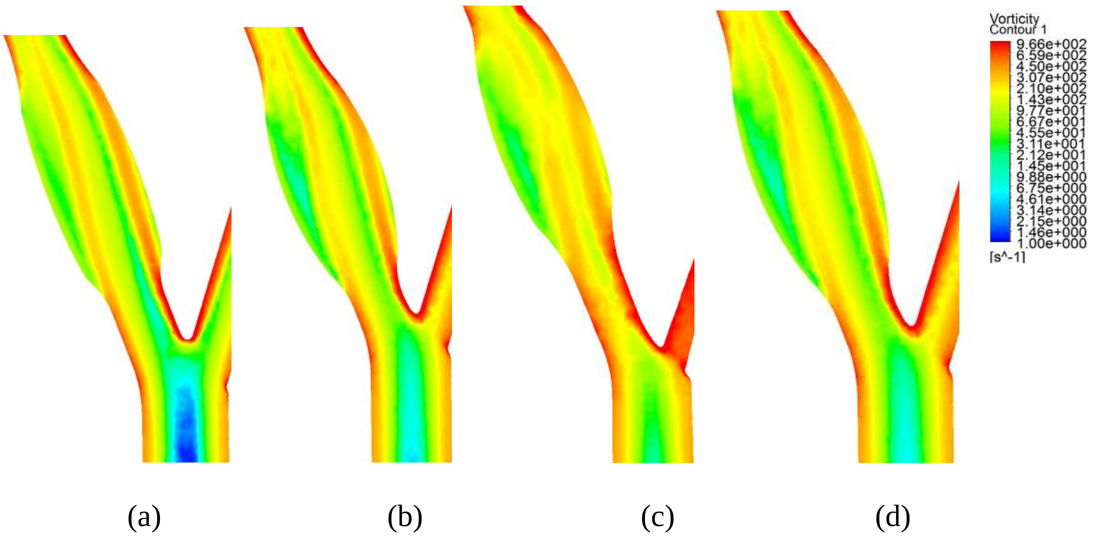


Figure 12. Vorticity contours in the medial plane of the sinus with off-plane angle of: a) 0^o b) 5^o c) 10^o d) 15^o

5. Discussion

Previous statistical studies are concluded that the coincidence of atherosclerosis, thermogenesis, and strokes is higher for elderly than younger people. Different reasons were reported as biological factors, stress factors, smoking, genetics, and mechanical factors. As the current results indicated, the unique geometry of the carotid artery is thought to be one of the key players that has major role in promoting mechanical factors that lead to the initiation of such diseases.(23) The effect of the geometry is, actually, due the direct changes that happen to the flow stream in this artery. As well known, blood related diseases are not random in the cardiovascular system. The areas with branching, curvature, and sudden area changes are more promoted to cause and create abnormal flow regimes that have direct effect on the vital cells either in the blood itself or the vessel wall.

According to previous experimental works that they used different image techniques, as the human being getting older and older three major changes can be observed in the geometry of the carotid artery; first, the radius of the common and the internal carotid arteries increases with age. Second, is that the bifurcation angle between the internal and the external carotid artery is also increasing with age. Finally, the off-plane angle of the internal carotid artery increases with age.

The current work is employed the CFD computations to study the higher coincidence of brain-blood hemodynamic diseases with aging. More focus is given to the off-plane angle of the IC. As what was expected, the results of this work improved that the normal changes in the off-plane angle leads to create more flow abnormalities with higher values. The presence of higher velocity blood stream near the apex, means higher shear stresses there, vital cells, mainly platelets and endothelial cells, are highly sensitive to such values. Consequently, the formation of abnormal plaques and atherosclerosis in such region.

In the sinus region, the formation of secondary flows and vorticities is observed for all cases, this formation is associated with separation and higher recirculation. All these abnormalities are getting higher and higher with increasing of the off-plane angle. Adjacent to the outer wall of the sinus, such flow patterns are thought to cause LDL, platelets, and other blood constituents to stay there for longer time. Thus, the exposure of such cells to abnormal shear values will cause them to get promoted, activated, and hence the formation un-preferable plaques there.

Post the formation of plaques and atherosclerosis in different regions in the carotid artery, the higher blood stream near the centerline of the sinus might cause fragmenting of small parts of the plaque. Consequently, dragging and moving small masses of lipids or other cells along with the main blood stream toward cerebral arteries in different places on the brain. Post of the sinus, the internal carotid artery will branch into smaller arterioles with small diameters. As the fragmented mass reaches smaller diameter arteriole, than its dimension, it will block this arteriole and, might cause stroke.

6. Conclusions

The current work showed that the flow complexity increases in older people who have higher internal carotid off-plane angle values. Consequently, higher probability of the formation of atherosclerosis or plaques in the sinus of the carotid artery and hence brain strokes. Referring to literature, the results of the current study showed good agreement with most of the previous experimental and computational studies that have been accomplished to date. Despite of that, more work is still needed to find a comprehensive model and understanding the mechanism of plaque formation under physiological conditions.

Finally, this study evoked many assumptions. For example, all walls were treated as rigid, the blood is treated as homogenous Newtonian fluid, and the flow regime is assumed to be laminar steady flow. In physiological conditions, the wall artery is distensible and react with blood wave through dilation and constriction. Moreover, the flow in the carotid artery is unsteady pulsatile and not steady. Thus, more studies are still needed to mimic the physiological flow regime with its values in the carotid artery.

7. Declarations

7.1. Ethical Approval and Consent to participate

The present paper does not report on or involve the use of any animal or human data or tissue

7.2. Consent for publication

The present paper does not contain data from any other works

7.3. Availability of supporting data

The data and results used and/or analyzed during the current study are available from the corresponding author on reasonable request

7.4. Competing interests

The authors declare that they have no competing interests

7.5. Acknowledgements

We would like to acknowledge The Hashemite University for its support.

7.6. Funding

This work is partially funded by the Hashemite University / Jordan.

7.7. Authors' contributions

T. A., L. K., and M. S, performed computational studies. T.A., wrote the manuscript, S. B., O.S., and S. A. revised figures, read, and approved the final manuscript.

7.8. Authors' information

None

References

[1] Chandran KB, Rittgers SE, Yoganathan, Biofluid Mechanics: The Human Circulation. CRC/Taylor & Francis: Boca Raton, 2007; 419.

[2] AlMomani T, UdayKumar H S, Marshall J S, and Chandran K B. Micro-scale Dynamic Simulation of Erythrocyte–Platelet Interaction in Blood Flow. Annals of Biomedical Engineering. 2008; 36: 905–920.

[3] Gijsen FJ1, van de Vosse FN, Janssen JD. The influence of the non-Newtonian properties of blood on the flow in large arteries: steady flow in a carotid bifurcation model. J Biomech. 1999; 32(6); 601-608.

[4] Thomas B. J, Antiga L., Susan L., Jaques S., Steinman D.A. H., Spence D., Rutt B. K., Variation in the Carotid Bifurcation Geometry of Young Versus Older Adults Implications for Geometric Risk of Atherosclerosis. Stroke. 2005; 36; 2450-2456.

[5] Samuel KC. Atherosclerosis and occlusion of the internal carotid artery. J Path Bact. LXXI. 1956; 391-401.

[6] Susan K Roepke, Elizabeth A Chattillion, Roland von Känel, Matthew Allison, Michael G Ziegler, Joel E Dimsdale, Paul J Mills Thomas L Patterson, Sonia Ancoli-Israel, Susan Calleran, Alexandrea L Harmell, and Igor Grant. Carotid plaque in Alzheimer caregivers and the role of sympatho-adrenal arousal. NIH public access. 2011; 73(2); 206-213.

[7] Chistiakov DA1, Orekhov AN, Bobryshev YV. Effects of shear stress on endothelial cells: go with the flow. Acta Physiol (Oxf). 2017 Feb;219(2):382-408.

[8] Federico Vozzi,
Jonica Campolo, Lorena Cozzi, GianfrancoPolitano, Stefano Di
Carlo, Michela Rial, Claudio Domenici, and Oberdan Parodi. Computing of Low Shear Stress-Driven Endothelial Gene Network Involved in Early Stages of Atherosclerotic Process. BioMed Research International. Sep 2018; 2018: 1-12.

[9] Gnasso A, Carallo C, Irace C, Spagnuolo V, De Novara G, Mattioli PL, Pujia A. Association between Intima-Media Thickness and Wall Shear Stress in Common Carotid Arteries in Healthy Male Subjects. Circulation. 94(12):3257-62.

[10] Shankaran H, Alexandridis P, Neelamegham S. Aspects of hydrodynamic shear regulating shear-induced platelet activation and self-association of von Will brand factor in suspension. Blood. 2002; 101(7):2637-45.

[11] Zhanjun MA, Kangquan S, Zonghuan LI, Chao I, Baiwen QI, Aixi YU. Negative pressure wound therapy promotes vessel destabilization and maturation at various stages of wound healing and thus influences wound prognosis. Exp Ther Med. 2016 Apr; 11(4): 1307–1317.

[12] Ming Chen, Baohua Huang, Fang Yang, Wei Shu, Zhenfeng Chen, Ming Chen. Functional Roles of Shear Stress in Vascular Endothelial Cells. Cellular Immunology and Serum Biology. 2017; 3(1):64-67.

[13] Zarins CK, Giddens DP, Balasubramianan K, Sottiurai V, Mabon RF, Glagov S. Carotid plaques localize in regions of low flow velocity and shear stress. Circulation. 1981; 64:44.

[14] Perktold K1, Resch M, Florian H. Pulsatile non-Newtonian flow characteristics in a three-dimensional human carotid bifurcation model. J Biomech Eng. 1991 Nov; 113(4):464-75.

[15] Karino T, Goldsmith HL. Flow behavior of blood cells and rigid spheres in an annular vortex. Philos Trans Roy Soc Lond B Biol Sci. 1977 Jun 10; 279(967):413-45.

[16] Stokholm R, Oyre S, Ringgaard S, Flaagoy H, Paaske W.P, Pedersen E.M. Determination of wall shear rate in the human carotid artery by magnetic resonance techniques. European Journal of Vascular and Endovascular Surgery 2000 Nov; 20(5):427-33.

[17] Fayad Z.A, Fuster V. The human high-risk plaque and its detection by magnetic resonance imaging. American Journal of Cardiology. 2001; 88(2A):42E-45E.

[18] Marshall I, Papathanasopoulou P, Wartolowska K. Carotid flow rates and flow division at the bifurcation in healthy volunteers. Physiological Measurement.2004; 25 (3): 691–697.

[19] Papathanasopoulou P, Zhao S, Kohler U, Robertson M.B, Long Q, Hoskins P, Xu X.Y, Marshall I. MRI measurement of time resolved wall shear stress vectors in a carotid bifurcation model, and comparison with CFD predictions. Journal of Magnetic Resonance Imaging. 2003; 17: 153–162.

[20] Younis H.F, Kaazempur-Mofrad M.R, Chan R.C, Isasi A.G, Hinton D.P, Chau A.H, Kim L.A, Kamm R.D. Hemodynamics and wall mechanics in human carotid bifurcation and its consequences for atherogenesis: investigation of inter-individual variation. Biomechanics and Modeling in Mechanobiology. 2004; 3 (1):17–32.

[21] Se Jeong Jeon, Hyo Sung Kwak, and Gyung Ho Chung. Widening and Rotation of Carotid Artery with Age: Geometric Approach. J Stroke Cerebrovasc Dis. 2018 Apr 19; 27(4):865-870.

[22] David N. Ku, Don P. Giddens, Christopher K. Zarins, and Seymour Glagov. Pulsatile Flow and Atherosclerosis in the Human Carotid Bifurcation Positive Correlation between Plaque Location and Low and Oscillating Shear Stress. Arteriosclerosis. 1985; 5 (3): 293-302.

[23] Estelle E. Seyman, Natan Bornstein, Eitan Auriel, Oren Cohen, Tania Nissel & Hen. Hallevi. Assessment of carotid artery ultrasonography in the presence of an acoustic shadow artifact. BMC Neurology 2019; 19: 178 – 186.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 149 - 154

Jordan Journal of Mechanical and Industrial Engineering

Assessment of Road Traffic Noise: A Case Study in Monastir City

Jalel Chebil[a], Jasim Ghaeb[b], Mohamed Anwer Fekih[a] and Mohamed Hadi Habaebi[c]

[a]*University of Sousse, Higher Institute of Transport and Logistics, NOCCS Laboratory, Sousse, Tunisia.*
[b]*University of Philadelphia, Department of Mechatronics Engineering, Amman, Jordan.*
[c]*Faculty of Engineering International Islamic University Malaysia, 53100 Gombak, Kuala Lumpur, Malaysia.*

Received 12 March 2019

Abstract

In the last two decades, Tunisia has witnessed a sharp increase in the number of vehicle users which gave rise to higher road traffic noise levels especially in urban areas. This type of noise causes adverse effect on humans, such as sleeping disorder, annoyance, hearing disorder, etc. Therefore, it is crucial to evaluate traffic noise levels in populated areas in order to highlight problematic areas, so that architects and engineers can find solutions to reduce noise impacts. This paper presents a case study of traffic noise levels at four main roads in the city of Monastir- Tunisia and it also investigates the performance of traffic noise models. In the first part, it is found that the measured noise is mostly due to traffic noise and its levels exceeded the limits recommended by both the World Health Organization (WHO) and the Tunisian environmental standards. In the second part, the study compares the performance of five known traffic noise models: Griffith and Langdon, Burgess, French CSTB (Centre Scientifique et Technique du Batiment), Fagotti and the Italian CNR (Consiglio Nazionale delle Ricerche). It found that the French CSTB model outperforms the others when applied to the measured sites.

Keywords: Road traffic noise; traffic noise; equivalent sound level; statistical noise level; noise exposure level;

1. Introduction

Road traffic noise has been identified as one of the main factors in the increase of noise pollution in urban areas and can cause serious degradation of quality of life [1, 2]. This problem is becoming more significant due to the continuous increase in traffic volumes and the use of heavy vehicles inside urban roads. In addition, the inadequate urban planning of many cities, especially in third world countries, makes things worse since many residential areas and business centers are built close to the main roads of the metropolitan without buffer zones or adequate sound proofing [3]. The harmful impacts of road traffic noise exposure are well-known [4]. It can include productivity losses due to poor concentration, interference with communication, reduction in enjoyment of activities. Moreover, there are psychological and physiological health effects, such as nuisance, loss of well-being, sleep disturbance, fatigue and hearing impairment, etc [4]. Many research studies have been conducted on road traffic noise in many countries [5- 17]. However, very few studies have investigated noise pollution in Tunisia, and none was done in the city of Monastir which is a known coastal city. Monastir is a medium size city with growing population from 67,730 in 2002 to 93,300 in 2014 [18-19]. It has expanded continuously in all directions in the last three decades. The city has been subjected to daily continuing road traffic and commercial activities due to overall increase in prosperity, fast development, and an increase in the economy growth. Therefore, it is important to assess the noise level in this city. This research conducted a case study at four main roads in the city of Monastir- Tunisia where the traffic noise levels were collected and analyzed. The study also investigates the performance of five common models that are for traffic noise estimation.

The paper is organized as follows. Section 2 introduces some background information about noise measurements and traffic noise descriptors and it also presents a brief description of five well-known traffic noise models. The methodology used in this study is described in Section 3 and the results are discussed in Section 4. Finally, the last section concludes the paper.

2. Background Information

Noise is defined as an unwanted sound that may interfere with normal activities [20]. It may also cause annoyance and harm to human beings [3]. However, traffic noise is due mainly to the sound produced by vehicles' engine, exhaust, and tires. Engine and exhaust noise are usually dominant if vehicle speeds are under 50 km/hr, otherwise, the tire noise is louder [20]. Thus, the traffic noise in highways is generally dominated by tire noise, while the traffic noise in local roadways is typically

* Corresponding author e-mail: chebil8@hotmail.com

dominated by engine and exhaust noise. In what follows, noise measurements and traffic noise descriptors were discussed. It is proceeded by presenting the Tunisian standards for noise levels. Finally, the five road traffic noise models that are used for comparison were described, which are namely: Griffith and Langdon, Burgess, French CSTB (Centre Scientifique et Technique du Batiment), Fagotti, and the Italian CNR (Consiglio Nazionale delle Ricerche).

2.1. Noise measurements and traffic noise descriptors

Noise is measured in terms of sound pressure using logarithmic scale with units of decibel (dB) since the range that the human listeners can detect is very wide [3]. For traffic noise studies, it is more common to use the A-weighted decibel (dBA) as the unit of measure in order to best approximate the way the average person hears sounds. Thus, the dBA scale reduces the strength of very low and very high audio frequencies which are not audible by humans [20]. Noise levels can range from 0 dBA, barely audible, to about 120 dBA which is painful to human ears. The typical range for human speech is the mid 60 dBA range while quiet suburban environments are often in the 40 to 50 dBA range [21]. Another feature of the measured noise is its fluctuation with time since it is never constant. To address the changes of noise over time, traffic noise is assessed using the equivalent sound level (L_{Aeq}) which is defined as the average noise level over a given period of time and it is measured in dBA [21]. When noise measurement is taken over a time period, it is common practice to add the measurement duration to the subscript. For example, the symbol $L_{Aeq,1}$ is used to denote a 1-hour L_{Aeq} measurement [4]. It is also possible to express the noise levels during times of the day, such as daytime (L_D), evening-time (L_E) or night-time (L_N). In this paper, the daytime is considered from 7:00 to 20:00, the evening time is from 20:00 to 22:00 and the night-time is from 22:00 to 7:00. Based on [13], the equivalent noise level for day, evening and night is denoted by L_{DEN} and is defined by

$$L_{DEN} = 10 \times log_{10}\left(\frac{1}{24}(13 \times 10^{L_D} + 2 \times 10^{L_E+5} + 9 \times 10^{L_N+10})^{1/10}\right) \quad (1)$$

It is the average noise exposure throughout a 24-hour day, and it can be used to reasonably represent the correlation of the community responses [12].

Another useful set of parameters for the traffic noise are the n-percent noise exceeded level, L_n, which are known as the statistical noise levels. L_n is defined as the sound pressure level exceeded for n percent of the time, and the widely used statistical noise levels are L_{10}, L_{50} and L_{90} [22]. L_{10} and L_{90} are commonly referred to as the average maximum level and the background level respectively [4]. L_{90} is also referred to an indicator of long-lasting noise in a specified area [12]. In addition, many researchers make use of the maximum and the minimum detected sound levels, L_{max} and L_{min}, in their analysis [12, 13].

2.2. Standards for noise levels

Several countries have established standards for noise levels that determine the maximum permissible limits for equivalent sound levels in different areas and for different time periods. Based on the standards adopted by the municipality of Tunis, the capital city of Tunisia [23], Table 1 shows the maximum permissible limits for selected areas. These limits are compatible with the recommendations of World Health Organization (WHO) [24] and will be used in the analysis.

Table 1. Maximum permissible noise limits in Tunis for selected areas [23]

Zones	Permissible limit for L_{Aeq} in dBA		
	Day	Evening	Night
Hospitals and rest areas	45	40	35
Suburban residential areas with low road traffic.	50	45	40
Urban residential areas	55	50	45
Urban or suburban residential areas near important roads and with some workshops, business center and shops	60	55	50
Areas with commercial activities.	65	60	55
Areas with predominantly heavy industry.	70	65	60

2.3. Road traffic noise models

Many models were proposed for the estimation of road traffic noise in different environments. In this study, the performances of five well known models are compared based on road traffic noise measurements. These models include: Griffith and Langdon, Burgess, CSTB, Fagotti and CNR. This section presents briefly the expression for each model.

The Griffiths and Langdon model [25, 26] proposed the evaluation of equivalent level starting from the percentile level as follow:

$$L_{eq} = L_{50} + 0.018(L_{10} - L_{90})^2 \quad (2)$$

where the statistical percentile indicator are expressed as:

$$\begin{cases} L_{10} = 61 + 8.4\log Q + 0.15\,P - 11.5\log d \\ L_{50} = 44.8 + 10.8\log Q + 0.12\,P - 9.6\log d \\ L_{90} = 39.1 + 10.5\log Q + 0.06\,P - 9.3\log d \end{cases} \quad (3)$$

Q is traffic volume in vehicles per hour and d is the distance from observation point to center of the traffic lane, in feet; and P is the percentage of heavy vehicles.

The Burgess model [25, 27] uses the following expression for the sound level

$$L_{eq} = 55.5 + 10.2\log Q + 0.3P - 19.3\log d \quad (4)$$

The CSTB model [25], suggested a predictive formula based on the average acoustic level (L_{50}) with the following expression:

$$L_{eq} = 0.65L_{50} + 28.8 \quad (5)$$

where the value of L_{50} is given by:

$$L_{50} = \begin{cases} 11.9\log Q + 31.4 \text{ , for urban road and highway with Q<1000 vehicles/hour} \\ 15.5\log Q - 10\log L + 36 \text{ , for urban road with buildings near the edge} \end{cases} \quad (6)$$

where L is the road width, in meters, near the measurement point.

Fagotti *et al.* proposed the following formula [25]

$$L_{eq} = 10\log(Q_L + Q_M + 8Q_P + 88\,Q_{Bus}) + 33.5 \quad (7)$$

where Q_L, Q_P,Q_P and Q_{Bus} are the traffic volumes of the light and heavy vehicles, motorcycles and buses respectively

For the CNR model, the equivalent sound level in dBA is given by [25]

$$L_{eq} = \alpha + 10\log(Q_L + \beta\,Q_P) - 10\log\left(\frac{d}{d_0}\right) + \Delta L_V + \Delta L_F + \Delta L_B + \Delta L_S + \Delta L_G + \Delta L_{VB} \quad (8)$$

where Q_L and Q_P represent the traffic flow in one hour for light and heavy vehicles respectively, d_0 is a reference distance of 25 meter and d is the distance between the lane center and the observation point on the road's edge. The parameters ΔL_V, ΔL_F, ΔL_B, ΔL_S, ΔL_G and ΔL_{VB} represent the correction factors which depend respectively on: mean flux velocity, presence of reflective façade near the observation point, or in opposite direction, types of pavement, road's gradient and presence of traffic lights or slow traffic. These correction factors are well defined in [25].

3. Methodology

Traffic noise measurements were carried out along four main important roads in Monastir: Avenue of 14 January 2011, Taieb Mhiri Avenue, Environment Avenue and Supreme Fighter Avenue. These roads cross at the roundabout El-Nafoura as shown in Figure 1, and they lead to many public buildings, government and company offices. Hence, it features high traffic flow during peak hours, which causes significant increase in noise emissions. In this study, zone 1 refers to the Supreme Fighter Avenue which extends for about 600 m from the El-Nafoura roundabout to the next main cross section. Similarly, zone 2, zone 3 and zone 4 refer respectively to the Avenue of 14 January 2011, Environment Avenue and part of Taieb Mhiri Avenue. These zones extend for 600m, 3 km and 300m respectively.

The measurements of the A-weighted instantaneous sound pressure level were made at one spot in each zone as shown in Figure 1. The instrument used is a programmable sound level meter of type TES-1352H made by the TES Electrical and Electronics Corporation. It was held at a height 1.5 m above the ground and with 2 m away from the side of the road and at a distance not less than 1 m away from any other reflecting object. In addition, the microphone was pointed at the circulating vehicles and the instrument was set on the *A* frequency weighting and slow time weighting. All measurements were done during working days and under acceptable weather conditions in February 2017. The data are recorded at intervals of one second and for 24 hours period for each road. For this study, the equivalent sound level is computed for every hour at each location in order to compare the noise variation during one full-day. In addition, the noise indicators L_D, L_E, L_N and L_{DEN} are calculated along with L_{10}, L_{50} and L_{90}.

In order to compare the performance of the five traffic noise models discussed earlier, measurements of the traffic volumes, *Q*, and percentage of heavy vehicles, *P*, were also carried out along the four zones for two periods of time: 11:00-12:00 and 19:00-20:00. These measurements were conducted at a distance *d*=5m from the center of the roads.

Figure 1. Map showing the four zones in the city of Monastir where the measurement were conducted [28].

4. Results and Discussion

The first part of this section presents the results obtained from the measured data and analyzes the computed traffic noise level statistics. In the second part, it investigates the performance of the five road traffic noise models based on the measured data.

4.1. Traffic noise level statistics

The 24-hour noise levels measured in the four zones are shown in Figure 2 along with the noise limits set by the Tunisian environmental standards. The figure shows that the noise levels are higher during the day specially from 07:00-8:00 and lower during the night. This is expected since most people go to work during the day and that there are less activities and mobility during the night. It is also observed that the measured noise level is higher than the Tunisian standards except during early morning around 2:00-4:00. The highest noise level is measured in zone 3 where it exceeds 75 dBA from 7:00-14:00. This is due to higher traffic density with vehicles exceeding an average speed of 40km/h in zone 3. However, vehicles in zones 1, 2 and 4 cannot exceed an average speed of 30 km/h since these roads have short distances less than 600 m and there are many shops along these roads. That explains why less noise is observed in these streets although the traffic is sometimes comparable to zone 3.

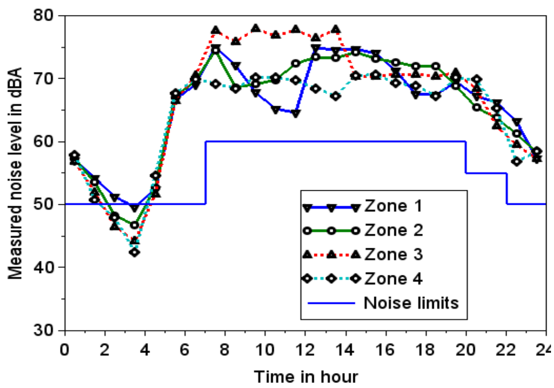


Figure 2. The plots of 24-hour noise levels at four streets in Monastir compared with Tunisian Environmental standards

The noise levels L_D, L_E, L_N, and L_{DEN} are shown in Table 2 for the four zones. For the daytime noise level, L_D varies from 69.4 to 75.3 dBA and is higher than the permissible level by 9.4 to 15.3 dB. It is also observed that L_D is higher in zone 3 than in the other zones. This is due to higher traffic densities and higher spatial velocity in zone 3. For the evening time, L_E varies from 64.7 to 68.1 dBA and is higher than the permissible level by 9.7 to 13.1 dBA. This variation is lower than the daytime since less traffic is observed. For the night time, L_N has almost the same value of 63 dBA in the four zones and it is higher than the permissible level by 13 dBA. This is an indication that the traffic density is quite low during the night and its impact is minimal. In general, the measured noise levels near these four streets exceed the permissible noise limits by at least 9 dBA but not more than 15.3 dBA. This might affect people living in the buildings and shops along these streets. In addition, the measured noise levels in Table 2 show clearly that zone 3 has the highest noise level during the day. During the evening, zone 4 has the highest level whereas during the night the noise levels are close in all four zones. The day-evening-night values, L_{DEN} varies from 71.4 to 74.3 dBA and it shows that the average noise exposure throughout a 24-hour day is higher in zone 3. These values are considered lower in comparison with the measured noise levels at many cities such as Cairo [13], Islamabad [13] and IKEJA [17].

Table 2. Computed values of L_D, L_E, L_N and L_{DEN} in the four zones in dBA

	L_D	L_E	L_N	L_{DEN}
Zone 1	72.0	66.7	62.6	72.2
Zone 2	72.2	64.7	62.9	72.3
Zone 3	75.3	66.4	63.0	74.3
Zone 4	69.4	68.1	63.0	71.4
Standard levels	60	55	50	62.14

The commonly used statistical noise levels L_{10}, L_{50} and L_{90} are computed for one-hour period during the afternoon and the results are listed in Table 3. For all these levels, zone 3 has the highest values and the sound levels exceed 81, 77.8 and 74.1 dBA for 10%, 50% and 90% of the time respectively. To improve the environmental conditions in zone 3, it is recommended to introduce some restrictions, such as lowering the speed limit, and banning horns and trucks.

Table 3. Statistical noise levels computed for one-hour period for the four zones

	L_{10}	L_{50}	L_{90}
Zone 1	77.9	74.1	70.9
Zone 2	76.9	72.2	67
Zone 3	81	77.8	74.1
Zone 4	74.1	69.8	66.2

4.2. Comparison between five models for the road traffic noise levels

This section compares the performance of five traffic noise models discussed in Section 2. For this purpose, the Root Mean Squared Error (RMSE) is used as a criterion for selecting the model that produces minimum RMSE. Based on the definition presented by [29], the RMSE can be expressed as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(L_{eq} - \hat{L}_{eq}\right)^2} \quad (9)$$

where L_{eq} and $\hat{L}_{eq}$ are the measured and the estimated road traffic noise levels respectively. The number of measurements is denoted by n. In order to evaluate this equation, measurements of Q, P and L_{eq} were conducted at a distance $d=5m$ from the centre of each road for two periods of time and the outcome is shown in Table 4. The RMSE is computed for each model using the corresponding equations from (2) to (8) and the measured values of Q, P and d. The result is shown in Table 5 which clearly indicates that the CSTB model outperforms the other four models. Therefore, the CSTB model can be used to estimate road noise levels in Monastir based on the available traffic data. This means, noisy areas can be identified in this city and the relevant experts can find solutions to reduce the impact of road traffic noise.

Table 4. Measured values of *Q*, *P* and L_{eq} at each zone for two periods of time.

	Time	*Q* (Veh/h)	*P* (%)	Measured L_{eq}(dBA)
Zone 1	11:00-12:00	660	1.6	74.98
	19:00-20:00	480	0.8	69.5
Zone 2	11:00-12:00	690	11.3	73.96
	19:00-20:00	510	10.6	68.79
Zone 3	11:00-12:00	830	23.4	78.65
	19:00-20:00	640	19.3	70.92
Zone 4	11:00-12:00	740	18.7	70.92
	19:00-20:00	580	15.4	70.13

Table 5. Computed RMSE for the five traffic noise models

	Griffith & Langdon	Burgess	CSTB	Fagotti	CNR
RMSE	2.47	1.94	1.23	3.40	2.73

5. Conclusion

This study conducted measurement of traffic noise in four main roads in the city of Monastir. It is found that the noise levels are generally higher than the limits set by the Tunisian environmental standards and the WHO. In addition, the measured noise levels show clearly that zone 3 and zone 4 have the highest noise level during the day and during the evening respectively, whereas the noise levels in all four zones have close values during the night due to extremely low traffic volume. It is recommended to introduce some restrictions, such as lowering the speed limit, and banning horns. The study also compares between five known traffic noise models and found that the CSTB model outperforms the other models. As a result, this model can be used to estimate road noise levels in the city using road traffic data.

Acknowledgment

The authors are grateful to the Higher Institute of Transport and Logistics of Sousse for supporting this project.

References

[1] D. Piccolo, G. C. Plutino, Evaluation and analysis of environmental noise of Messina, Italy, Applied acoustics, No. 66 (2005), pp. 447-465.

[2] S. I. Korfali, , and M. Massoud, Assessment of community noise problem in Greater Beirut Area, Lebanon, Environ. Monit. Assess. 84 (2003), pp. 203–218. https://doi.org/10.1023/A:1023322507415.

[3] D. Pal and D. Bhattacharya; Effect of road traffic noise pollution on human work efficiency in government offices, private organizations, and commercial business centres in agartala city using fuzzy expert system: a case study; Advances in Fuzzy Systems, Volume 2012 (2012), http://dx.doi.org/10.1155/2012/828593.

[4] Austroads, Modelling, Measuring and Mitigating Road Traffic Noise, Report No. AP-R277/05, Austroads Incorporated, Sydney, 2005.

[5] A.I. El-Sharkawy, and A. A. Aboukhashaba, Traffic Noise Measurement and Analysis in Jeddah, Applied Acoustics, Vol. 16 (1983), pp. 41-49.

[6] Bjorkman M., Maximum Noise Levels in Road Traffic Noise, Journal of Sound & Vibration, 127(3) (1988), pp. 583-587.

[7] Brown A.C. and Cliff D. I., National Exposure to Road Traffic Noise, Acoustics Australia, Vol. 17 (1) (1989), pp.7-11.

[8] Hammad RNS and Abdel Aziz M.K., Measurements and Analysis of the Traffic Noise in Amman, Jordan and Its Effects, Applied Acoustics, 21 (1987), pp. 309-320.

[9] Suksaard Thanaphan, Road Traffic Noise Prediction Model in Thailand, Internoise Congresses 94, Yokhana-Japan., August 1994.

[10] D. Chakrabarty, S.C. Santra, A. Mukherjee, B. Roy and P. Das, Status of road traffic noise in Calcutta metropolis-India, The Journal of the Acoustical Society of America, 101 (1997), pp. 943-949. https://doi.org/JASMAN101.

[11] Onuu, M. U., Road Traffic Noise in Nigeria: Measurements, Analysis and Evaluation of Nuisance, Journal of Sound and Vibration, 233(3) (2000), pp. 391-405.

[12] L. Dai, J. Cao, L. Fan and N. Mobed, Traffic Noise Evaluation and Analysis in Residential Areas of Regina, Journal of Environmental Informatics, 5(1) (2005), pp. 17-25.

[13] Zekry F. Ghatass, Assessment and Analysis of Traffic Noise Pollution in Alexandria City- Egypt, World Applied Sciences Journal, 6 (3) (2009), pp. 433-441.

[14] Salem Dahech, Fethi Rekik, Trafic routier et pollution sonore à Sfax (Tunisie méridionale) : étude pluridisciplinaire, Pollution Atmospherique, July 2012.

[15] F. Uysal and E. Tinmaz, Evaluation and analysis of traffic noise from the main roads in the Trakya region of Turkey, WIT Transactions on The Built Environment, Vol. 77 (2005), pp. 479-486.

[16] M. Abdmouleh and S. Dahech, Répartition spatiale de la pollution sonore dans l'agglomération de Sfax de la mesure à la modélisation, Journées Interdisciplinaires de la Qualité de l'Air, 10 - 11 février, Villeneuve d'Ascq, 2014. http://www.jiqa.fr/doc/2014 /Article/ABDMOULEH.pdf.

[17] O. O. Oluwasegun,; M. U. Onuu, and O. E. Oyenekan, Study of Road Traffic Noise Pollution and Impacts on Residents of Ikeja Local Government Area of Lagos State, Nigeria, International Journal of Scientific & Engineering Research, 6 (5) (2015), pp.1108-1117.

[18] Wikipedia, Monastir-Tunisia, https://en.wikipedia.org/wiki/Monastir,_Tunisia, 2017 (accessed 21.10.17).

[19] Mattar, Philip, Encyclopedia of the Modern Middle East and North Africa, Macmillan Library Reference; 2nd Edition, 2004.

[20] Arizona Department of Transportation information planning, Noise Information Sheet Facts About Traffic Noise, https://www.azdot.gov/docs/default-source/planning/ noise_facts_about_noise.pdf, 2017 (accessed 10.02.17).

[21] Oregon Department of Transportation, Chapter 3 Fundamentals of Highway Traffic Noise, http://www.strongsville.org/content/documents/ODOT_Ch_3 _Fundamentals_of_Highway_Traffic_Noise.pdf, 2017 (accessed 10.02.17).

[22] Environmental Protection Department of Hong Kong, Noise Assessment - Assessment Methodologies", http://www.epd.gov.hk/epd/noise_education/web/ENG_EPD _HTML/ m2/types_4.html, 2017 (accessed 30.09.17).

[23] Municipality of The city of Tunis, **Lutte contre les nuisances sonore,** http://www.commune-

tunis.gov.tn/publish/content/article.asp?id=173), 2017 accessed (19.07.17).

[24] Berglund, T. Lindvall, and D. H. Schwela, Guidelines for Community Noise, World Health Organization (WHO), Geneva, Switzerland, 1999.

[25] J. Quartieri, Nikos Mastorakis, Gerardo Iannone, etal. A Review of Traffic Noise Predictive Noise Models, Conference: Recent Advances in Applied and Theoretical Mechanics, January 2009.

[26] D. Griffiths and F. J. Langdon (1968), Subjective Response to road traffic noise, Journal of Sound and Vibration 8, 16-32.

[27] M.A. Burgess, Noise prediction for Urban Traffic Conditions-Related to Measurement in the Sydney Metropolitan Area, Applied Acoustics, vol. 10, pp 1-7, 1977.

[28] WorldMapFinder, Map of Monastir, http://www.worldmapfinder.com/BingMaps/Ar_Africa_Tunisia_Monastir.html, 2017 accessed (12.11.17).

[29] Stefanie Jachner, K. Gerald van den Boogaart and Thomas Petzoldt. Statistical Methods for the Qualitative Assessment of Dynamic Models with Time Delay (R Package qualV), September 2007, Volume 22, Issue 8.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 155 - 160

Jordan Journal of Mechanical and Industrial Engineering

Enhanced Boiling Heat Transfer on Surfaces Covered with Microstructural Mesh Coatings

Łukasz J. Orman

Faculty of Environmental, Geomatic and Energy Engineering, Kielce University of Technology, al. Tysiaclecia P.P.7, 25-314 Kielce, Poland

Received 3 September. 2019

Abstract

The paper deals with the issue of pool boiling heat transfer enhancement. It presents the experimental test results of distilled water and ethanol boiling at ambient pressure on heater surfaces covered with microstructural porous copper coatings. The coatings are made of 1 to 5 layers of fine metal meshes sintered together to form a porous structure. The obtained values of the heat transfer coefficient from these surfaces were compared to the smooth surface test results. A considerable enhancement of heat transfer due to the application of the analyzed additional coatings was recorded, especially in the area of low superheat values.

Keywords: boiling heat transfer, enhancement, porous coatings;

1. Introduction

Boiling heat transfer is a phase – change phenomenon. It is highly efficient in dissipating significant heat fluxes at low temperature differences. Its applications in engineering and technology are wide, and they cover areas of refrigeration, electronics cooling, heat pipes, among many other usages. In the design of power engineering and energy systems, there is currently a growing need to search for more effective methods of the removal of heat. One of the techniques to intensify the heat exchange in phase change processes (boiling, condensation) is the application of additional coatings on heat exchangers. In the case of boiling, such structures could be made in the form of metal wire meshes, fibrous layers, flame spraying coatings and others. They might considerably enhance boiling heat transfer. Compared to the smooth surface, without any coating, heat flux values for the same superheat could be even several times higher. Up to now there exists no efficient model of boiling heat transfer on structural coatings that could successfully predict the performance of a coated surface based on the physical and chemical parameters of the system. Similarly, data on the impact of structural parameters of the coatings found in literature is sometimes incongruent or even contradictory. Although it is generally stated that the application of microstructures enhances heat transfer, some scientific reports indicate the opposite effect for some coatings. Moreover, typically data available in literature focuses on isothermal surfaces, while non – isothermal heaters in the form of a fin are very rarely considered. In the present work an analysis of the thermal performance of a non – isothermal heater covered with up to five mesh layers will be done.

Tsay et al. [1] presented experimental results of water boiling on a horizontal smooth and rough surface covered with a single stainless-steel mesh layer. The authors reported that the application of a single mesh layer led to enhanced heat transfer compared to the smooth reference surface at superheats (which is a difference between the surface temperature and the saturation temperature) above 6 K. The visualization studies enabled to conclude that the number of bubbles grown on the heater covered with the coating was higher than on the smooth reference surface. Brausch and Kew [2] provided test results of water boiling at ambient pressure on vertical surfaces with stainless steel mesh wicks. The porous coatings had one, three and five layers of mesh. The application of a single layer of mesh led to increased heat fluxes in relation to the smooth surface at low superheats. While at higher superheats more heat was dissipated from the smooth surface. This phenomenon was explained by the authors by the coalescence of vapor bubbles within the porous structure, which creates local dry out and the creation of an insulating film. This film decreases the heat transfer coefficient. Additional mesh layers increased thermal resistance of the microstructure and reduced the value of the heat transfer coefficient. Gerlach and Joshi [3] investigated boiling of PF 5060 dielectric liquid at atmospheric pressure on a horizontal surface with single copper and bronze meshes attached to a copper block by soldering. The tests were conducted for the cases when the mesh was open to a liquid pool and when the structure was covered with a plate and bubbles exited sideways. The application of the porous coating as it was open to the pool

* Corresponding author e-mail: orman@tu.kielce.pl

generally led to an increase in heat flux in comparison with the smooth surface. However, when the meshes were covered with the plate, a lower heat transfer than in pool boiling was recorded. The authors state that the results published by others might be contradictory, possibly due to different clamping methods as well as variation in spacing and contact at the mesh to the heater surface. Franco et al. [4] conducted experimental tests of pool boiling of the dielectric refrigerant R141b. The heating surface was covered with copper, aluminum, brass and stainless-steel mesh structures. It was found that mesh layers of small aperture enhance heat transfer comparing to the smooth surface for small heat fluxes. According to the authors, heat transfer is influenced by the structure height (number of layers). On the one hand, they increase active nucleation sites density, but on the other the hydraulic resistance of vapour flow rises. However, according to the experimental investigations by Rannenberg and Beer [5] on R11 and R113 boiling on surfaces covered with two to nine layers of mesh, there is no apparent impact of the number of layers on heat transfer. Similar findings were presented by Smirnov and Afanasiev [6].

Wire mesh structures can be used as an internal coating in heat pipes. Wong and Kao [7] analyzed performance of heat pipes with a two – layered mesh wick covering under water evaporation/boiling conditions. It was found that the coarse mesh nucleate boiling was absent at low heat loads while the heat was dissipated by surface evaporation on the water–vapor menisci within the wires. The fine mesh was reported to provide more nucleation sites. Investigations of multi-layer copper meshes sintered to the heat pipe surface were also performed by Liou et al. [8] with water as the working fluid. A development regarding the use of meshes has been recently proposed by Pastuszko [9], who investigated structures made of microfins and meshes sintered to them for further boiling heat transfer enhancement. The works of the author also confirm significant possibilities of boiling heat transfer enhancement with the use of microstructural coatings (for example [10, 11]).

It needs to be noted that the performance of microstructure coated heaters is tested on the isothermal surfaces. However, the use of the thermovision technique enables to perform tests on the non-isothermal surfaces of fins (which are more often found in practical applications). This measuring technique is only used at Kielce University of Technology.

2. Material and method

The tests have been designed to analyze the boiling performance of heaters covered with one to five mesh layers made of pure copper. The meshes have been joined together with the base surface (in form of the fin) using the sintering technology in order to produce durable bonds. The sintering temperature was ca. 900 ^{o}C and it occurred in the reduction atmosphere to prevent oxidation of the samples. The research has been done with two boiling agents: distilled water and ethanol (99.8% purity) under ambient pressure. The coatings have been located on the fin's surface of 4 mm thickness, 12 mm height and 90 mm length. The porosity of microstructures ranged from 51% to 73%, while their height from 0.4 mm to 1.25 mm.

The experimental analysis has been performed on the stand presented in Fig. 1. The main element of the experimental set-up is a copper fin placed within the wall of the vessel. On one side it is in contact with the boiling liquid. This side of the fin was earlier covered with the porous microstructure. On the other side the fin is open to the atmosphere and observed with the long-wave (8 – 14 μm) thermovision camera. In order to ensure correct temperature recordings, the fin is painted black on this side. Before the actual measurements, the emissivity of the black paint was determined on a separate experimental stand and amounted to 0.97. The stand consisted of the electrically heated surface covered with the investigated paint and observed with the thermovision camera. The temperature of the heater was determined with the K-type thermocouple located within the surface. The measurements were done with increasing heat flux at different temperatures (the temperature ranges reflected those encountered during the actual measurements on the fin). The discussion of the measurement errors during infrared testing can be found in the work by Orzechowski and Orman [12]. Considering the accuracy of the temperature readings, the errors of determination of the heat transfer coefficient using the described technique have been assessed to amount to 8%.

Heat is supplied to the base of the fin with an electric heater. Consequently, a temperature gradient along the element is created. The thermovision camera is applied to record the temperature along the fin. The measured temperature distribution is then used to determine local values of the heat transfer coefficient. An auxiliary heater (a resistance wire) is installed at the bottom of the vessel to maintain stable pool boiling conditions. The boiling process takes place on the inner side of the fin inside the vessel. This part is covered with the microstructure. A detailed description of the apparatus and the testing method used in the present study has been given in [13].

The obtained temperature distribution along the fin is numerically differentiated to determine local values of the heat transfer coefficient according to the method presented by Orzechowski [13]. Based on the experimental reports of pool boiling heat transfer on microstructure coated surfaces (e.g. [5, 15], it has been assumed that the heat transfer coefficient (α) between the fin's surface and the liquid in the vessel depends exponentially on superheat (Θ) according to the following equation:

$$\alpha = a\,\theta^{n} \tag{1}$$

where a, n are constants. Their experimental determination leads to a formula for the boiling curve as a dependence of the heat transfer coefficient or heat flux. The values of n have been found to differ depending on the kind of the microstructural coating and the range of superheat under investigation (e.g. [5. 15]). This concept of the exponential dependence is based on experimental observations of pool boiling heat transfer phenomenon.

Having considered (1) and the simplifying assumptions for one - dimensional heat conduction, a formula for temperature distribution along the fin is produced:

$$\frac{d^2\theta}{dx^2} = m^2\,\theta^{n+1} \tag{2}$$

This issue was analyzed by Ünal in [16].

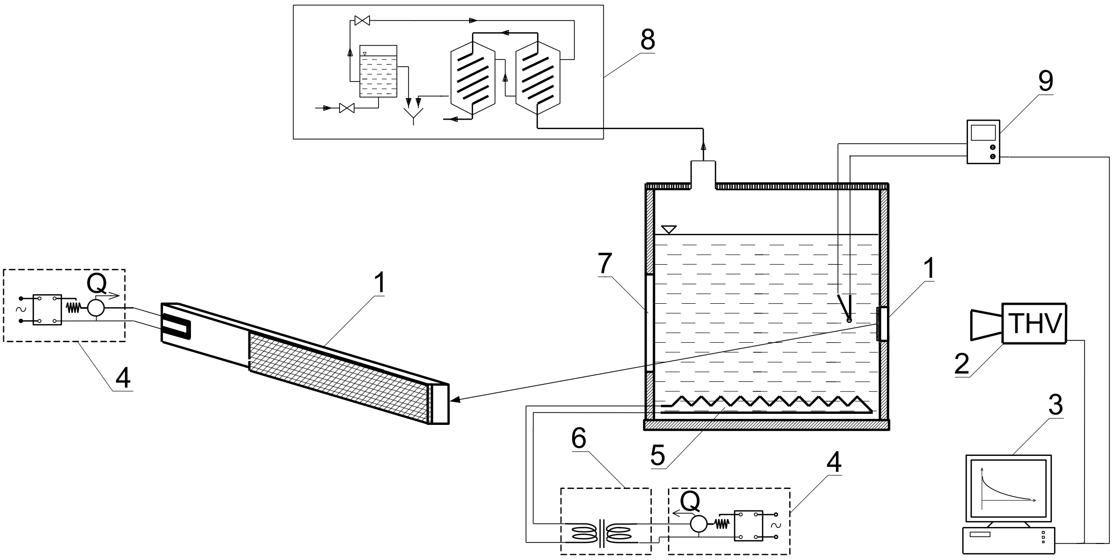


Figure 1. Schematic of the experimental set - up:

1 – copper fin with the porous microstructure, 2 – thermovision camera, 3 – data acquisition unit, 4 – autotransformer, 5 – auxiliary heater, 6 – electrical current separation unit, 7 – observation window, 8 – cooling and condensate retrieval unit, 9 – temperature measuring device [14]

After differentiation of equation (2) the superheat gradient in logarithmic coordinates is calculated:

$$\ln\left(\frac{d\theta}{dx}\right)^2 = \ln\left(\frac{2m^2}{n+2}\right) + (n+2)\ln\theta \tag{3}$$

where $n \neq 2$ and m^2 is defined in the following form:

$$m^2 = \frac{aP}{\lambda F} \tag{4}$$

In the above equation P and F are the circumference and surface area of the fin, respectively. λ is the thermal conductivity of the material of the fin. For fins of considerable length, as in the analyzed case, no heat transfer at the tip can be assumed. In such a case an integration constant in (3) was 0.

The temperature distribution along the fin obtained with the thermovision camera and its numerical differentiation enabled to determine the constants: a and n from (3). It is done by applying the linear fitting to research data. Consequently, with the use of equation (1) it is possible to draw boiling curves which are a function of the heat transfer coefficient or heat flux vs. wall superheat. In order to more precisely determine the heat transfer coefficients for ethanol the measurement results have been analyzed assuming the non - linear dependence for the heat transfer coefficient (as proposed by Orzechowski [17]. The applied data acquisition technique based on thermovision measurements for the analyses of non – isothermal surfaces in the boiling mode is used only at Kielce University of Technology [18].

3. Results and discussion

The tests have been conducted for distilled water and ethanol as boiling liquids at atmospheric pressure. Before the measurements the surface of the fin was covered with the special black paint of high emissivity. This side of the heater was observed with the thermovision camera and temperature distributions were recorded. The inside part of the fin was located inside the vessel were boiling took place. The recorded temperature gradients along the x axis of the horizontally located fin made it possible to provide data on thermal performance of the heaters with different porous coatings. Fig. 2a and 2b present the test results for two fluids and different numbers of layers as the dependence of the superheat gradient vs. superheat.

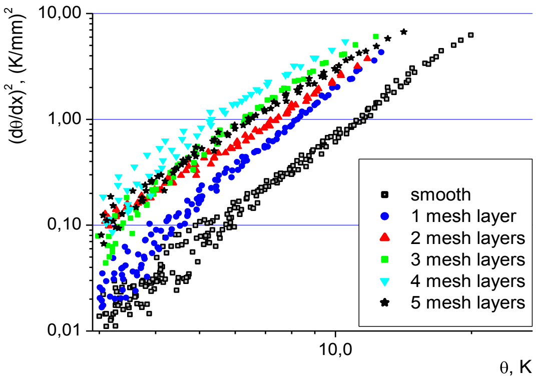


Figure 2a. Superheat gradient vs. wall superheat for distilled water: 1 – 5 mesh layers and the smooth surface test results.

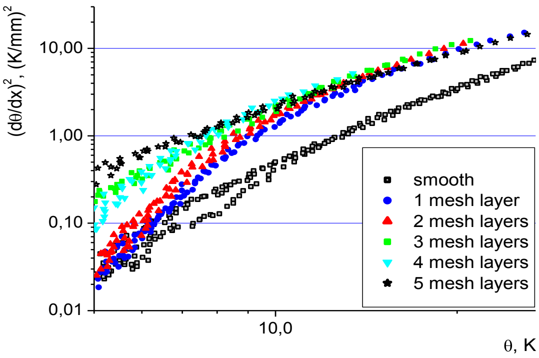


Figure 2b. Superheat gradient vs. wall superheat for ethanol: 1 – 5 mesh layers and the smooth surface test results.

The analysis of the above figures indicate that the performance of the meshed surfaces improves as the number of meshes increases. The best performance for distilled water has been recorded for the four mesh layers, while in the case of ethanol the most advantageous has been the coating that consisted of five meshes up to the superheat value of ca. 8 K, and four and three meshes for the higher superheats.

Based on the above test results, it is possible to draw boiling curves, which more precisely show the heat transfer performance of each kind of surface. The results for wall superheat values in the range of 5 to 12 K have been given in Fig. 3a and 3b for distilled water and ethanol, respectively.

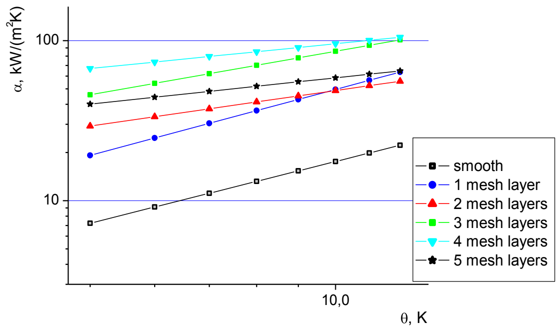


Figure 3a. Heat transfer coefficient for distilled water for different mesh layers and wall superheat 5 – 12 K.

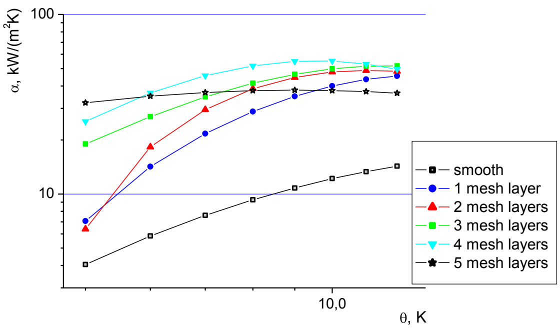


Figure 3b. Heat transfer coefficient for ethanol for different mesh layers and wall superheat 5 – 12 K.

The presented results indicate that the application of mesh layers enhances boiling heat transfer in comparison with the smooth surface. The heat transfer coefficient for the same wall superheat for the surface with the additional layers has been much higher than for the reference surface without the meshes. In can be noted that, generally, the heat transfer performance is improved as the number of mesh layers increases up to four – which seems to be the optimal number – especially for distilled water. Adding the fifth coating results in lowering the value of the heat transfer coefficient. It can be explained by higher hydraulic resistance of the flow of the created vapour out of the microstructure and of fresh liquid flow inside the microstructure for vaporization, so worse conditions of mass transfer within the coating.

In the case of ethanol at low superheats, the 5-layer coating provided the highest enhancement. It might be related to increased density of nucleation site for this working fluid already active at low superheats. At higher wall superheats vapor could be permanently present within the structure which might result in smaller heat fluxes dissipated from the fin. It needs to be noted that the two fluids considered in this paper have different surface wetting properties (surface tension for water is a few times higher), which might be a factor considerably influencing boiling heat transfer in microstructural coatings.

The impact of the number of mesh layers for two values of superheat (5 K and 12 K) has been presented in Fig. 4 in order to precisely determine the enhancement produced by the application of the mesh microstructures. The ratio of the heat transfer coefficient for the microstructure coated surface (α_m) and the smooth surface without any coating (α_s) is considered as a dependence of the number of mesh layers. The results have been presented for both distilled water and ethanol.

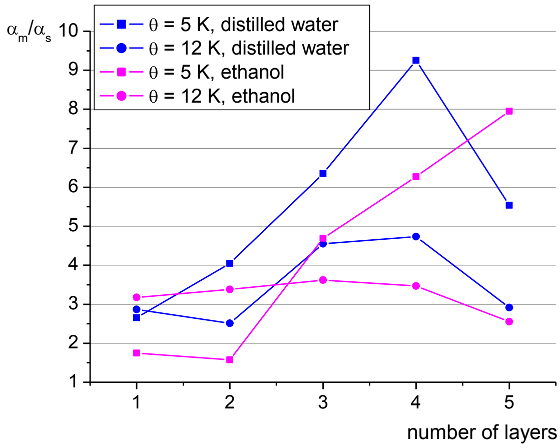


Figure 4. The ratio of heat transfer coefficient for the microstructure coated surface (α_m) and the smooth surface (α_s).

As can be seen in Fig. 4, the value of the heat transfer coefficient can be over nine times higher in comparison with the smooth surface if the additional mesh layers are added onto the heat exchanging surface for water boiling. In the case of ethanol about eight times higher values have been recorded. Generally, more significant enhancement has been observed for low superheats of 5 K than for 12 K. This phenomenon might be explained by a higher number of active nucleation sites for microstructures, which are activated at low superheats. In the case of the smooth surface the number of nucleation sites increases with temperature, while this growth is not so rapid for the structural coatings. That could be the reason why the enhancement for the higher superheat of 12 K might not be so significant as at low temperature differences.

The investigation of pool boiling on the non-isothermal surface in the presented experimental set-up has its limitations. Thus, in order to more precisely consider the impact of the number of mesh layers on the thermal performance of the samples, the boiling curves have been determined on the isothermal experimental facility, described in detail by the author in [10]. The samples have been located horizontally, which as opposed to the vertical location in the case of the non-isothermal testing set – up. The boiling performance of the samples has been presented in the form of heat flux dependence on wall superheat in Figs. 5a and b for distilled water and ethanol, respectively. The mesh layers had the porosity in the range of 64 – 70%.

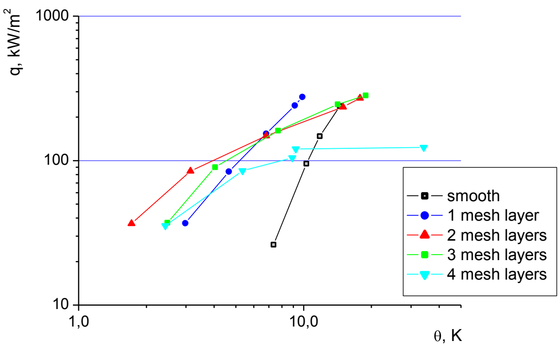


Figure 5a. Heat flux for distilled water for different mesh layers.

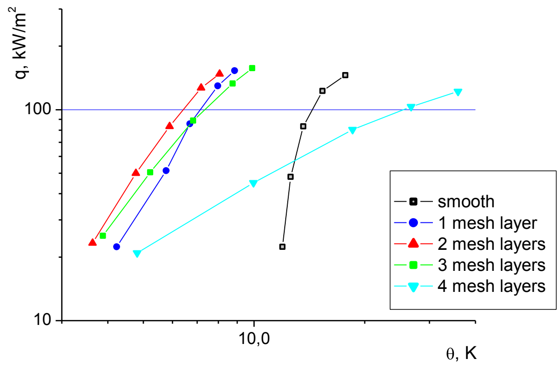


Figure 5b. Heat flux for ethanol for different mesh layers.

The analysis of the above figures confirms that there is an optimal number of meshes, which provides the best performance for the certain boiling liquid and the considered superheat range. The porous layers with the highest number of meshes (four in figures 5a and b) might produce the unfavorable effect of blocking the flow of liquid and vapor, thus, limiting the transferred heat fluxes. In the case of water boiling, the application of the 4-mesh layer led to a rapid transition to film boiling, while in the case of ethanol, the microstructure transferred heat in the nucleate boiling mode, but its performance was significantly reduced (even worse than the smooth surface without any coating). This might be related to the fact that water vapor bubbles are larger, and the highly packed structure produced the blocking effect, while in the case of ethanol the bubbles are smaller and the two-phase flow was still possible, although hampered.

4. Conclusions

The application of microstructural coatings enhances boiling heat transfer for the analyzed mesh layers located on the analyzed fin. The heat transfer coefficient can be several times higher if such additional coatings are used. The four-layer structure has been most efficient for water boiling, producing over nine times higher value of the heat transfer coefficient in comparison to the smooth surface. In the case of a different kind of mesh and the isothermal heating, four mesh layers proved to be least effective. This phenomenon proves that the thermal performance of phase change heat exchangers is significantly affected by the geometrical parameters of the coatings as well as the properties of the working fluid. Thus, the design of such heat exchangers for industrial applications should consider several parameters.

The application of an increasing number of mesh layers generally leads to elevated heat transfer coefficients. However, a considerable number of meshes could result in the opposite effect, namely the reduction of the dissipated heat flux. It seems to be related to difficulties in mass transport as vapor needs to flow out of the structure and liquid into it. Some vapor could permanently remain in the voids of the multi-layered coatings, especially at high superheats, thus, reducing the heat transfer coefficient in relation to surfaces with a lower number of meshes.

More tests in this field could help to produce a reliable model of boiling heat transfer on microstructure covered surfaces and provide information for the design of efficient phase change heat exchangers.

References

[1] J.Y. Tsay, Y.Y. Yan, T.F. Lin, “Enhancement of pool boiling heat transfer in a horizontal water layer through surface roughness and screen coverage”. Heat and Mass Transfer, Vol. 32, 1996, 17-26.

[2] A. Brausch, P.A. Kew, “Examination and visualisation of heat transfer processes during evaporation in capillary porous structures”. Applied Thermal Engineering, Vol. 22, 2002, 815-824.

[3] D.W. Gerlach, Y.K. Joshi, “Boiling performance of fluorinert PF 5060 on confined and unconfined wire meshes soldered to the substrate”. International Mechanical Engineering Congress and Exposition, Orlando, USA, 2005.

[4] A. Franco, E.M. Latrofa, V.V. Yagov, “Heat transfer enhancement in pool boiling of a refrigerant fluid with wire nets structures”. Experimental Thermal and Fluid Science, Vol. 30, 2006, 263-275.

[5] M. Rannenberg, H. Beer, “Heat transfer by evaporation in capillary porous wire mesh structures”, Letters in Heat and Mass Transfer, Vol. 7, 1980, 425-436.

[6] G.F. Smirnov, B.A. Afanasiev, “Investigation of vaporisation in screen wick – capillary structures”. International Heat Pipe Conference: Advances in Heat Pipe Technology, London, UK, 1982.

[7] S.-C. Wong, Y.-H. Kao, “Visualization and performance measurement of operating mesh – wicked heat pipes”. International Journal of Heat and Mass Transfer, Vol. 51, 2008, 4249-4259.

[8] J.-H. Liou, C.-W. Chang, C. Chao, S.-C. Wong, “Visualization and thermal resistance measurement for the sintered mesh – wick evaporator in operating flat – plate heat pipes”. International Journal of Heat and Mass Transfer, Vol. 53, 2010, 1498-1506.

[9] R. Pastuszko, „Pool boiling heat transfer on micro-fins with wire mesh Experiments and heat flux prediction”, International Journal of Thermal Science, Vol. 125, 2018, 197-209.

[10] Ł.J. Orman, “Enhancement of pool boiling heat transfer with pin-fin microstructures”. Journal of Enhanced Heat Transfer, Vol. 23, No 2, 2016, 137-153.

[11] R. Chatys, Ł.J. Orman, “Technology and properties of layered composites as coatings for heat transfer enhancement”. Mechanics of Composite Materials, Vol. 53, No 3, 2017, 351-360.

[12] T. Orzechowski, Ł.J. Orman, “Measurement errors in thermovision diagnostics of machinery”, Science Report, CEEPUS Project PL-0007: Modern Metrology in Quality Management Systems. Eds: S. Adamczak, P.H. Osanna, Kielce University of Technology, 2006, 209-216.

[13] T. Orzechowski, Wymiana ciepła przy wrzeniu na żebrach z mikropowierzchnią strukturalną, Kielce: Publishing office of Kielce University of Technology; 2003.

[14] Ł.J. Orman, T. Orzechowski, “Analysis of boiling heat transfer on the non-isothermal surface covered with selected wire mesh structures”. 7th World Conference on Experimental Heat Transfer ExHFT-7, Krakow, Poland, 2009.

[15] N.H. Afgan, L.A. Jovic, S.A. Kovalev, V.A. Lenykov, “Boiling heat transfer from surfaces with porous layers”, International Journal Heat Mass Transfer, Vol. 28 (2), 1985, 415-422.

[16] H.C. Ünal, “Determination of the temperature distribution in an extended surface with non – uniform heat transfer coefficient”, International Journal Heat Mass Transfer, Vol. 28, 1985, 2279-2283.

[17] T. Orzechowski, “Local values of heat transfer coefficient determination on fin’s surface”, Experimental Thermal and Fluid Science, Vol. l31 (8), 2007, 947-955.

[18] Ł.J. Orman, “Measurements of boiling heat transfer on a single fin”, Structure and Environment, Vol. 3(1), 2011, 47-51.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 161 - 174

Jordan Journal of Mechanical and Industrial Engineering

Assessing the Thermal Effectiveness of Implementing Green Roofs in the Urban Neighborhood

Dana Al Jadaa[a], Abeer Abu Raed[a*], Hanan Taleb[b]

[a] *Department of Architecture, American University of Ras Al Khaimah, Ras Al Khaimah, UAE*
[b] *Faculty of Engineering, The British University in Dubai, Dubai International Academic City, UAE*

Received 8 September. 2019

Abstract

The United Arab Emirates demonstrated abundance attention to modern developments, especially when it comes to the ostentatious projects of architecture, thermal comfort, energy efficiency and smart cities. Producing one of the world's most astonishing cities is not the only concern for the government of the United Arab Emirates. The United Arab Emirates paid great attention to implementing various policies and strategies targeting to become among the most sustainable countries in the world. One of these strategies is the implementation of green roofs in residential and commercial spaces.

Researches and studies indicate that green roofs are one of the strategies that result in an ecosystem when it is applied in urban regions. It can reduce the heat island effect that is formed in urban spaces which results in reducing air temperature and pollution in the environment, reducing greenhouse gas emission, energy conservation and many other benefits. This research proved the ability to mitigate heat island and improve thermal conditions when integrating these structures into the new developments and neighborhoods. This is achieved by examining the sequences resulting from integrating green roofs to the rooftops of residential units at Mina Al Arab compound by measuring the thermal parameters and investigating the changes that will occur using ENVI-met simulation program. The simulation compares a green roof system and a conventional roof system at 1:00 pm on a summer day and a winter day.

The results gained from the simulation of Mina Al Arab demonstrated that the green roof system has excellent potential to mitigate heat island effect and improve thermal conditions by decreasing relative humidity, mean radiant temperature, and coefficient of heat exchange.

Keywords: Green roof, Sustainability, Simulation, UAE, Environment;

1. Introduction

The quick expansion in economic and social growth contributed to an obvious increase in the ecological footprint that leads to the heat island phenomenon. Heat island effect is a result of the increase of concrete structures and reduction in open spaces and green surfaces that filters the air and play a magnificent rule in reducing air temperatures around. As a result, a clear increase in pollution and air temperature in urban areas compared to adjacent rural spaces is witnessed.

This increase in air temperatures raises the demand for cooling which requires an abundant amount of energy and power. Studies showed that 70% of the electricity demand is required for cooling in a building in hot dry regions like UAE. Heat island effect increases three times the demand for electrical power, it increases two times the cooling load and decreases the performance of air conditioning machines in a building [1]. An urban heat island profile is presented graphically in figure 1.

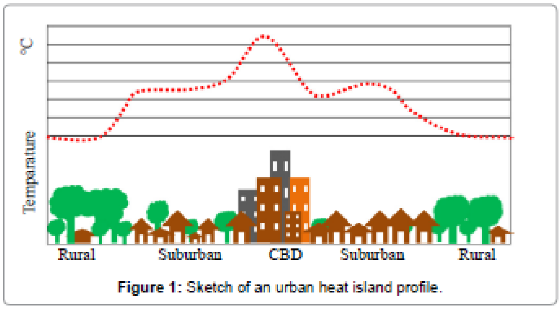


Figure 1. A sketch of an urban heat island profile [2]

Growing plants on the top of buildings roofs is an old idea used by people since the past. People used to grow their herbs and plants for eating and drinking and even for leisure on top of their roofs for centuries. They also planted their roofs in order to protect them from high radiation that caused high temperatures inside their houses [3]. At the beginning of the nineteenth century, before the

* Corresponding author e-mail: abeer.raed@aurak.ac.ae

use of green roofs, people used to build water features on the top of roofs in order to cool rooftops and allow the forming of a cool breeze. A few decades later many European countries including Germany and Sweden begin to introduce green roofs into their houses in order to increase vegetation areas that are considered very valuable. Green roofs were used not only for planting, but also as a solution for controlling storm water by absorbing part of it and preventing heat from penetrating to the indoor spaces.

The concept of green roof is reused nowadays to enhance buildings energy performance. In a model research done by Liu and Minor [4], a comparison between a green roof and a normal roof was conducted. Results showed that when a green roof is installed at the top of the building, the energy demand for cooling and heating is up to 3 times less. Green roofs served as a barrier protecting the building's top from solar radiation.

Parizotto and Lamberts [5] examined the thermal effect of a green roof when integrated into a family household in Brazil. The researchers did experiments using different types of roof finishing. These materials are: metal roof, ceramic tiles, and a green roof. By measuring the indoor air temperature, they found out that the lowest indoor temperature was recorded when they used a green roof system, followed by ceramic tiles and metallic system. This is due to the low heat gain property of soil compared to ceramic tiles or metal. Also, the layers of soil and the drainage system work as a buffer zone or as an extra insulation layer that prevents heat transfer to the indoor spaces. Moreover, they found out that the shaded area formed by the higher plants was even less in temperatures. As a result, the demand for cooling is less and the cooling performance is enhanced in the study area.

Alcazar S., Bass B. [6] also investigated the performance of green roofs and its effect on thermal comfort in indoor spaces according to many parameters including soil layer thickness, irrigation and planting density in addition to types of plants used and the area of plants leave. According to his investigation green roofs are one of the greatest methods of insulation or finishing to be used for building roofs in order to reduce heat transfer to buildings. He also recommended using vegetation that has wide leaves area and that grows and expands horizontally in order to maximize the reflection of solar radiation and increase the shaded area.

In a research paper that has been done by Rosheidat and Bryan [7], a simulation study using ENVI-met on a selected area in the city of Phoenix in Arizona was conducted. The aim of their simulation was to calculate the reduction in temperatures in the space. They found out that the temperature is decreased by 5°C in vegetated areas. Moreover, they found out that changing in the surface material and using materials that have high reflectivity also reduced the surface temperature reasonably. Moreover, the difference between materials that have high reflecting abilities and green roofs is compared and found out that green roofs have a higher albedo level and was able to reflect solar radiation more than reflective materials, although these results varied between different types of vegetation types and layer thickness. They also included in their research simulation of water content within layers and indicated that water content leads to a reduction in surface temperatures during the daytime period and that the quantity of temperature reduction is proportional to the quantity of moister content in the green roof layer.

A quantitative case study accomplished by Kumar and Kaushik [8] in India indicated that the thermal performance of a green roof system is influenced by leaf area index and vegetation height by acting as a shading layer that prevents solar radiation. Consequently, heat gain to indoor spaces and indoor air temperatures were reduced.

One of the main challenges that limit the expansion of the green roof system is the lack of policies and regulations that encourages the use of these systems in current and new developments. This can be related to the limited number of researches and studies in this area. Another factor to consider is the lack of experience in the design, implementation and maintenance of the green roof systems. Also, manufacturers of green roof products are located and concentrated only in some countries, thereby increasing the cost of exporting the material to the required place. For instance, the application of green roof system in hot arid regions improves the thermal comfort and reduces the load on the cooling system, while green roof system on the other hand in humid and rainy areas protects construction composition by regulating and handling storm water. It could be a difficult job to alter the present roofing technologies used by developers and contractors, so encouraging new practices can begin through the government for government structures to show their intention to find more viable alternatives and to be a reference for future changes in the building industry [9].

The installation costs are deemed to be one of the factors why green roof systems are unpopular in some areas. However, the availability of a broad range of new products and techniques makes it more feasible and more demanding at reduced market prices. As researchers have proven the advantages of using green roofs, the demand for them has increased by owners of spaces who are prepared to pursue additional sustainable alternatives and by manufacturers and vendors who supply these techniques. As researches and evidence proved the benefits of using green roofs, the demand for it increased by owners of spaces who are willing to go for extra sustainable solutions and by producers and suppliers that are providing these technologies. Consequently, the price of these techniques decreased and became more viable for customers [10].

Abu Dhabi's government has established new policies and guidelines aimed at a sustainable green future and at controlling and reducing city emissions. One of the most important strategies that have been set by the government of Abu Dhabi is Estidama guidelines that aim to a new sustainable future and encourages to adopt new methods, strategies and technologies to face the challenges of the future. Although landscaping and vegetation in Abu Dhabi are challenging due to its hot climate and dusty soil characteristics, in addition to the lack of water resources, however Estidama new guidelines encourage the integration of green roofs and vegetation spaces in order to control and reduce heat island effect in the city.

The theoretical base of this paper is to investigate the potential benefits of green roofs in the urban areas of a hot arid region like the United Arab Emirates that is facing a huge increase in the population size and footprint yearly. This research aims to examine the ability to minimize and control the environmental thermal issues associated with urban developments and to evaluate the positivity of integrating green roofs in the buildings. This will be done by simulating outdoor thermal limitations including surface temperature, relative humidity (RH), mean radiant temperature (MRT) and wind speed. The research findings will certify the contribution and advantages of incorporating green roof systems into urban buildings. One of the objectives of this work is to guide construction architects and designers to enhance their designs to obtain maximum thermal comfort and energy savings in any project according to their requirements.

2. Research study area and methodology

Site examination and computer simulation are the most common methodologies that have been used to study green roofs. The positive point of computer simulation is time and cost-saving by the ability to test different locations, during different seasons of the year using different weather files. Moreover, simulation programs allow for studying and comparing different parameters with the least difficulties and it gives the chance for applying changes on the module and all different parameters. It also saves a lot of time and effort required to measure and observe using sensors and testing machines needed for experimental methodology. Even the module scale, number of floor height, location, and finishing material can be changed and updated. Although simulation methodology is considered very flexible and has many positive aspects, it also has some negative aspects like human errors during input of data or even in the programming of the software that is designed by humans.

In this work, simulation methodology will be used to evaluate the performance of green roof over an urban compound community. The creation of a simulation prototype that integrates a green roof into a building in a social community is a dynamic way to detect its performance and interaction. It gives the chance to observe the impact of weather conditions, the material used in the system, wind and temperature interaction on the prototype. All data obtained from simulation programs will be saved and presented graphically in forms of tables and charts. The software for simulation to be used in this research is ENVI-met. The simulation will provide information to be analyzed so that an ideal solution can be found. The simulation software does not have all local vegetation data of UAE. Therefore, some manual records will be required

The study is conducted in Ras Al-Khaimah which is in the northern area of UAE. Retaining a rich archaeological culture that dates back to 50centuries, Ras Al-Khaimah features a treasure of landscapes and environments, ranging from huge mountains to seaside areas and red sandy desert. Due to its topography Ras al-Khaimah is considered one of the most desirable growing tourist destinations of all the emirates and gulf area, it contains a great range of amenities and resorts that attracts visitors from the surrounding emirates and from all over the world, where many developments, resorts, museums, malls, golf courses and other facilities are available.

The area that will be investigated is a residential area located at Mina Al Arab which is a waterfront luxury development established by RAK properties. Mina Al Arab consists of six districts that have unique characteristics and attractive atmosphere. The urban structure is designed carefully to accommodate different facilities and amenities that please residents and visitors. In addition to being a residential compound, Mina Al Arab contains many luxury resorts and hotels for visitors. The master plan Mina Al Arab is presented in figure 2.

The selected area that will be studied is Bermuda. Bermuda consists of residential villas and townhouses that vary between 2-6 bedrooms with private and shared gardens, open spaces and beach sides as listed in Table 1 and presented in Figure 3.

Figure 2. Mina Al Arab master plan. Source: Rak Properties, 2017.

In this study the simulation of green roof integration will be applied to the 5 different types of units to investigate the effect of the green roof and how does it interact with adjacent buildings and units. All the units are consisted of 2-floor heights and are adjacent to private open gardens, spaces and shared amenities that have different surfaces and Albedo levels. Figures 4, 5 and 6 show perspective, ground and first-floor plan of type D type villa. Each villa has a private roof and a private garden. The building envelope and roof of all types (A, B, C, D and E) are built of concrete.

Table 1. Bermuda villa types and area. Source Rak Properties, 2017

Villa Type & Sizes Sr. No	Model	No. of units	Bedrooms	Build up Area	Approximate Plot Area
1	Type E	48	2	2572.3	2184
2	Type D	61	3	3773.5	3432
3	Type C	23	4	4768.2	6240
4	Type B	20	5	5530.6	6456
5	Type A	5	6	6741.1	7747

Figure 3. Bermuda key plan. Source: Rak Properties, 2017.

Figure 4. Type D villa back view. Source: Rak Properties, 2017.

The Envi-met academic version is used to simulate the required area. The simulation process involves different stages and procedures. The stages include setting up the inputs, modeling the area and assigning materials, setting up a simulation file that contains all settings and visualizing the outputs using Leonardo visualization that is included in the program as shown in figure 7.

Figure 5. Type D ground floor plan. Source: Rak Properties, 2017.

Figure 6. Type D first-floor plan. Source: Rak Properties, 2017.

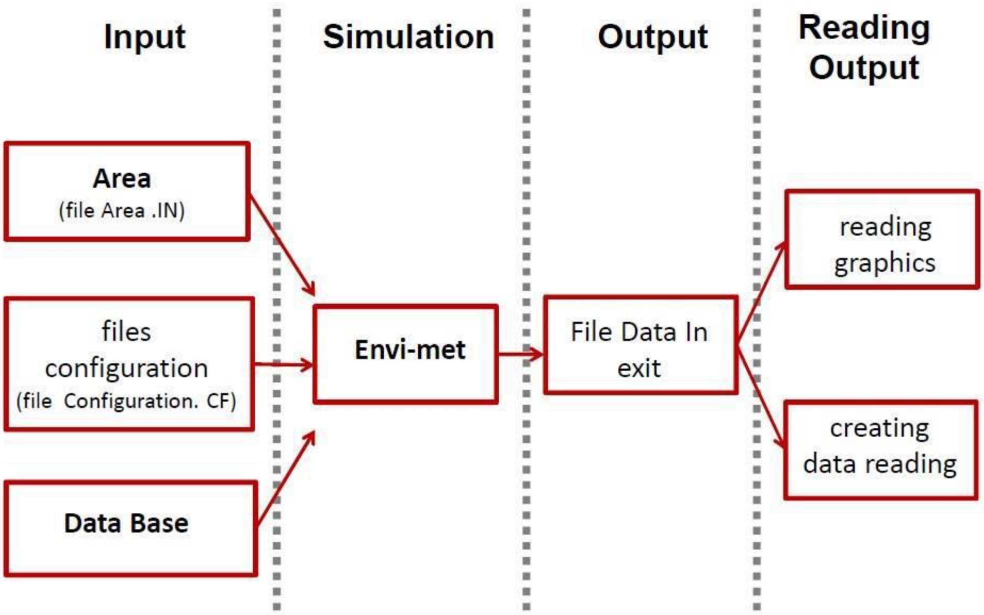


Figure 7. Process of ENVI-met simulation. Source: ENVI-met learning, 2017.

  - Volume 13, Number 3 (ISSN 1995-6665)

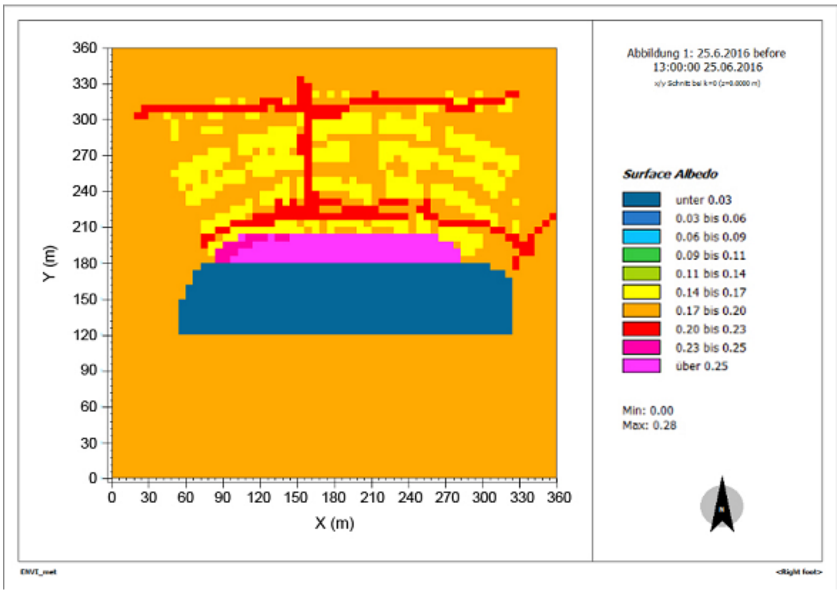
3. RESULTS AND DISCUSSION

The simulation was conducted during summer (25th of June 2016 at 13:00 pm) where sun exposure, radiation and temperatures are at its maximum. The study during winter was conducted on the 25th of January at 13:00 pm. The studies were set in order to investigate and evaluate different parameters. These parameters include surface albedo, heat exchange coefficient, relative humidity, mean radiant temperature and wind speed. Analyzing the results obtained from the tested parameters is performed.

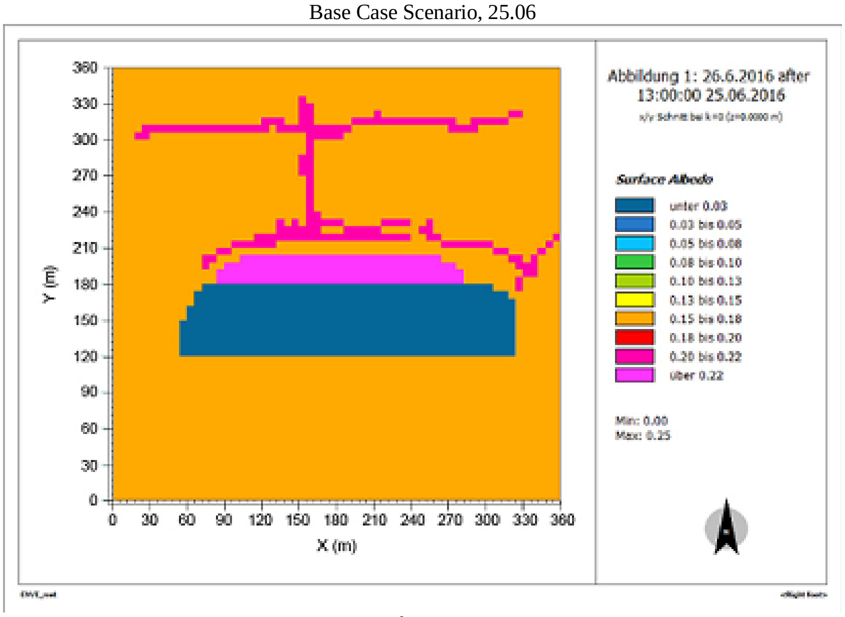
4. Surface albedo

A simulation for the surface albedo was conducted using ENVI-met to evaluate the changes that occur on the albedo level of different surfaces in winter and summer for conventional roof and green roof systems. To conduct the simulation, different inputs were used for the four cases and simulated in order to produce the output files that contain the results of the simulation. The last step is to present the results graphically using Leonardo program

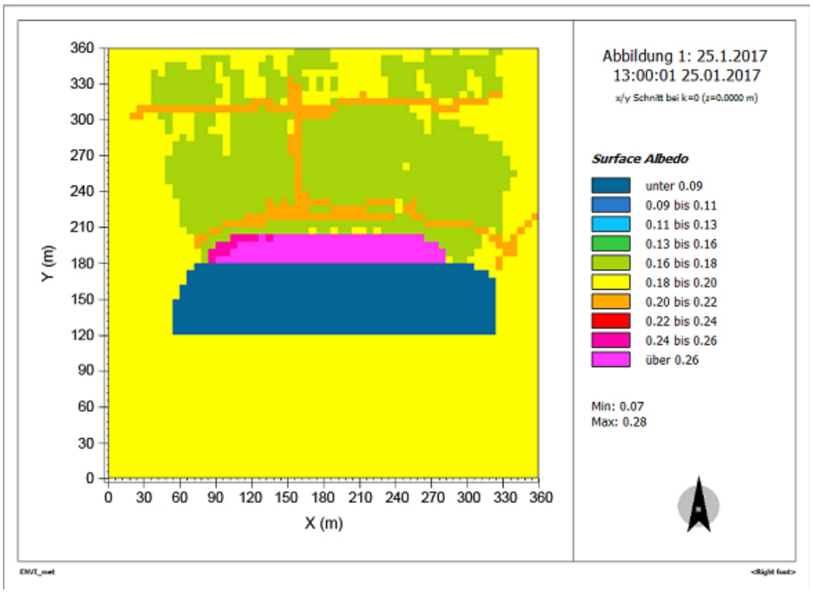
Figure 8 presents the graphical analysis of surface albedo for the four different cases. It is noticed that surface albedo maximum and minimum ranges did not change between summer and winter time, however it did slightly change between the conventional roof and green roof scenarios. According to the simulation done for Mina Al Arab community, the results do not show any potential for green roofs to reduce surface temperatures that are emitted to the surrounding space. Further studies on a larger area with additional information including the type of the installed materials to the green roof system, the area of leaves and the water content may increase the changes of the simulation results.



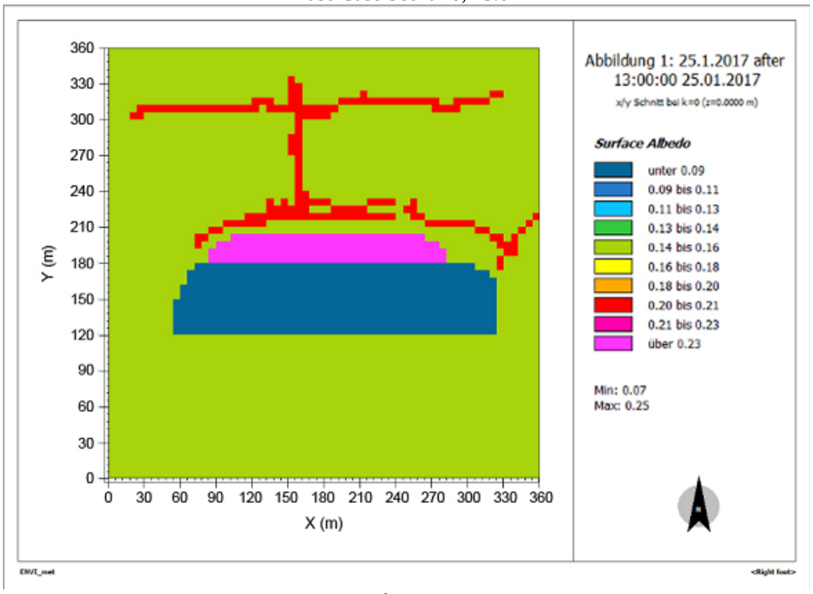
Base Case Scenario, 25.06



Green Roof Scenario, 25.06



Base Case Scenario, 25.01



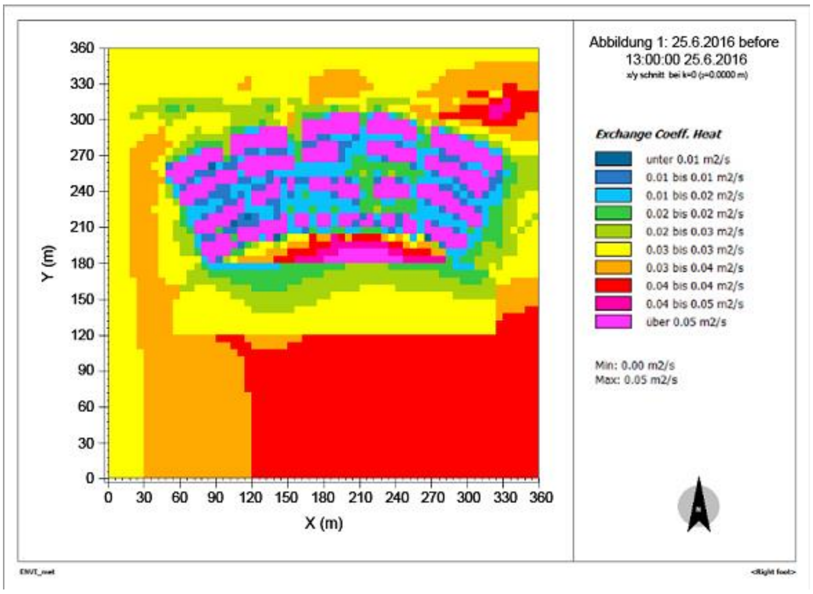
Green Roof Scenario, 25.01

Figure 8. The graphical analysis of surface albedo

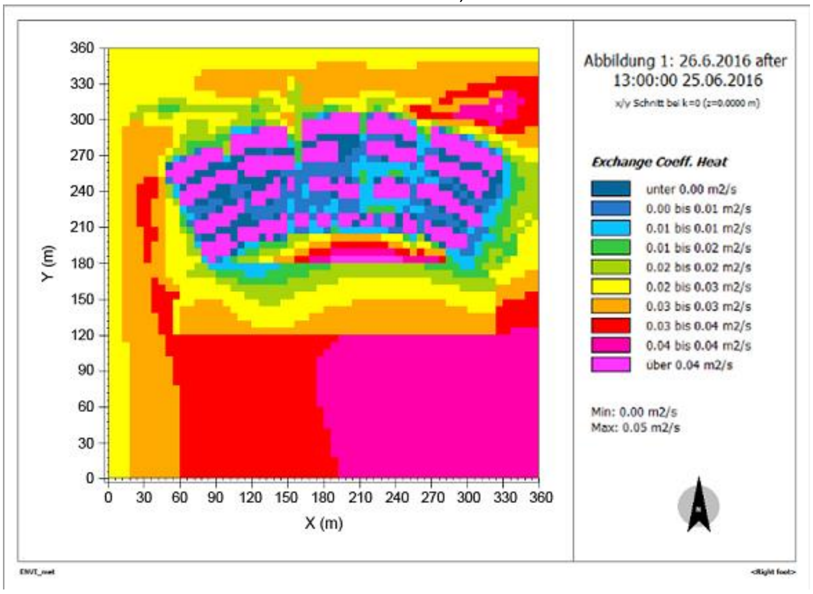
5. Heat exchange coefficient

A simulation for the heat exchange coefficient was done to simulate and evaluate the variation of heat exchange levels in the four different cases of roofing systems in the different seasons of the year using the ENVI-met simulation software. As in the previous parameter a simulation was conducted for the different cases producing an output file. The output file is presented using Leonardo rendering software to obtain results and conclusions.

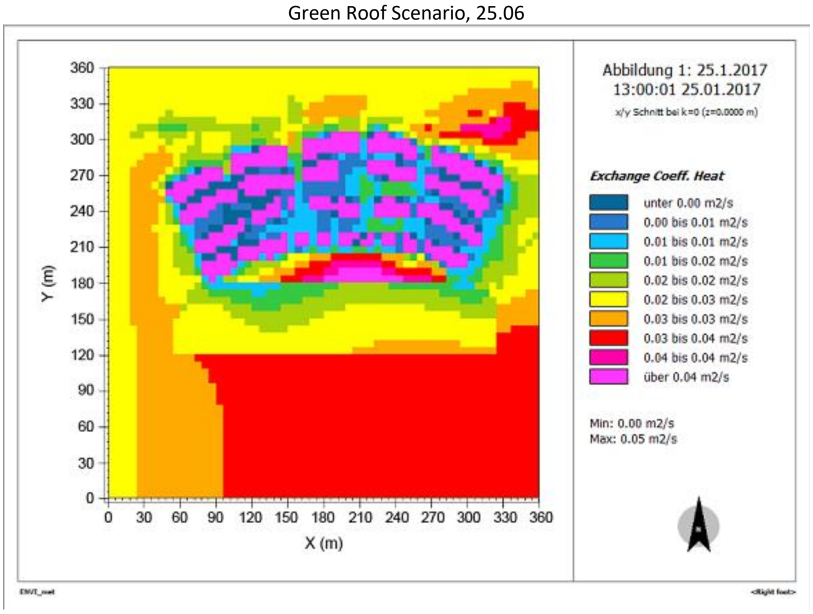
According to the graphical rendering of the heat exchange coefficient simulation results for the different cases shown in Figure 9, the heat exchange coefficient for both the conventional roof and the green roof did not change from summer to winter on the top of the roofs. However, it did change around the building units which means that it affects the surrounding environment and reduces the heat exchange between surfaces and objects. As a result, the potential of using green roofs to reduce the heat exchange coefficient must be considered. It affects the thermal environment of the surrounding area by increasing the comfort zone and it affects the interior spaces or indoor environment.



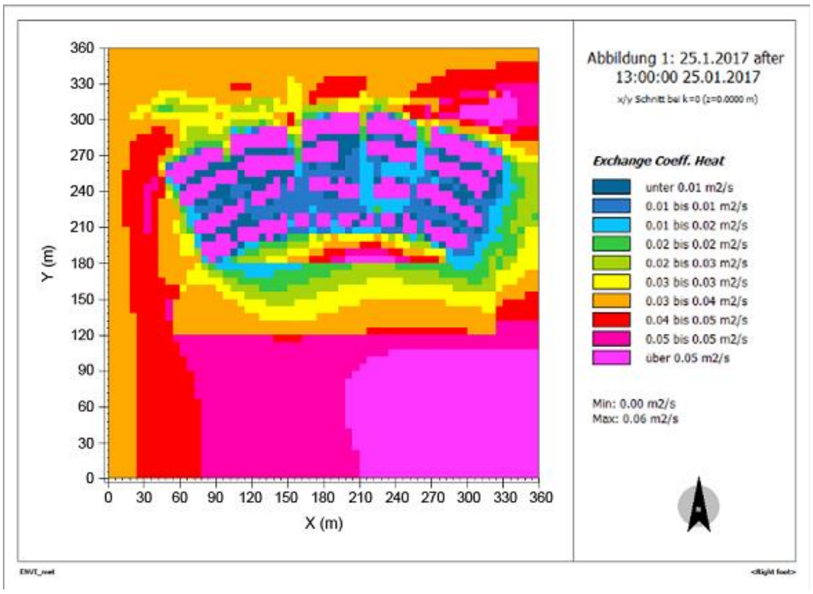
Base Case Scenario, 25.06



Green Roof Scenario, 25.06



Base Case Scenario, 25.01



Green Roof Scenario, 25.01

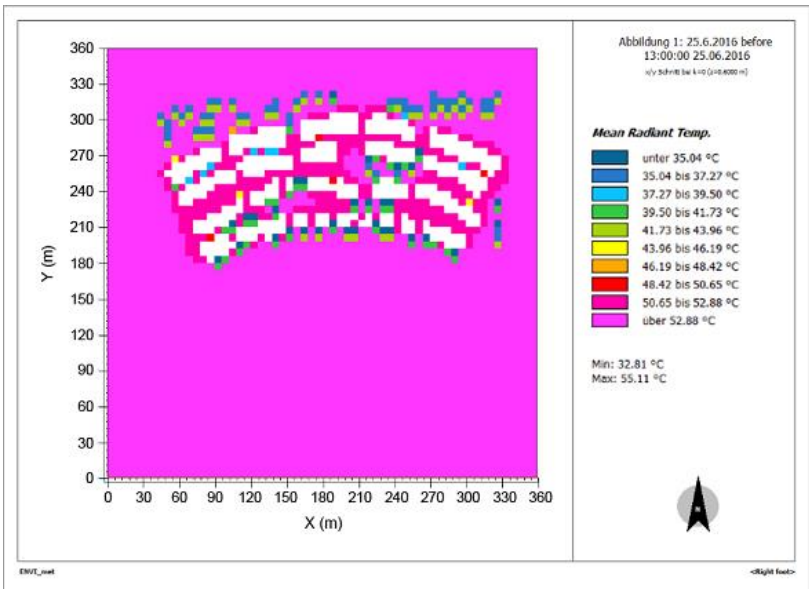
Figure 9. The heat exchange coefficient simulation results

It is essential to consider also that vegetation type and leaf area index can also contribute to the changes in the heat exchange coefficient levels. Vegetation types that are higher in height and have very dense foliage or wide leaves area deliver enough shading on the rooftop that limits or reduces the amount of solar radiation that can reach the rooftop.

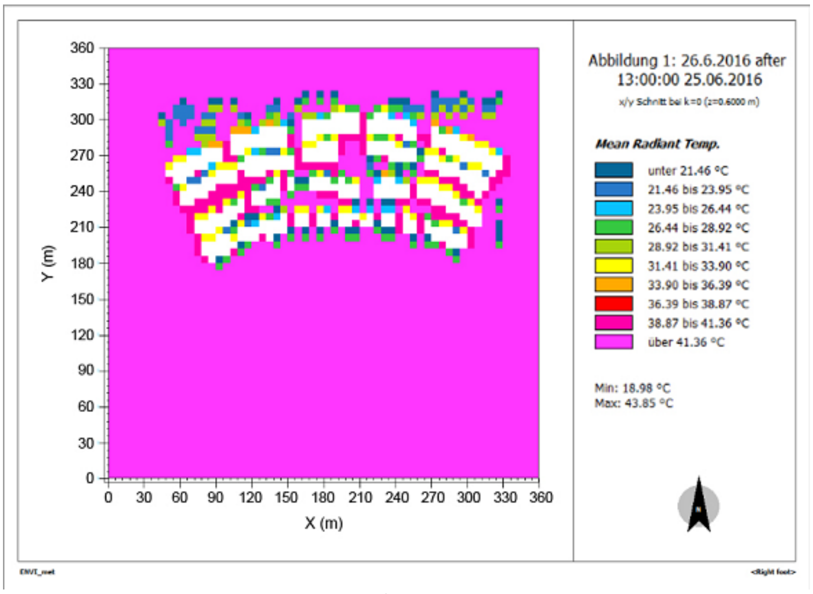
6. Mean radiant temperature

Analyzing the mean radiant temperature of the case studies by simulation using ENVI-met software showed the variation in levels between summer and winter, green roof and conventional roof case. To achieve results the simulation was applied to the different cases producing an output file previewed by Leonardo software as shown in figure 10 that illustrates the findings.

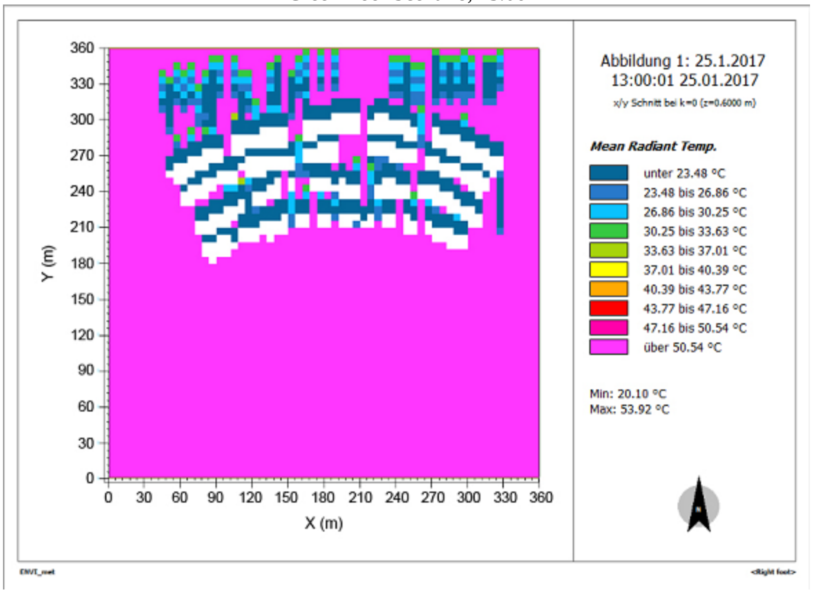
The simulation of the 4 cases recorded noticeable changes in the mean radiant temperature levels. The radiant temperature during the summer season ranged from 32 ° C to 55 ° C, but with the integration of the green roof system the radiant temperature ranged from 18 ° C to 43 ° C on the same day, time and weather. For the conventional roof in the winter season mean radiant temperature was ranged between 20°C – 53°C however with the integration of green roof system the percentage of mean radiant temperature dropped to be in the range of 36°C – 52°C. According to the results, the potential of using a green roof system in urban spaces have a great effect on reducing the mean radiant temperature in an urban space. This greatly affects the thermal comfort in the area by reducing the solar waves that are emitting and radiating between surfaces, spaces and the surrounding air. This drop-in means radiant temperature levels, since vegetation surfaces installed on the top of the buildings, emits shorter radiation waves when compared to the compacted hard surfaces and objects like concrete roof used in buildings.



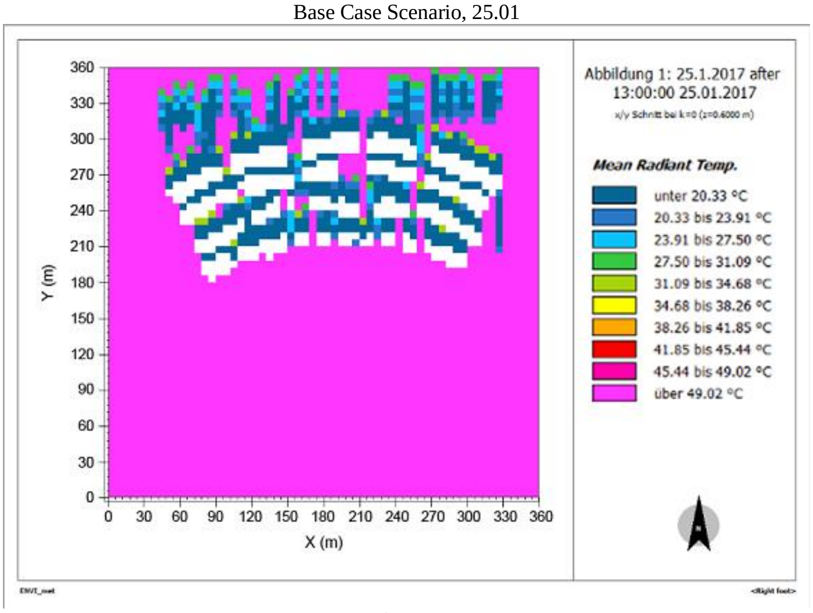
Base Case Scenario, 25.06



Green Roof Scenario, 25.06



Base Case Scenario, 25.01

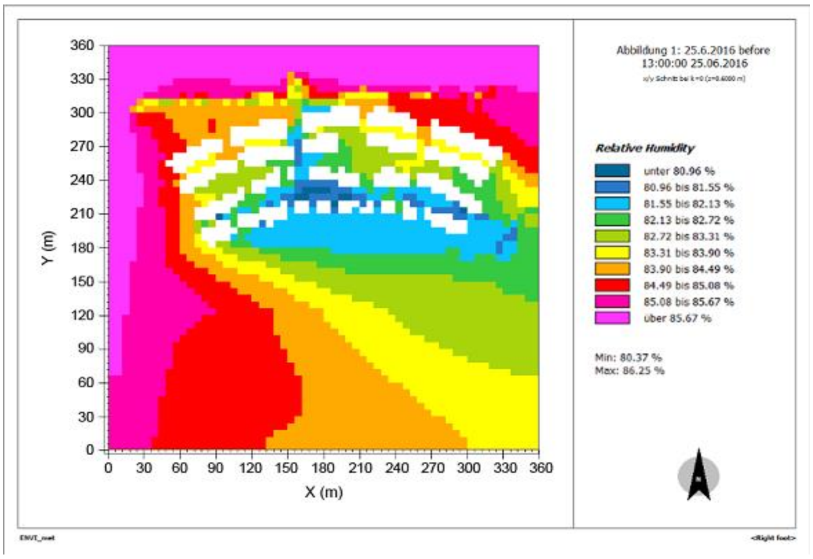


Green Roof Scenario, 25.01

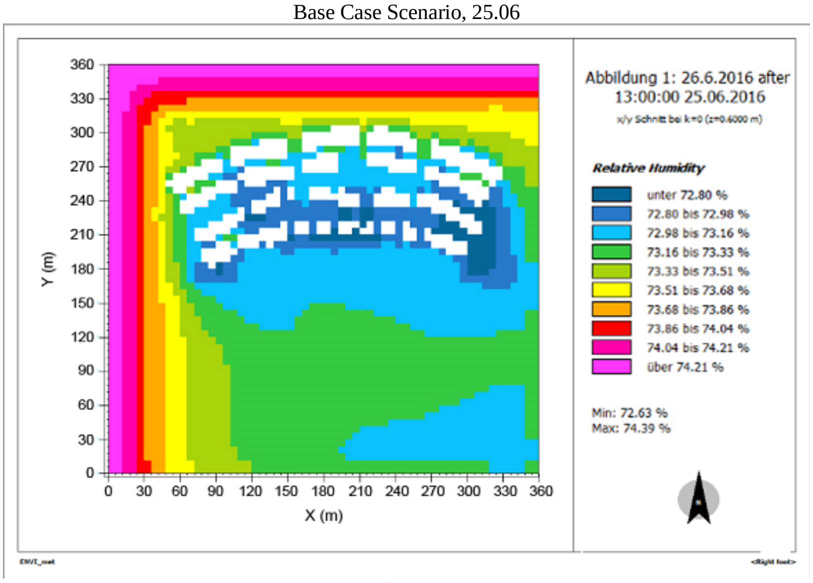
Figure 10. Mean radiant temperature simulation results

7. Relative humidity

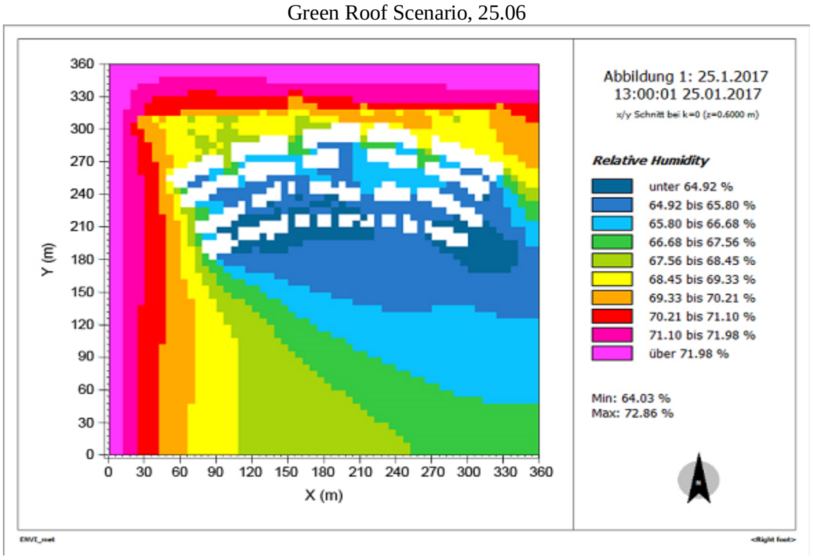
According to figure 11 that illustrates the relative humidity levels for the 4 different cases. Relative humidity varied between summer and winter, with conventional roof and with the green roof system.



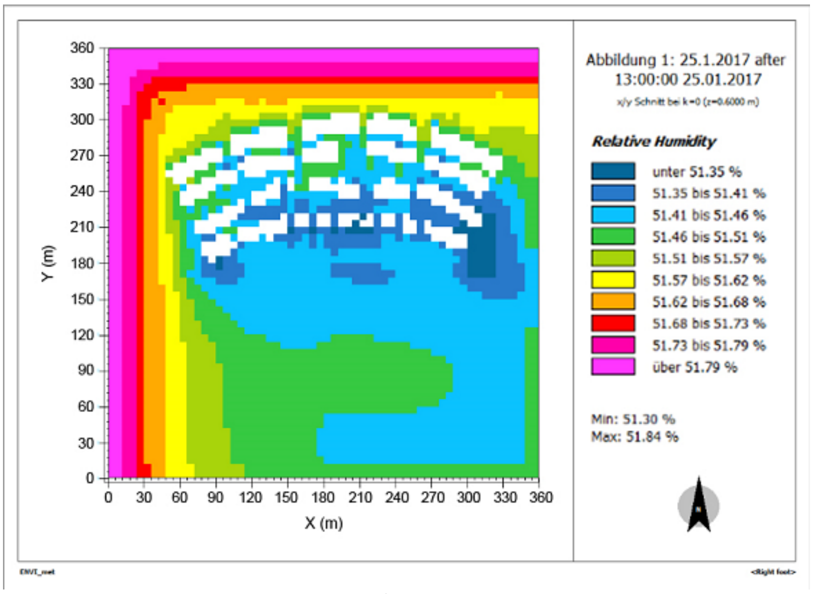
Base Case Scenario, 25.06



Green Roof Scenario, 25.06



Base Case Scenario, 25.01



Green Roof Scenario, 25.01

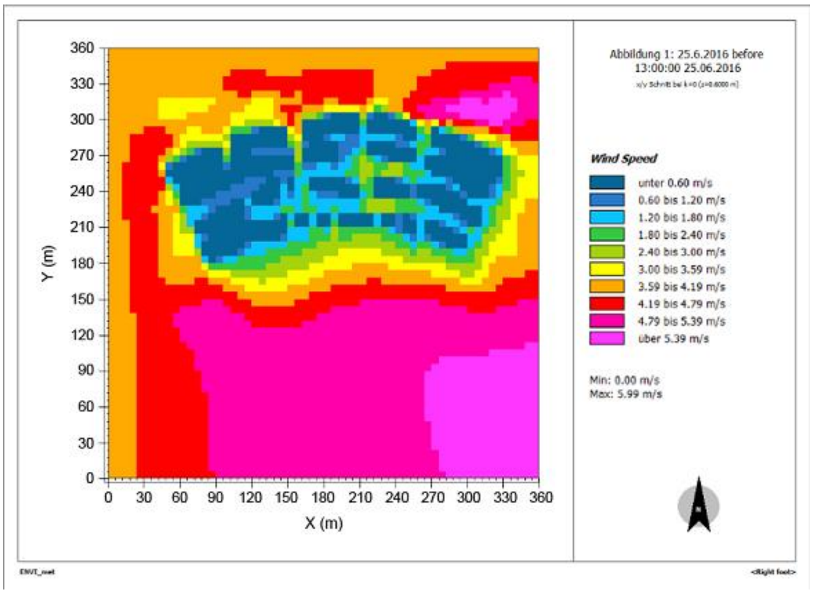
Figure 11. the relative humidity levels simulation results

In the summer season with the conventional roof scenario, the relative humidity ranged between 80 – 86% which records an average of 83%. During the same summer day, time and location but with the integration of green roof system instead of the conventional roof system, relative humidity ranged between 72 – 74% which records an average of 73%. As a result, the relative humidity in summer dropped 10% when the green roof system was integrated. This change in the records increases the potential for using a green roof system to control thermal conditions in an urban area especially during the summer period of the year.

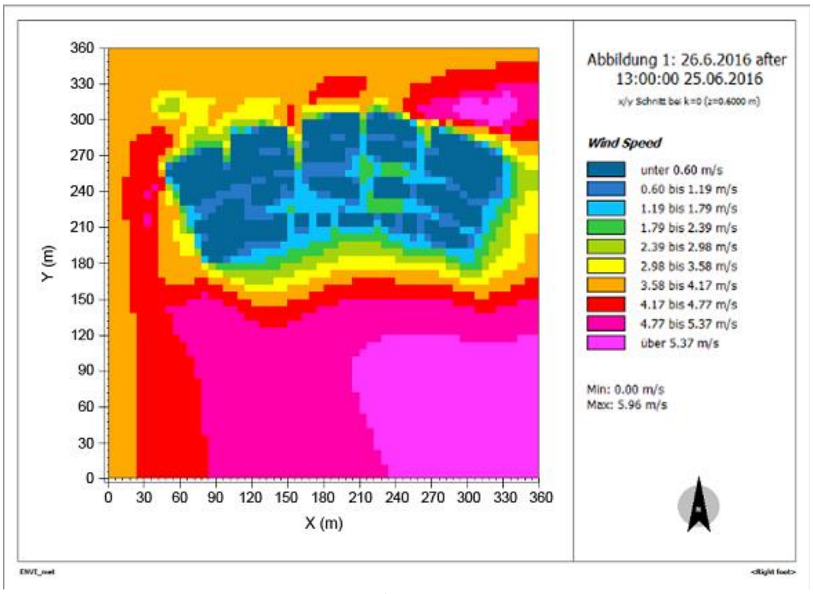
During the winter season, with the conventional roof scenario, relative humidity ranged between 64 – 72% which records an average of 68%. During the same summer day, time and location, with the integration of the green roof system instead of the conventional roof system, relative humidity recorded an average of 51%. According to the results, the relative humidity in winter dropped around 17% when the green roof system is integrated with the urban area selected at 1:00 pm.

8. Wind speed

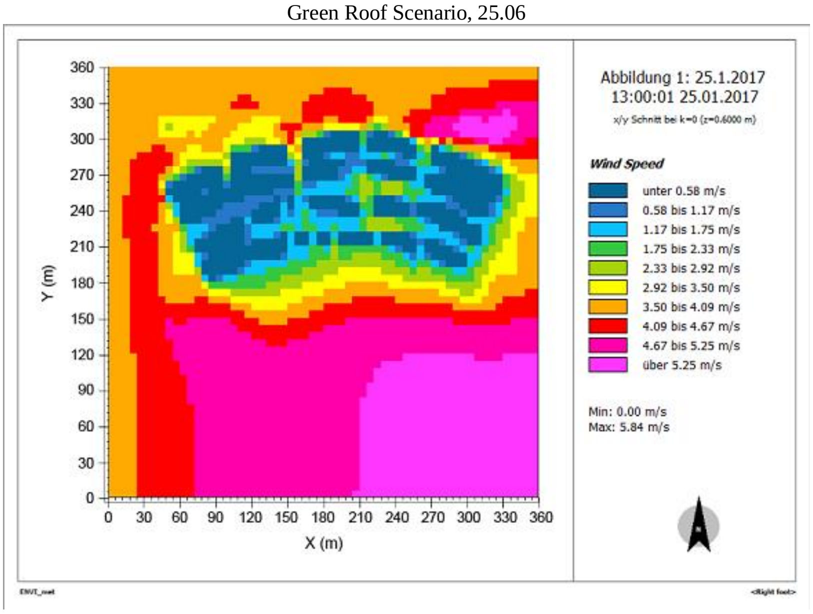
As shown in the previous simulation of the variable parameters, a simulation for the wind speed parameter using ENVI-met simulation was conducted to show the changes in the levels between summer and winter. The comparison between the different cases presented in figure 12 indicates that wind speed levels didn’t change between summer and winter, however it changed slightly when comparing the records between conventional roofing system and green roof system.



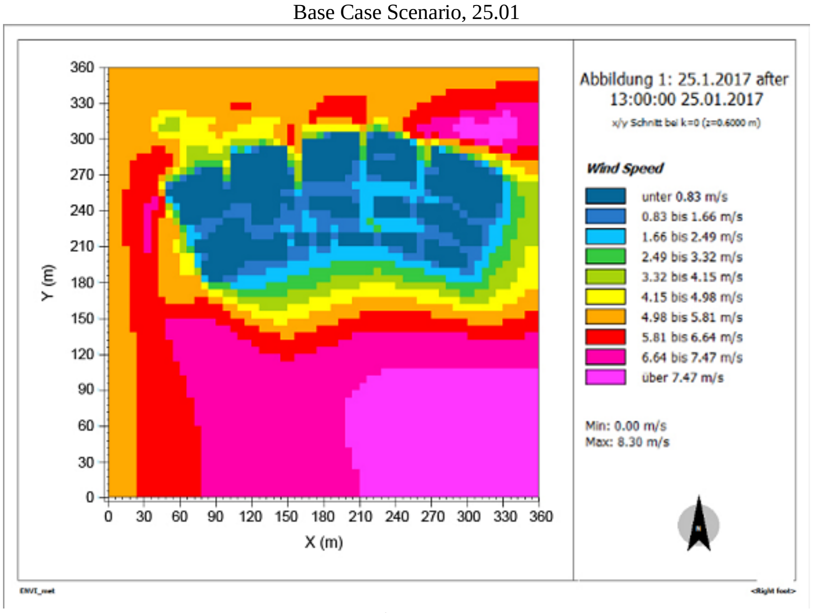
Base Case Scenario, 25.06



Green Roof Scenario, 25.06



Base Case Scenario, 25.01



Green Roof Scenario, 25.01

Figure 12. Wind speed simulation results

Wind flow and speed affect the humidity levels which also effects the mean radiant temperature, all these parameters play a marvelous role in controlling thermal comfort in the urban space.

Variation in the floor heights in the urban space also affects the wind speed, flow and pattern. However, this point cannot be considered or determined in this case study where all the units are composed of 2 levels with an average of 6 meters height.

The results of the conceptual model that was conducted for Mina Al Arab urban compound presented evidence and proofs that integration of green roof system highly effects on different parameters that contribute to an enhanced and improved outdoor thermal condition for users.

9. CONCLUSION

This research investigated the integration of vegetation in regions that have hot arid weather like the UAE through ENVI-met simulation software. The simulation using ENVI-met simulation program detected the influence of using green roofs in hot arid climates of Mina Al Arab compound and the potential of integrating this system to the selected area. All the parameters investigated in the research and the outcomes of this research are essential for developers, architects and engineers to promote new innovative methodologies to reduce heat island phenomenon and improve the thermal conditions using green roof systems.

The outcomes that were obtained from the simulation of Bermuda compound area in Mina Al Arab proofed that green roof system has a great potential in mitigating heat island effect and improving thermal conditions by increasing albedo levels of the urban structure surfaces, reducing relative humidity, mean radiant temperature and heat exchange coefficient.

Green roof systems work as an insulation material that can insulate the main roof layer from direct solar radiation as a result reducing longwave radiations and moderate mean radiant temperature through the reduction of exchange heat coefficient. Green roofs when it functions as an insulation system for the rooftop prevent radiation from penetrating the indoor spaces as a result it reduces the demand for cooling and regulates the amount of energy power required for that.

This research showed that when different thermal parameters are evaluated, designers can enhance and improve their designs in order to obtain the maximum levels of thermal comfort and energy saving in any project depending on its requirements and functions.

REFERENCES

[1] Jim, C. (2015). Assessing climate-adaptation effect of extensive tropical green roofs in cities. Landscape And Urban Planning, 138, 54-70.

[2] Bhargava A, Lakmini S and Bhargava S, Urban Heat Island Effect: It's Relevance in Urban Planning, Journal of Biodiversity and Endangered Species, Volume 5, Issue 1, 2017.

[3] Carter, T.L. and Rasmussen, T.C. (2005) Use of green roofs for ultra-urban stream restoration in the Georgia Piedmont (USA). In Proceedings of the 3rd International Greening Rooftops for Sustainable Communities Conference; 2005 May 4-6; Washington, DC.

[4] Liu. K, Minor J. 2005. Performance evaluation of an extensive green roof. Paper presented at the Third Annual Greening Rooftops for Sustainable Communities Conference, Awards and Trade Show; 4–6 May 2005, Washington, DC.

[5] Parizotto, S. & Lamberts, R. (2011). Investigation of green roof thermal performance in temperate climate: A case study of an experimental building in Florianópolis city, Southern Brazil. Energy and Buildings.

[6] Alcazar, S., Bass, B. (2006.) Life cycle assessment of green roofs—case study of an eight-story residential building in Madrid and implications for green roof benefits. In: Proceedings of the Fourth North American Green Roof Conference: Greening Rooftops for Sustainable Communities, Boston, MA. The Cardinal Group, Toronto, 11–12 May.

[7] Rosheidat, Akram and Harvey Bryan. "Optimizing the effect of vegetation for pedestrian thermal comfort and urban heat island mitigation in a hot arid urban environment". http://www.ibpsa.us. N.p., 2010. Web. 28 Feb. 2017.

[8] Kumar, R. and Kaushik, S.C., (2005) Performance evaluation of green roof and shading for thermal protection of buildings. Building and Environment. Vol. 40 (11), pp 1505-1511.

[9] Chen, C. (2013). Performance evaluation and development strategies for green roofs in Taiwan: A review. Ecological Engineering; 52, pp.51--58.

[10] Getter K.L. and Rowe D.B., (2006) the role of extensive green roofs in sustainable development. Hort Science 41 (5), 1276-1285.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 175 - 190

Jordan Journal of Mechanical and Industrial Engineering

A New Framework of Reliability Centered Maintenance

Islam H. Afefy*[a], A. Mohib[a], A. M. El-kamash[b], M. A. Mahmoud[c]

[a] Industrial Engineering Department, Faculty of Engineering, Fayoum University, Egypt
[b] Professor, Hot Laboratory, Atomic Energy Authority of Egypt
[c] P.G. Student, Industrial Engineering Dept., Faculty of Engineering, Fayoum University, Fayoum, Egypt

Received 28 July. 2018

Abstract

In this paper, a novel framework of reliability centered maintenance (RCM) is proposed. The objective of this RCM model is to overcome the shortcomings of the existing RCM models. Current RCM models neither propose any actions with non-critical equipment nor propose any maintenance metrics. Moreover, they do not present maintenance work order flow. Both the classical and streamlined RCM models are adopted to formulate the proposed model to be more effective and efficient, i.e. it is focusing on the main functions of the system for the sake of preventing or eliminating the maintenance actions that are not necessary and identify effective maintenance tasks. In addition, a general Excel-based algorithm is proposed to perform criticality analysis and classification of machines/equipment into different categories. The proposed framework of RCM has been applied and evaluated in a real case study; namely is Fayoum Sugar Works Company in Egypt which produces Sugar. The results of applying RCM on A-Sugar line show that corrective and preventive maintenance downtime decreased by 55.77% and 52.17%, respectively. This reduction in downtimes leads to a saving in the total maintenance cost by 52.17%, which means that the proposed RCM saved about 6.19×106 L. E (Egyptian Pound) in total maintenance cost. Moreover, the results reveal that the availability increase from 57.1% to 90.74% and reliability increased from 99.73% to 99.88% as well.

Keywords: RCM, Availability, Maintenance metrics, Framework;

1. Introduction

Over the last twenty years, maintenance has changed. The changes are due to a huge increase in the number and variety of facilities (plant and equipment) that must be maintained throughout the world, more complex designs, new maintenance techniques and changing views on maintenance organization and responsibilities. RCM is a systematic approach to determine the maintenance requirements of plant and equipment in its operation [1]. It employs preventive maintenance, predictive maintenance (PdM), real-time monitoring (RTM), run to failure and proactive maintenance. These techniques are an integrated manner to increase the probability that a machine or component will function in the required manner over its design life cycle with a minimum of maintenance [2]. The aim of RCM is to create such maintenance strategy that helps minimize the total operating costs while increasing reliability of the system [3]. Diego Piasson et al [4] introduces a new approach for reliability-centered maintenance programs in electric power distribution systems based on a multi-objective genetic algorithm. A framework for application of reliability centered maintenance in the lead oxide production system was studied by Nafis Ahmad et al [5]. Implementation of failure mode and effect criticality analysis (FMECA) and fishbone techniques in reliability centred maintenance planning presented by Tamer El-Dogdog et al [6]. RCM analysis of process equipment is studied by Majid et al [8]. This study focused on RCM analysis applied to process equipment with heat exchangers as a case study. Islam [9] introduced maintenance planning based on computer-aided preventive maintenance policy. Selvik and Aven [10] explained a framework for reliability and risk centered maintenance. The purpose of their paper is to motivate Reliability and Risk Centered Maintenance (RRCM) methodology and describes its main features.

An expert system for reliability- centered maintenance in the chemical industry presented by Fonseca and Knapp [11]. Dacheng et al [12] explained study and application of reliability-centered Maintenance considering radical maintenance. The researchers in this paper made a combination of radical maintenance (RM) and traditional RCM to improve the quality of maintenance strategies. Strategic maintenance-management in Nigerian industries was discussed by Mark et al. [13]. Implementation of the RCM methodology on the example of city waterworks is introduced by Zoran Petrović et al [14]. Dawane and

* Corresponding author e-mail: islamhelaly@yahoo.com

Sedani introduced to study and investigations of RCM methodology in the manufacturing industry to minimize breakdown maintenance [15]. A framework for identification of maintenance significant items in reliability centered maintenance introduced by Yang Tang et al [16]. Samuel et al [17] investigated RCM study for an individual section-forming machine. Effectiveness of RCM program for power transformer is discussed by Burnet O'Brien Mkandawire et al [18]. In our previous works, Islam et al [19] investigated a model of reliability, availability, and maintainability (RAM) for industrial systems evaluations and implementation of framework RCM Made Simple approach is introduced [20].

In this section, we have reviewed various maintenance types applied to manufacturing and process industries. Most of the researchers have focused on classic RCM. Some researchers used software and others didn't. All results lead to the importance of the planned maintenance and preference of RCM. These become important issues in situations when applied the proposed framework RCM instead of classic RCM to a real case study. Moreover, excel program is used to perform maintenance program that would minimize maintenance cost and improve system availability and reliability of industrial processes. It is used in criticality analysis, functional failure analysis, and FMECA.

In this paper, a proposed framework of RCM model is presented to overcome some RCM classic shortcomings, such as no action for non-critical equipment, non-taking maintenance metrics, and not including maintenance work order flow. Furthermore, RCM classic should consider all factors that lead to failures, this increases the maintenance downtime. It should be noted that the major contribution of this work is to develop the framework of RCM implementation from which analysis of maintenance metrics and downtime metrics. Integrating between RCM Group and maintenance group to overcome some shortcoming in maintenance tasks leads to improvement in maintenance effectiveness and maintenance key performance indicators (KPIs). The proposed RCM and framework of RCM implementation has been applied and evaluated in a real industrial process, illustrating the effectiveness of the proposed RCM.

In addition, the framework of proposed RCM methodology describes all necessary steps thoroughly. This detailed description does not exist in classic RCM and is not clear in standard RCM. It selects the critical items based on system functions and functional failures analysis besides the failure history analysis, which helps the readers and researchers to implement the RCM in general. As for the difference between framework of proposed RCM methodology and the standard RCM technique, the standard RCM technique does not clarify the methodology of implementation.

2. Proposed RCM Model

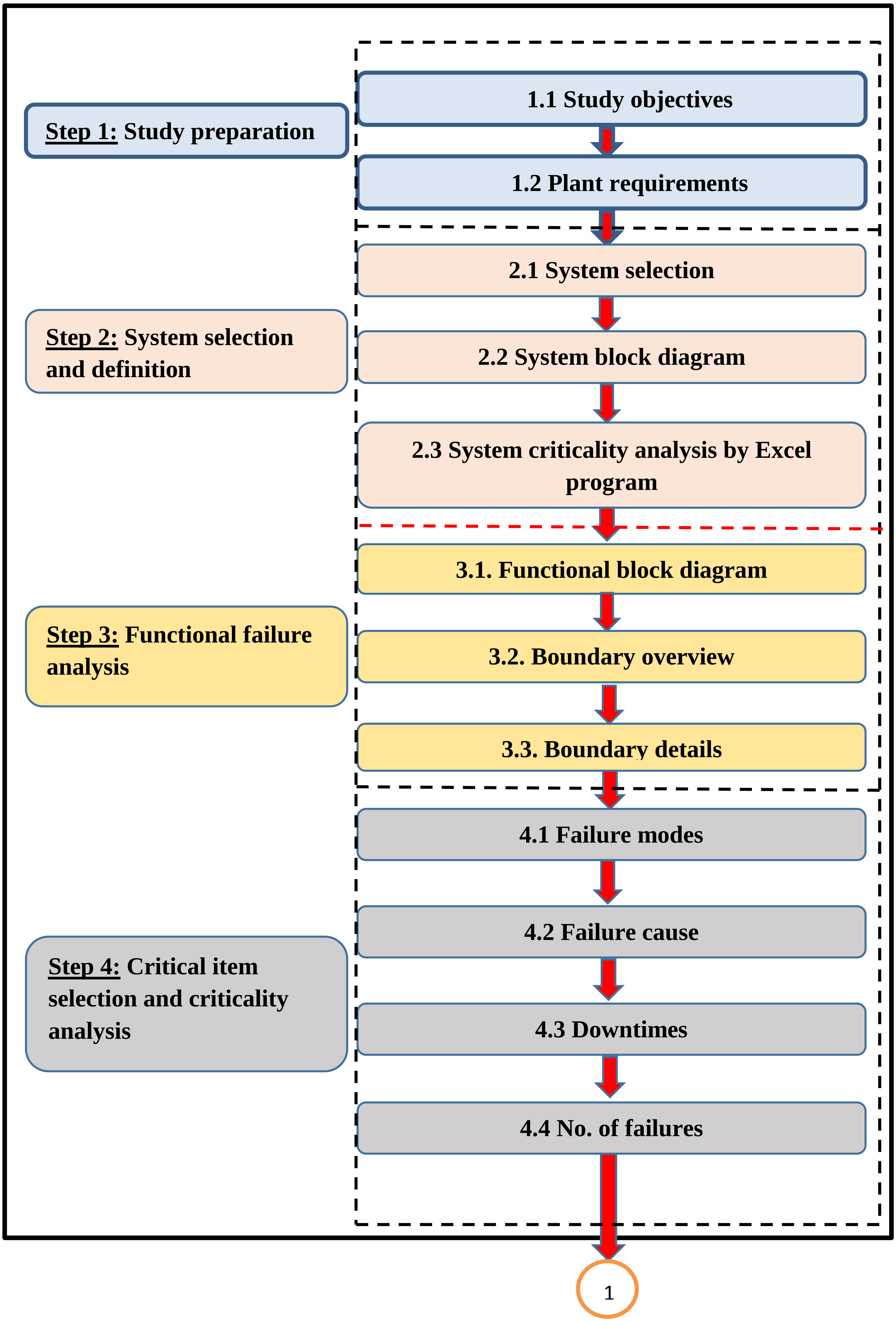
The proposed RCM methodology steps are discussed in addition to some of the analysis tools used in the real case study and applied in large industrial processes as follows:

1. Study preparation.
2. System selection and definition
3. Functional failure analysis (FFA)
4. Critical item selection and criticality analysis.
5. Data collection and analysis
6. Failure mode and effect criticality analysis (FMECA)
7. Logic tree analysis (LTA)
8. Selection of maintenance actions
9. Determination of maintenance intervals
10. Default action for non-critical equipment
11. Planned maintenance comparison
12. Framework of RCM Implementation.

Figure 1 shows that the proposed policy of RCM starting with study preparation of system works plant then select the most important plant, defined its boundary and critical equipment selected by applying criticality analysis. Functional block diagram (FBD) for the plant is carried out. All critical equipment failures in last years collected and analysed with respect to failures and downtime. Why-Why technique is used to analyse critical item failures. FMECA is carried on for failures, then LTA applied to select the suitable maintenance approach for each failure and used in maintenance action selection. After that, the default action for non-critical items is illustrated and planned maintenance comparison is performed. Finally, a proposed RCM plan is implemented and gets results.

2.1. Study Preparation

In this step, a proposed RCM project group is established. The project group must define and clarify the scope and the objectives of the analysis. RCM project group must define and clarify the objectives and the scope of the analysis.


Step 1: Study preparation
1.1 Study objectives
1.2 Plant requirements
Step 2: System selection and definition
2.1 System selection
2.2 System block diagram
2.3 System criticality analysis by Excel program
Step 3: Functional failure analysis
3.1. Functional block diagram
3.2. Boundary overview
3.3. Boundary details
Step 4: Critical item selection and criticality analysis
4.1 Failure modes
4.2 Failure cause
4.3 Downtimes
4.4 No. of failures
1

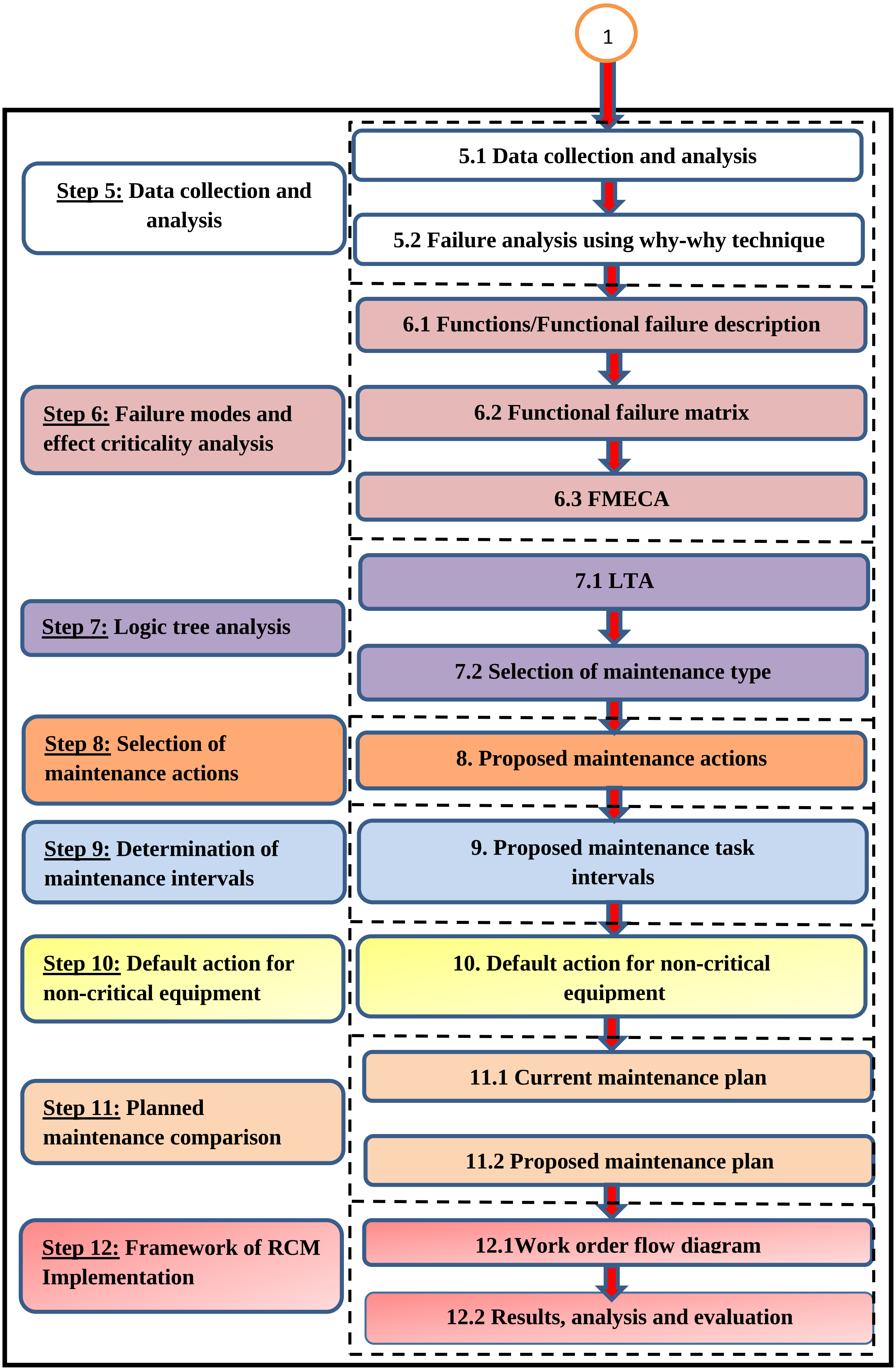


Figure 1. Framework of proposed RCM methodology

2.2. System Selection and Definition

A system is defined as a group of equipment, components, or facilities that support an operational function. These operational functions are identified by criticality analysis (See Figure 2) [21].

2.3. Functional Failure Analysis (FFA)

Functional failures explain the ways in which a subsystem or system can fail to meet the functional requirements designed for the equipment [22]. The main objectives of FFA step are as follows;

- To identify the ways in which the system might fail to function
- To list input interfaces necessary for the system to operate.
- To describe and identify the system's required functions and performance criteria.

2.4. Functional Block Diagrams (FBD).

The FBD displays all components of a system, their functional relationships to one another, and in and out interfaces with other systems. The FBD is also useful as a basis for the FMECA in the RCM analysis process. It is generally not required to establish the FBD for all the system functions. The diagrams are, however, efficient tools to illustrate the input interfaces to a function.

2.5. Root Cause Failure Analysis

The goal of RCFA is to identify the contributing causal factors that have led to a performance problem. Also finding the real cause of the problem and dealing with it rather than simply continuing to deal with the symptoms. When problems occur, use the tools to understand and gain insight into the causes before making changes based on assumptions. [23]. A fishbone diagram is a visual way to look at cause and effect. It is a more structured approach than some other tools available for brainstorming causes of a problem (e.g., the Five Whys tool). Five whys strategy worksheet is introduced in table 1.

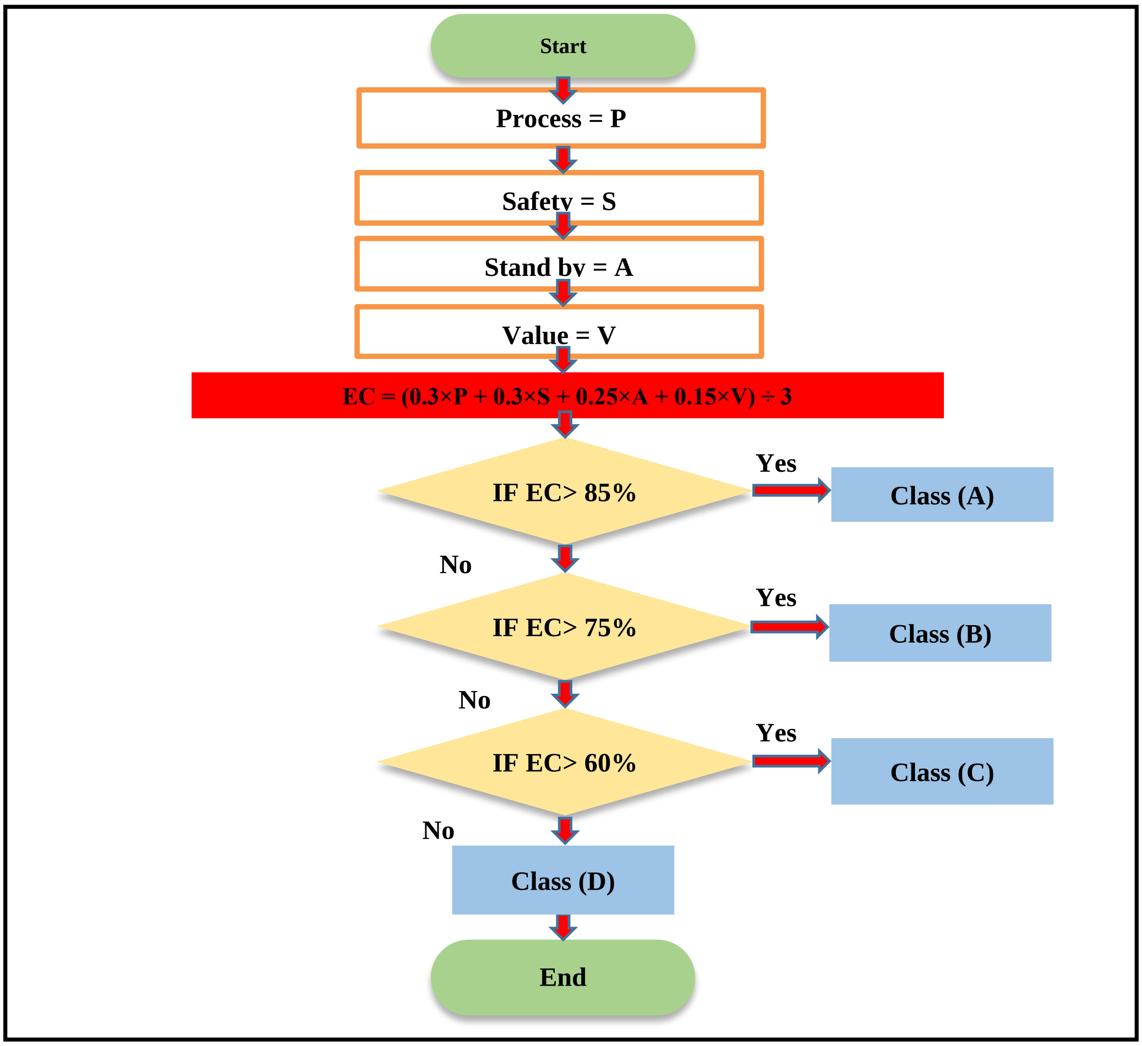


Figure 2. General algorithm diagram used to calculate the criticality value [21]

2.6. Critical Item Selection

The objective of this step is to identify the analysis items that are potentially critical with respect to the functional failures identified in FFA step.

2.7. Data Collection and Analysis

The purpose of this step is to establish a basis for both the qualitative and quantitative analysis. The data necessary for the RCM analysis may be categorized into the following three groups:

- Group 1: Design data
- Group 2: Operational and failure data
- Group 3: Reliability data

2.8. FMECA

FMECA is the procedure, which consolidates a failure mode and effects analysis and a criticality analysis. It is utilized to document and identify the functions, functional failures, failure modes and failure effects of an item. Besides that, it determines the significance of functional failures in terms of operations, economics, safety, and environment. It also classifies the severity of every failure effect according to criteria of established severity classification and provides failure rate information.

2.9. Logic Tree Analysis

LTA process is the step used to determine the most applicable, cost-effective maintenance tasks for a component [23]. LTA is used to assess the relationship between failure mode and each part with a high maintenance priority [24]. The input to LTA is the failure modes from the FMECA. The main idea is for each failure mode to decide whether a preventive maintenance task is suitable, or it will be best to let the item deliberately run to failure and afterward carry out a corrective maintenance task. RCM logic tree is shown in figure 3.

2.10. Selection of Maintenance Actions

The maintenance task selection process uses various forms of logical decision making to arrive at conclusions in a systematic manner. The outcomes can include:

- Preventive maintenance.
- Condition monitoring.
- Inspection and functional testing.
- Run to Failure.

2.11. Determination of Maintenance Intervals

Maintenance actions are divided into corrective maintenance (CM) and preventive maintenance (PM). CM actions are unscheduled and are intended to restore a system from a failed state to a working state through either replacement or repair of failed components. On the other hand, PM is scheduled active and can be carried out either to reduce the likelihood of a failure or to improve the availability and reliability of the system [25].

2.12. Default Action for Non-Critical Equipment

It is important to include the non-critical machines or items in the maintenance policy. The catalogues of these machines are enough to prevent sudden failures because of the high total cost of maintenance. The sudden failure of these machines has no effect on production and safety.

2.13. Planned Maintenance Comparison

The outputs of the analysis will result in a change to the maintenance program. It is important that such changes are consistent with the maintenance philosophy of the plant and with regulatory and social obligations. For this reason, it is important that the process and it is outcomes be subjected to a final review.

2.14. Framework of RCM Implementation

As shown in figure 4, framework of RCM Implementation is presented. Once the proposed framework of RCM implementation approved, the final step is to implement the proposed maintenance tasks in a real case study. The maintenance tasks are then fed into suitable maintenance planning and control systems, while revised operating procedures are usually incorporated into standard operating procedure manuals. Furthermore, the framework includes maintenance metrics such as mean time between failures (MTBF), mean time to repair and mean downtime.

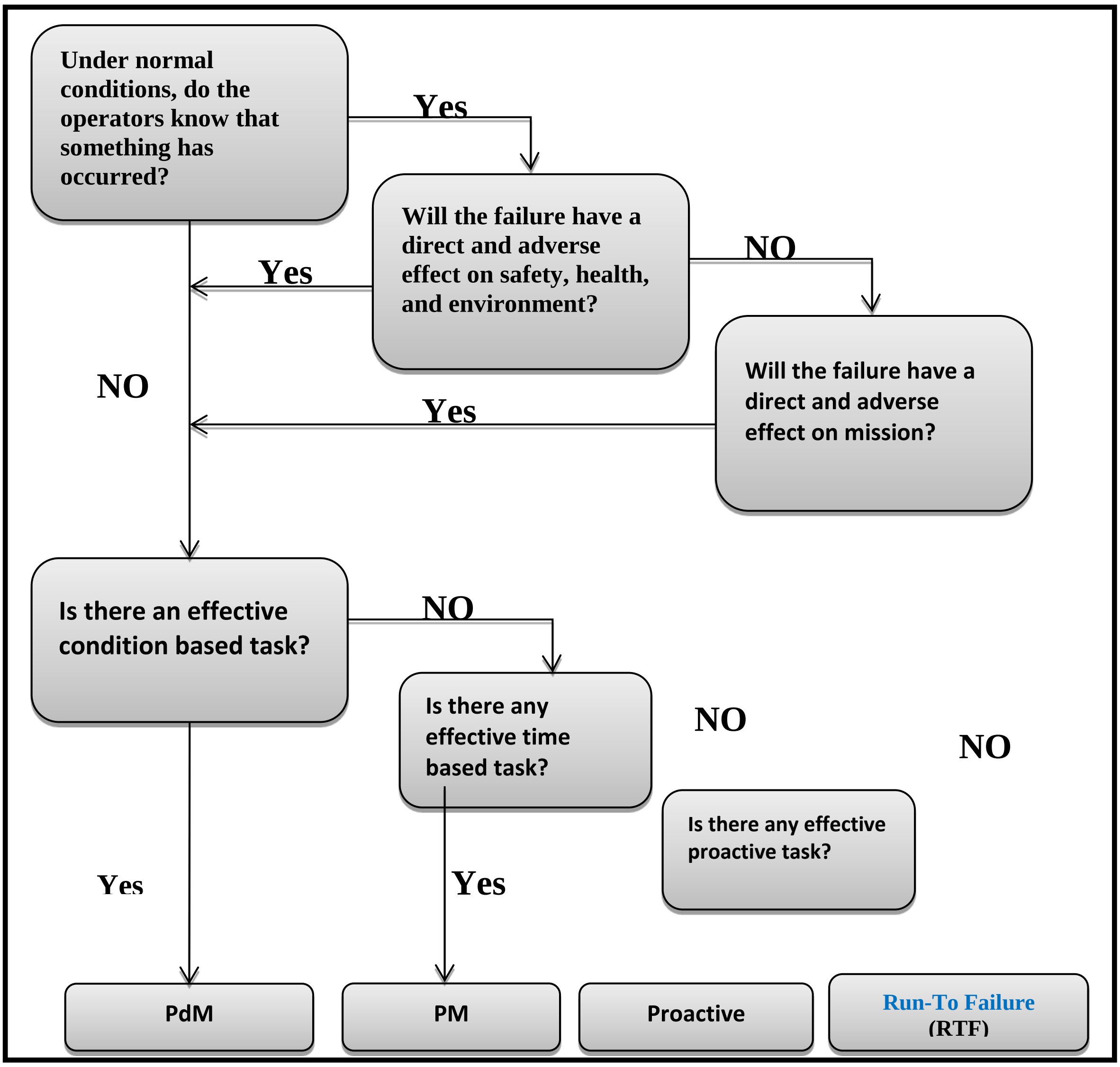


Figure 3. RCM logic tree

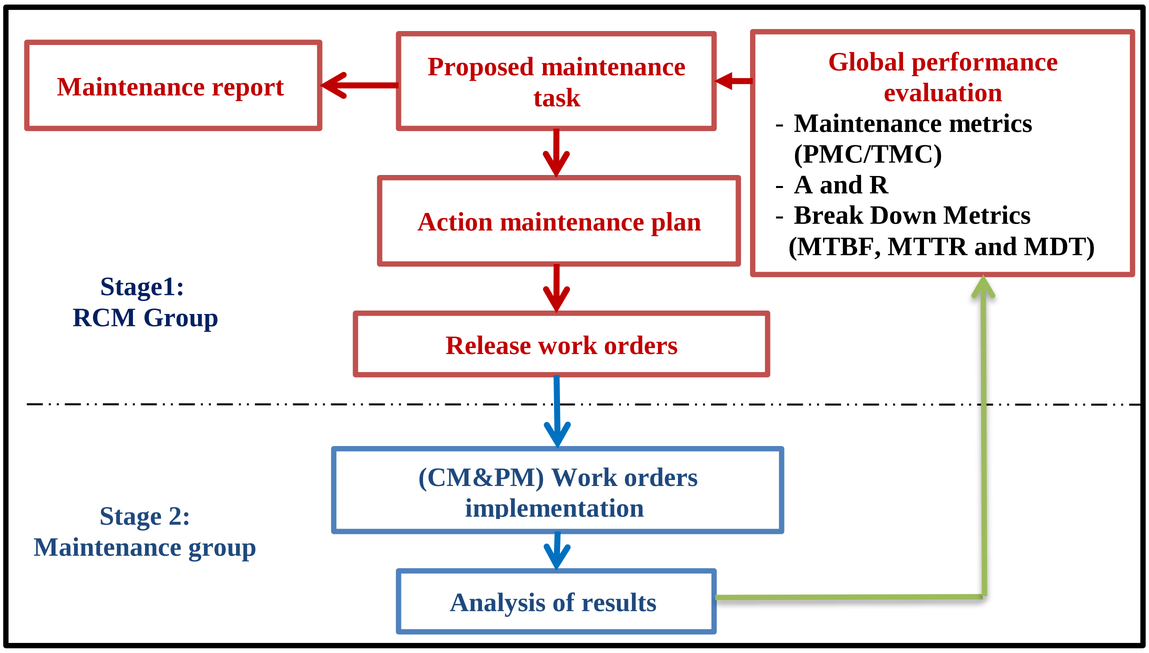


Figure 4. Framework of RCM Implementation

3. Modeling of Availability and Mean Time between Failure

Availability is the probability that a material, component, equipment, system or process is in its intended functional condition at a given time and therefore is either in use or capable of being used in a stable environment [9 &28]. It is a measure of the degree to which an item is in an operable state and can be committed at the start of a mission when the mission is called for at an unknown point in time and it is defined mathematically as:

$$A=\frac{MTBF}{(MTBF+MDT)} \quad (1)$$

$$MTBF=\frac{AT-(N_f \times MDT)}{N_f} \quad (2)$$

$$\text{Failure rate } (\lambda)=\frac{N_f}{AT-(N_f \times MDT)} \quad (3)$$

Where, N_f is the number of failure, AT is the available time, MTBF is mean time between failure and MDT is mean downtime.

4. Modeling of Reliability

The characteristic of an item expressed by the probability that it will perform a required function under a stated condition for a stated period of time [28] and it is defined mathematically as:

$$R=\frac{\text{Planned time}-\text{Unplanned time}}{\text{Planned time}} \quad (4)$$

Where, planned time is the total operated time, and unplanned time is the total of corrective maintenance time.

5. Modeling of the Total Maintenance Cost

Total maintenance cost (TMC) can be calculated by the follow equation:

$$TMC = CMC + PMC + DTC \quad (5)$$

Where, CMC is the corrective maintenance cost, PMC is the preventive maintenance cost and DTC is the downtime cost.

6. Modeling of the PM Worker Size.

A specific formula is present to define the preventive maintenance worker size (WS_{PM}) as:

$$WSPM = MDa / WOC \quad (6)$$

Where, MD_a is the total preventive maintenance annual man-day, and W_{OC} is the workers operating conditions (day/year).

The total preventive maintenance annual man-day can be calculated as:

$$MDa= Fa \times tdu \times Wn \quad (7)$$

Where F_a is the annual frequency per preventive maintenance type, t_{du} is the duration time, and W_n is the number of workers per preventive maintenance type.

7. Case study

The present study has been applied for the practical Sugar-end plant of the Fayoum Sugar Works Company, EL-Fayoum, Egypt. The main products of the company are Sugar, Molasses and Beet Pulp. The main objective of the sugar-end plant is to crystallize sucrose present in the thick juice to granulated-refined sugar and molasses.

7.1. System Block Diagram

As shown in figure 5, the block diagram of Fayoum Sugar Works Company is introduced. This figure shows the main equipment of the system. It shows that the process flow of juice between equipment of the A-sugar line.

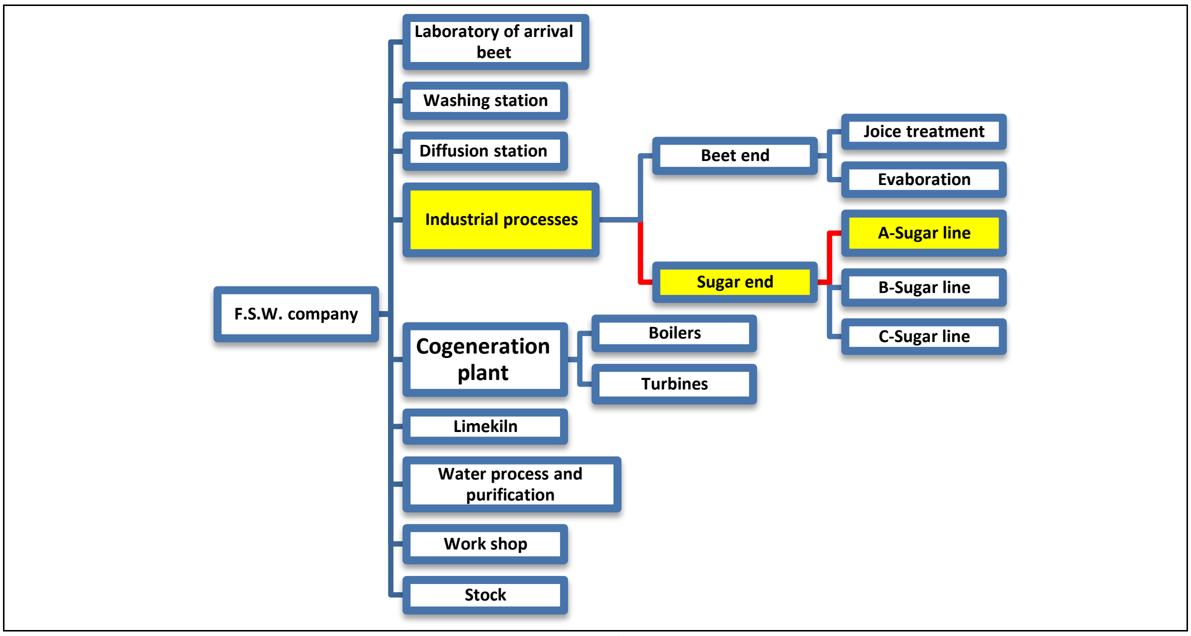


Figure 5. Block diagram of FSW Company.

7.2. System Criticality Analysis

General algorithm diagram used to calculate the criticality value (see figure 2). This algorithm shows the calculation steps of the equipment criticality [2]. Excel program established to perform criticality analysis for system equipment and grouping classified. Table 1 presents the result of system criticality and equipment selection

Table 1. System criticality analysis and equipment selection

Item No.	Component	P	S	A	V	EC%	Group
4.11	Wet sugar Bucket elevator	3	2	3	3	90	A
4.1	Sugar dryer	3	2	3	3	90	A
16.15	Wet sugar Vibrator (Gras hopper)	3	2	3	2	85	B
4.9	Dry sugar Bucket elevator	3	2	3	2	85	B
16.01	A-Feed Solution Tank	3	1	3	2	75	B
16.13	A-Sugar Distributor	3	1	3	2	75	B
16.04	A-Seed pump F500	2	2	3	1	70	C
4.12	Wet sugar Conveyor	3	1	3	1	70	C
4.8	Dry sugar Conveyor "dryer exiting"	3	1	3	1	70	C
4.08	Wet sugar Conveyor "dryer entering"	3	1	3	1	70	C
16.11	A-Massecuite pump F1000	2	1	3	2	65	C
4.16	Standard liquor tank	2	1	3	1	60	C
16.14	Centrifugal machineG1500 for A-Sugar No.1~6	2	2	1	2	58.3	D
16.05	A-Feed Solution Pump No.1, 2	2	2	1	2	58.3	D
16.02	A-Seed Pan	2	1	2	2	56.67	D
4.7	Dry sugar Vibrator	2	1	2	2	56.67	D
4.10	Fin Fan	1	1	3	2	55	D
16.06	A-VKT	2	1	1	3	53.3	D
16.03	A-Seed receiver tank	2	1	2	1	51.67	D
17.01	Under Scales conveyor (line No.1, 2, 3, 4)	1	2	2	1	51.67	D
17.02	Delivery conveyor (line No.1, 2, 3, 4)	1	2	2	1	51.67	D
16.19	A-Wash syrup Pump No.1, 2	1	2	1	2	48.3	D
4.18	Standard liquor tank Pump No.1, 2	1	1	1	1	33.3	D

7.3. Functional Block Diagram (FBD) of System

FBD of the system is shown in figure 6. It is the relation between equipment of the A-Sugar line of sugar end plant, which selected as a system to applied proposed RCM on it.

7.4. Critical Items Selection

Identifying the most critical equipment and the critical items in the system is based on total downtime. Table 2 shows downtimes and number of failures for the system. In addition to, the downtimes and number of failures for components of system equipment are introduced in table 3 and figure 7.

Table 2. Relation between no. of failures and downtimes for A-sugar line equipment

Equipment	No. of failure	Downtime (Min.)	% of downtime
Wet Sugar bucket elevator	2	90	5.14
Sugar dryer	2	55	3.14
Wet Sugar vibrator	9	343	19.60
Dry sugar Bucket elevator	8	165	9.43
A-Feed Solution tank	1	34	1.94
A-Seed pump F500	4	620	35.43
Wet sugar Conveyor	4	90	5.14
Dryer entering conveyor	3	60	3.43
Dryer exiting conveyor	4	93	5.31
A-Massecuite pump (F1000)	2	200	11.43
Total	39	1750	100

 © 2019 Jordan Journal of Mechanical and Industrial Engineering. All rights reserved - Volume 13, Number 3 (ISSN 1995-6665)

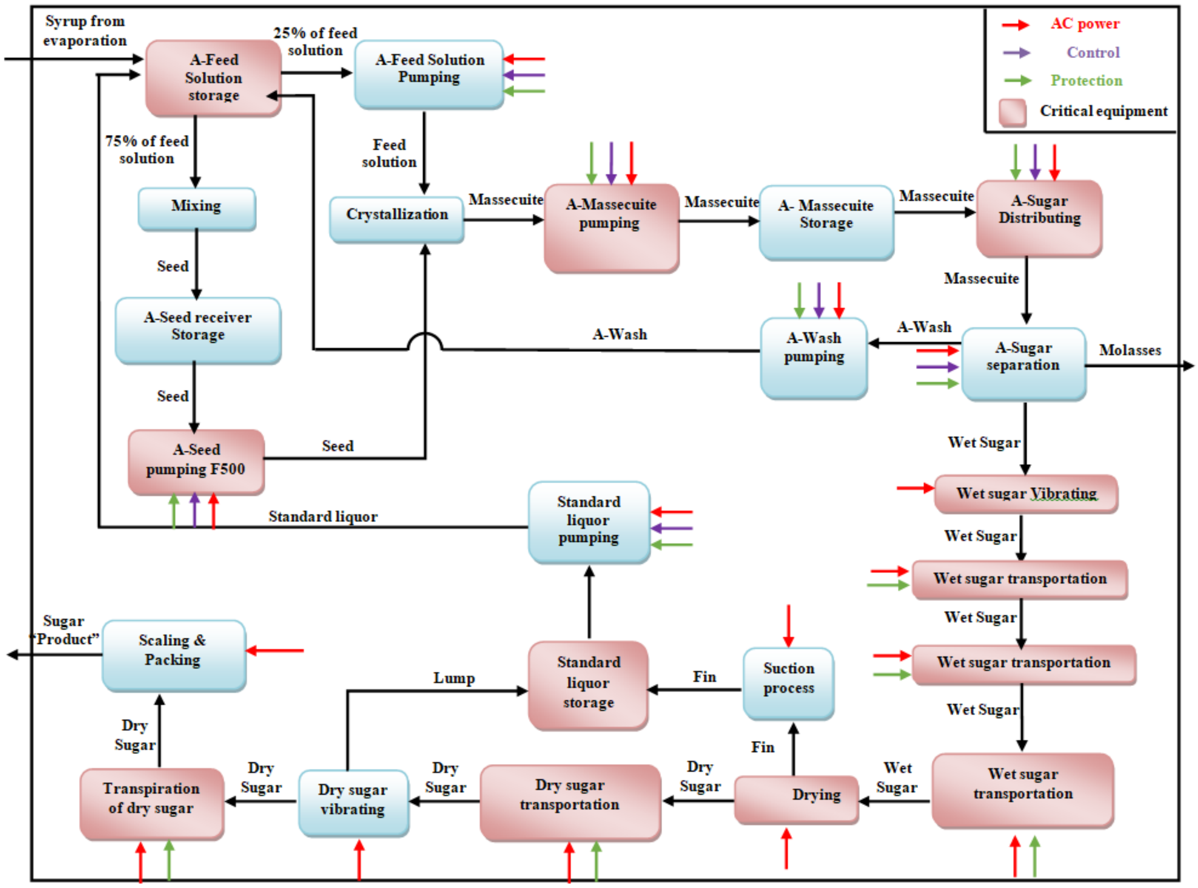


Figure 6. Functional block diagram of A-Sugar line End plant

Table 3. Relation between no. of failures, downtimes and MTBF for components of A-sugar line equipment.

Equipment	Component	No. of failure	Downtime (Min.)	% of downtime	MTBF (hr.)
Wet Sugar bucket elevator	Motor (over load)	1	60	3.49	180
	Level device	1	30	1.74	180
Sugar dryer	Motor (over load)	1	25	1.45	180
	Level device	1	30	1.74	180
Wet Sugar vibrator	Trough	7	298	17.33	25.7
	Bolts	1	15	0.87	3
Dry sugar Bucket elevator	Main conveyor (elongation)	6	90	5.23	30
	Main conveyor (slipping)	1	45	2.62	180
	Level device	1	30	1.74	180
A-Feed Solution tank	Motor	1	34	1.98	180
A-Seed pump F500	Piston clearance	4	620	36.05	45
Wet sugar Conveyor	Main conveyor (slipping)	1	15	0.87	180
	Motor's conveyors	2	60	3.49	90
	Rollers	1	15	0.87	180
Dryer entering conveyor	Motor's conveyors	1	30	1.74	180
	Rollers	2	30	1.74	90
Dryer exiting conveyor	Motor's conveyors	2	58	3.37	90
	Main conveyor (elongation)	1	20	1.16	180
	Rollers	1	15	0.87	180
A-Massecuite pump (F1000)	Piston clearance	1	180	10.47	180
	Spring	1	20	1.16	180

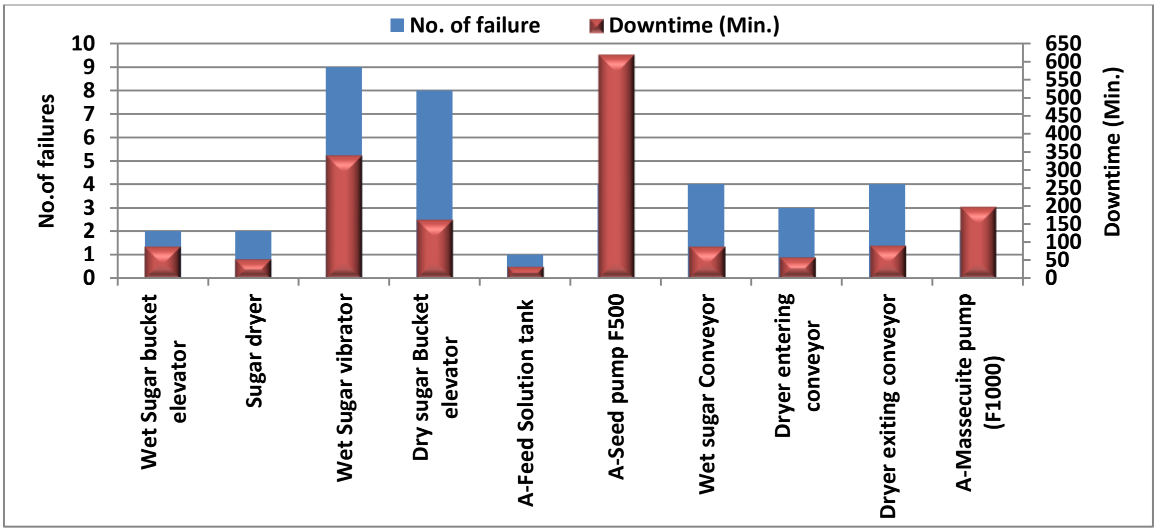


Figure 7. Relation between no. of failures and downtime

7.5. Failure Analysis Using Why-Why Technique

Tables from (4) to (8) shows that the equipment and its function, failure mode and levels analysis.

Table 4. Failure modes at different level of details for wet sugar vibrator

- Equipment: Wet Sugar vibrator
- Function: Transfer and distribute sugar from Centrifuges to wet sugar conveyor
- Failure mode: High sound

Level 1	Level 2	Level 3	Level 4	Level5
Trough unit operate with high vibration level	Crack in Trough	Over load	Sequence of centrifuges isn't correct	Fault in operation
		Material of trough		
	Bolts looseness	Assembly error	Poor maintenance	

Table 5. Failure modes at different level of details for dry sugar bucket elevator

- Equipment: Dry sugar Bucket elevator
- Function: Transfer and distribute sugar from Centrifuges to wet sugar conveyor
- Failure mode: Irregular feeding with low capacity and low speed

Level 1	Level 2	Level 3	Level 4	Level5
Main conveyor Slipping	Loosens of conveyor	Elongation of conveyor	Tension	Assembly error
Main conveyor Rupture	Slipping of conveyor and not stopped	Sensor not working		

Table 6. Failure modes at different level of details for A-Seed pump F500

- Equipment: A-Seed pump F500
- Function: Pumps are employed for delivery massecuite
- Failure mode: Delivery massecuite by flow rate less than specified

Level 1	Level 2	Level 3	Level 4	Level5
Piston high clearance	Wear	Erosion	Suspend solid	
Viscosity higher than specified	Low temperature	Fault in operation		
Line resistance too high	Fossilized sugar	Low temperature	Fault in operation	

Table 7. Failure modes at different level of details for A-Massecuite pump (F1000)

- Equipment: A-Massecuite pump (F1000)
- Function: Pumps are employed for delivery massecuite
- Failure mode: Low flow rate

Level 1	Level 2	Level 3	Level 4	Level5
Piston high clearance	Wear	Erosion	Suspend solid	
Viscosity higher than specified	Low temperature	Fault in operation		
Line resistance too high	Fossilized sugar	Low temperature	Fault in operation	

Table 8. Failure modes at different level of details for dryer exiting conveyor

- Equipment: Dryer exiting conveyor
- Function: Conveys dry sugar from dryer
- Failure mode: Fails to transfer some sugar

Level 1	Level 2	Level 3	Level 4	Level5
Motor's conveyors Consumed	High friction between pulley and conveyors	High tension	Poor maintenance	
Excessive wear on bottom side of belt.	Slippage between belt and drive pulley	Low tension	Assembly error	
			Conveyor elongated	

7.6. FMECA

FMECA work sheet represented in table 9, there are twelve failure modes to apply FMECA on it. FMECA shows the effect of every failure mode on the equipment, system and plant. In addition to, calculate risk priority number (RPN= S × O × D) of them. Where, S is the rank of the severity of the failure mode, O is the rank of the occurrence of the failure mode and D is the rank of the likelihood the failure will be detected.

7.7. Logic Tree Analysis

The LTA information of system is introduced in table10.

Table 9. FMECA worksheet

FF#	Equip. #	Equipment description	FM#	Failure mode	FC#	Failure cause	Failure effect			S	O	D	RPN
							Local	System	Plant				
3.1.1	3	Wet sugar vibrator	3.01	Crack in vibrator trough	03.01.1	Over load	High sound	Stopped	Stopped	5	8	8	320
			3.02	Loosening of bolts	03.02.1	Poor assembly	High sound	Stopped	Stopped	6	4	8	192
	4	Dry sugar Bucket elevator	4.01	Bearing deterioration	04.01.1	Age/wear out	Inoperative	Stopped	Stopped	3	4	6	72
			4.02	Belt slipping	04.02.2	Belt elongation	Slipping	Low performance rate	Reduced production	3	6	4	72
1.1.1	6	A-Seed pump F500	6.01	High piston clearance	06.01.1	Wear	Erratic pump operation	Reduced flow	Reduced production	3	3	6	54
			6.02	Bearing deterioration	06.02.1	Age/wear out	Erratic pump operation	Reduced flow	Reduced production	3	3	6	54
			6.03	Mechanical seal failed	06.03.1	Leakage	Trip	Low performance rate	low performance rate	3	5	5	75
3.1.1	9	Dryer exiting conveyor	9.01	Belt slip	09.01.1	Insufficient traction between the belt and pulley	Erratic operation	Reduced flow	Reduced production	5	3	4	60
				Belt elongation	09.01.2	High tension	Stopped	Stopped	Stopped	3	4	4	48
2.1.1	10	A-Massecuite pump (F1000)	10.01	High piston clearance	10.01.1	Wear	Erratic pump operation	Reduced flow	Reduced production	3	3	4	36
			10.02	Bearing deterioration	10.02.1	Age/wear out	Erratic pump operation	Reduced flow	Reduced production	3	3	4	36
			10.03	Mechanical seal failed	10.03.1	Leakage	Trip	Low performance rate	low performance rate	3	5	5	75

7.8. Selection of Maintenance Actions

In this step, set of actions performed to failure modes depends on the maintenance strategy type, which were selected in the past step by using LTA. Table 10 shows the proposed maintenance task.

7.9. Determination of Maintenance Intervals

In this step, selection of maintenance intervals for A-sugar line is carried out. After choosing the type of maintenance and tasks suitable for each failure mode it should be determined the suitable maintenance intervals daily, weekly, monthly or yearly and it depends on MTBF and RPN. In table 10, determination of maintenance intervals of A-sugar line presented.

7.10. Default Action for Non-Critical Equipment

A remaining question is what to do with the items that are not analyzed. For equipment already has a maintenance program, it is reasonable to continue this program. If there is no maintenance program the default action for is RTF based maintenance. The catalog troubleshooting analysis of these machines is very enough to follow it to prevent suddenly failures. The maintenance policy of non-critical equipment based on the catalog is very enough.

7.11. Planned Maintenance Comparison

After the task selection and determination of intervals has been completed and reviewed, the maintenance actions arising from the task selection between "process" and "compared" against the current maintenance practices. The purpose of this comparison is to identify the changes needed to the maintenance program and the impact on resources and other commitments. Comparison between current and proposed maintenance tasks are shown in table 10.

Table 10. Logic tree analysis, current and proposed maintenance task.

Failure mode	Failure cause	Source of Failure	LTA						Current maintenance task	Proposed maintenance		
			1	2	3	4	5	6		Maintenance type	Task	Interval
Crack in vibrator trough	Over load	Sequence of centrifuges isn't correct	N	-	-	N	Y	-	RTF	PM	- Visual inspection - Check sequence - Welding cracks	W
Loosening of bolts	Poor assembly	Poor maintenance	N	-	-	N	N	Y	RTF	Proactive	Check &Retighten	M
Bearing deterioration	Age/wear out	Incorrect assembly	N	-	-	Y	-	-	RTF	PdM	Vibration measure	M
Belt slipping	Belt elongation	High tension	Y	N	Y	N	N	Y	RTF	Proactive	Check belt & tension	W
High piston clearance	Wear	Suspend solid	N	N	Y	N	Y	-	RTF	PdM	Flow rate measure	D
Bearing deterioration	Age/wear out	Incorrect assembly	N	-	-	Y	-	-	RTF	PdM	Vibration measure	M
Mechanical seal failed	leakage	Deterioration	Y	N	N	-	-	-	RTF	RTF	RTF	-
Belt slip	Insufficient traction between the belt and pulley	Incorrect assembly	N	-	-	N	N	Y	RTF	Proactive	Check belt & tension	W
Belt elongation	High tension	High tension	Y	N	Y	N	N	Y	RTF	Proactive	Check belt & tension	W
High piston clearance	Wear	Suspend solid	N	N	Y	N	Y	-	RTF	PdM	Flow rate measure	D
Bearing deterioration	Age/wear out	Incorrect assembly	N	-	-	Y	-	-	RTF	PdM	Vibration measure	M
Mechanical seal failed	leakage	Deterioration	Y	N	N	-	-	-	RTF	RTF	RTF	-

8. Results

Table 11 presents a comparison between results of before and after applying a proposed policy of RCM on the real case study. The result shows that corrective maintenance downtime decreased from 9.7 hours/year to 4.3 hours/year, therefore, its cost decreased from 3.88×10^6 L.E/year to 1.72×10^6 L.E/year. In addition, preventive maintenance is decreased from 20 hours/year to 10 hours/year and its cost decreased from 8×10^6 L.E./year to 3.97×10^6 L.E./year. This means that proposed RCM saved 6.19×10^6 L.E in total maintenance cost. System availability and reliability are improved as well. Availability is increased from 57.1% to 90.74% and reliability increased from 99.73% to 99.88%.

Figure 8 clearly depicts the difference between corrective maintenance downtime before and after applying a proposed policy of RCM in A-Sugar line. Also, the comparison between corrective maintenance cost before and after applying a proposed policy of RCM in A-Sugar line is presented in figure 9. As shown in figure 10, the comparison between preventive maintenance downtime before applying RCM and preventive maintenance downtime after applying a proposed policy of RCM for A-Sugar line is illustrated. Moreover, in figure 11, preventive maintenance cost before and after applying RCM for A-sugar line is discussed. Figure 12 shows the comparison of total maintenance cost before and after applying a proposed policy of RCM in A-Sugar line. As in figure 13, the availability analysis for A-sugar line is introduced and reliability analysis is presented in figure 14.

Table 11. Results of applying proposed RCM framework on A-Sugar line

No	Item	Before Current	After Applying RCM	Improvement %
	Corrective maintenance down time (hrs./year)	9.7	4.3	55.77
	Corrective maintenance cost $\times10^6$ (L.E/year)	3.89	1.72	55.77
	Preventive maintenance down time (hrs./year)	20	10	50
	Preventive maintenance cost$\times10^6$ (L.E/ year)	8	3.97	50.4
	Total maintenance cost $\times10^6$ (L.E)	11.89	5.69	52.17
	Availability (%)	57.1	90.47	36.98
	Reliability (%)	99.73	99.88	0.15

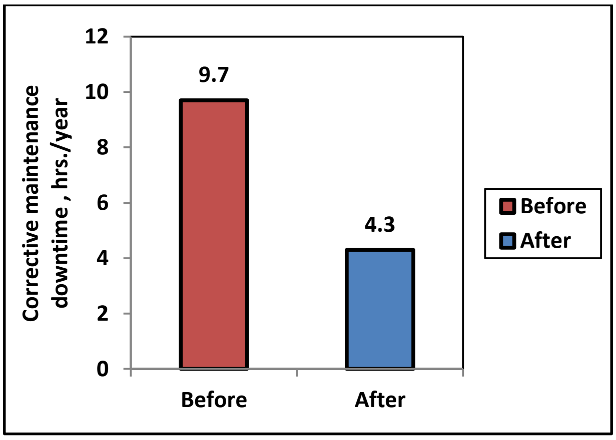


Figure 8. Corrective maintenance downtime for system

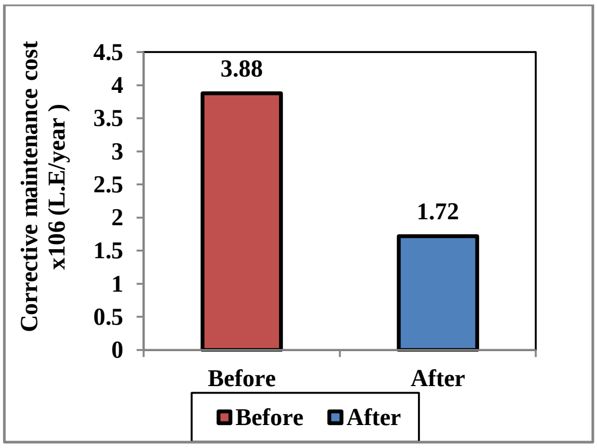


Figure 9. Corrective maintenance cost for system

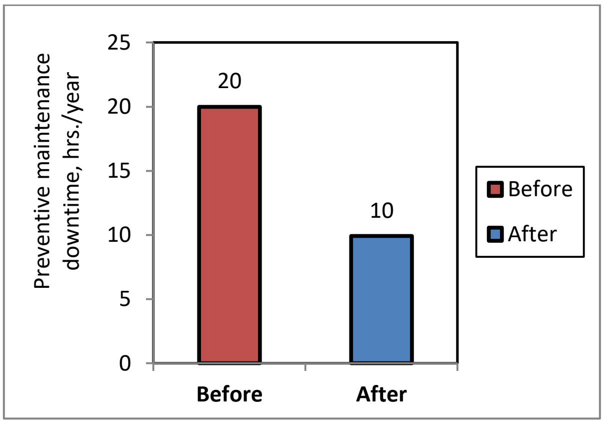


Figure 10. Preventive maintenance downtime for system

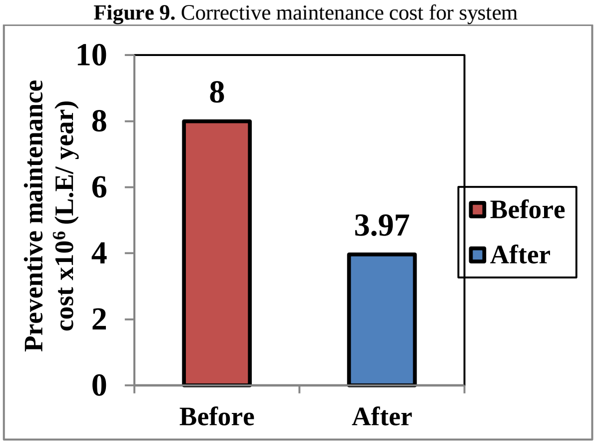


Figure 11. Preventive maintenance cost for system

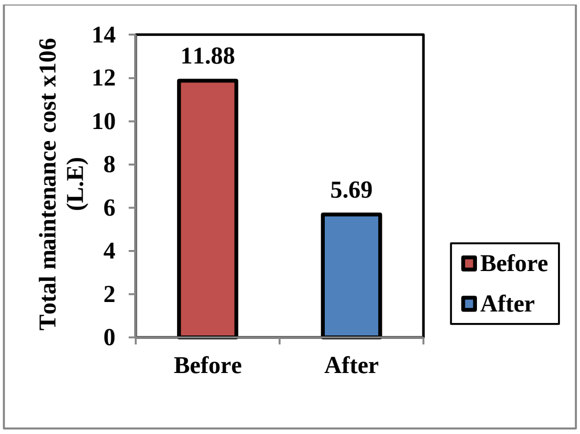


Figure 12. Total maintenance cost for system

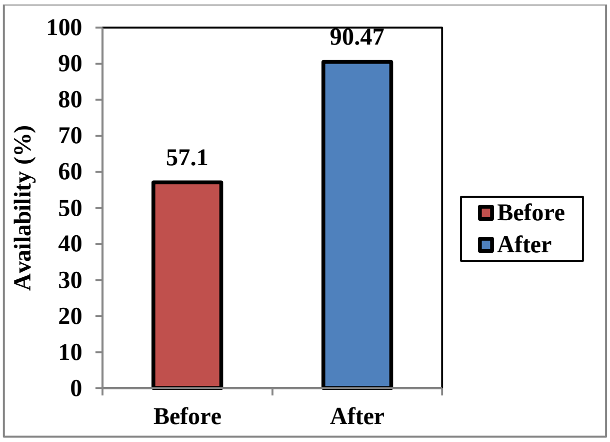


Figure 13. The availability analysis for system

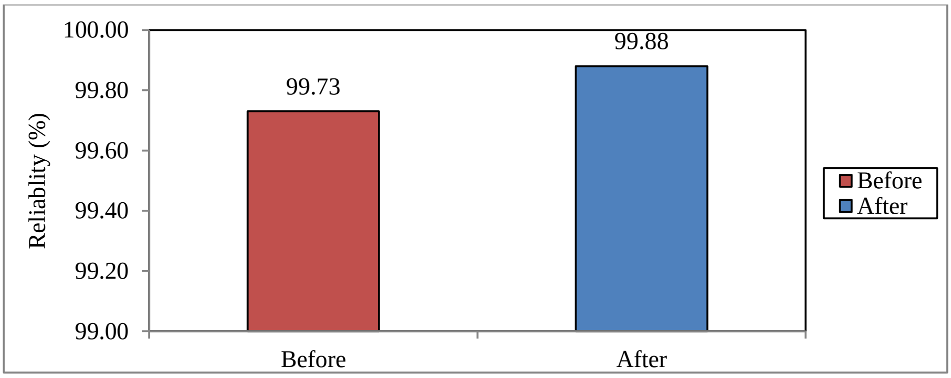


Figure 14. The reliability analysis for system

A comparison between the proposed RCM framework and other methods are introduced in table 12. The other methods include simple mode RCM [20], rational RCM [29], classical RCM [30], and group maintenance method [31]. Each of those methods has been validated in a different application, e.g., power distribution systems, manufacturing plant and CNC machine. The comparison between the proposed RCM framework and those methods is based on downtime reduction (%), reliability, availability (%), and maintenance cost reduction (%). We conduct the comparison with the available published results in their respective studies where not all measures are reported in those studies. The comparison indicates that the other methods succeeded to reduce the downtime between 33.15% and 38% while the proposed RCM framework realized a downtime reduction of 55.77%. The other methods succeeded to achieve a reliability that ranges between 87.2% and 92.1 while the proposed RCM framework improved the reliability to 99.88%. The maintenance cost reduction has also improved by 34.5% under the previous methods while the proposed RCM framework reduced the maintenance cost by 52.17%. As such, this comparison implies that the proposed RCM framework outperforms the other existing methods.

Table12. A comparison between the proposed RCM framework and relevant methods

No	Item	Relevant Methods				Proposed Framework RCM in present work
		RCM simple mode approach [20]	Rational RCM (RRCM) approach [29]	Classic RCM [30]	Group maintenance method [31]	
1	Application	-	Power distribution systems	Manufacturing plant in Poland	CNC Machine	Sugar Works Company in Egypt
2	Reduction downtime, %	Average 33.15		38	-	55.77
3	Reliability, %	-	87.2	-	Max. 92.1	99.88
4	Availability ,%	-	Mean availability 84	-	-	90.47
5	Reduced maintenance cost,%	-	34.5	-	-	52.17

9. Conclusions

A novel framework of RCM has been proposed and applied to a sugar–end plant in Fayoum Sugar Works Company. The new finding of this work shows that corrective maintenance downtime decreased from 9.7 hours/year to 4.3 hours/year and its cost decreased from 3.88×10^6 L.E/year to 1.72×10^6 L.E/year. In addition, preventive maintenance downtime decreased from 20 hours/year to 10 hours/year and its cost decreased from 8×10^6 L.E./year to 3.97×10^6 L.E./year. This means that proposed RCM saved 6.19×10^6 L.E in total maintenance cost. System availability and reliability are improved as well. Availability increased from 57.1% to 90.74% and reliability increased from 99.73% to 99.88%. New points of research can be investigated in future through various directions, e.g. the integration of the proposed methodology with other maintenance techniques, the proposed policy can be applied to other large industrial processes.

10. Acknowledgment

The authors would like to express the sincere appreciation for the staff members of the Fayoum Sugar Works Company, Kasr El Basel, Itsa, El Fayoum, Egypt, for their support during carrying out this work.

References

[1] Dixey M., "Putting Reliability at the Center of Maintenance", Professional Engineering, Vol. 6, pp. 23-25, 1993.

[2] Islam H. Afefy, "Reliability-Centered Maintenance Methodology and Application: A Case Study", Engineering, Vol.2, pp. 863-873, 2010.

[3] David L., et al. "Software solution design for application of reliability centered maintenance in preventive maintenance plan". Electric power engineering, 18 th international scientific conferences on. IEEE, pp. 1-4, 2017.

[4] Diego, et al. "A new approach for reliability-centered maintenance programs in electric power distribution systems based on a multi objective genetic algorithm". Electric power systems research, Vol. 137, pp. 41-50, 2016.

[5] N. Ahmad and M. A. U. Karim, "A framework for application of reliability centered maintenance in the lead oxide production system", Applied mechanics and materials, Vol. 860, pp. 123-128, 2017.

[6] Tamer M. El-Dogdog, Ahmed M. El-Assal, Islam H. , Ahmed A. El-Betar, "Implementation of FMECA and Fishbone Techniques in Reliability Centered Maintenance Planning", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, pp.18801-18811, 2016.

[7] International atomic energy agency" Application of reliability centered maintenance to optimize operation and maintenance in nuclear power plants" , 2007.

[8] Majid M. A., Muhammad M., and Yem N.I., "RCM Analysis of Process Equipment: A Case Study on Heat Exchangers" Journal of Applied Sciences, Vol. 11, pp. 2058-2062, 2011.

[9] Islam H. Afefy, "Maintenance Planning Based on Computer-Aided Preventive Maintenance Policy", International Multi Conference of Engineers and Computer Scientists, Vol. 2, pp.1378-1383, Hong Kong, 2012.

[10] Selvik J.T., Aven T., "A Framework for Reliability and Risk Centered Maintenance", Reliability Engineering and System Safety, Vol. 96, pp. 324-331, 2011.

[11] Fonseca D.J., Knapp G.M., "An Expert System for Reliability Centered Maintenance in the Chemical Industry", Expert Systems with Applications, Vol. 19, pp. 45-57, 2000.

[12] Dacheng Li., Jinji G., "Study and Application of Reliability-Centered Maintenance Considering Radical Maintenance", Journal of Loss Prevention in the Process Industries, Vol. 23, pp. 622-629, 2010.

[13] Mark C. E., Ogaji S.O.T., Probert S.D., "Strategic Maintenance Management in Nigerian Industries", Applied Energy, Vol. 83, pp. 211-227, 2006.

[14] .Zoran P., Zlatan C., Branko R., Leon Š., Vladan G., "Implementation of the RCM Methodology on the Example of City Waterworks", VIII International Conference "Heavy Machinery-HM", Vol.8, pp. , 2014.

[15] Dawane A. P., Sedani D. C. M, "Study and Investigations of RCM Methodology in Manufacturing Industry to Minimize Breakdown Maintenance", International Journal of Innovations in Engineering and Science, Vol. 1, pp.25-33, 2016.

[16] Yang Tang, et al. "A framework for identification of maintenance significant items in reliability centered maintenance". Energy, Vol. 118, pp. 1295-1303, 2017.

[17] Samuel et al. "A Reliability-Centered Maintenance Study for an Individual Section-Forming Machine", *Machine Journal*, Vol.6, 2018.

[18] Abdulrohim S. A., Salih O. D., and Raouf A., "RCM Concepts and Application: A Case Study", International Journal of Industrial Engineering, Vol. 7, pp. 123-132, 2000.

[19] Mkandawire B.O., N.M. Ijumba and A.K. Saha, "Transformer Risk Modelling by Stochastic Augmentation of Reliability-centered Maintenance," Journal of Electric Power Systems Research (Elsevier/Science Direct), Vol. 119, pp. 471-477, 2015.

[20] Islam H. A., Mohamed F. A., Ragab K. A., Eman K. A., "A Comprehensive Model of Reliability, Availability, and Maintainability (RAM) for Industrial Systems Evaluations", Jordan Journal of Mechanical and Industrial Engineering, Vol. 12, pp. 59-67, 2018.

[21] Islam H. A., Mohamed F. A., Ragab K. A., Eman K. A., "An Efficient Maintenance Plan Using Proposed Framework of RCM Made Simple Approach", Journal of Industrial Engineering & Management System, Vol. 18, pp. 222-233, 2019.

[22] 41. Anthony M. Smith and Glenn R. Hineheliffe, "RCM: Gateway to World Class Maintenance", Elsevier Inc., USA, 2004.

[23] Marvin R., "Reliability Centered Maintenance", Reliability Engineering and System Safety, Vol. 60, pp.121-132, 1998.

[24] Islam H. Afefy,"Hazard Analysis and Risk Assessments for Industrial Processes Using FMEA &Bow-Tie Methodologies" Journal of Industrial Engineering & Management Systems (IEMS), Vol.14, 2015.

[25] Dhillon B.S., "Engineering Maintainability: how to design for reliability and easy maintenance", Gulf Professional Publishing, 1999.

[26] Woohyun K., Jaechul Y., and Suneung A., "Determining the Periodic Maintenance Interval for Guaranteeing the Availability of A System with a Linearly Increasing Hazard Rate", International Journal of Industrial Engineering, Vol. 16, pp.126-134, 2009.

[27] Ronald T. A., Lewis N., "Reliability Centered Maintenance: Management and Engineering Methods", Elsevier science publishers ltd, New York, 1990.

[28] Ronald T. A., Lewis N., "Reliability Centered Maintenance: Management and Engineering Methods", Elsevier science publishers ltd, New York, pp.359, 1990.

[29] Yssaad B and Abene A., "Rational reliability Centered Maintenance optimization of power distribution system", Journal of Electrical Power and Energy System , Elsevier, Vol. 73, pp. 350-360, 2015.

[30] Jasiulewicz-Kaczmarek M., "Practical Aaspects of the Application of RCM to Select Optimal Maintenance Policy of the Production Line", Journal Engineering Safety and Reliability: Methodology and Applications, Taylor & Francis Group, London, pp. 1187- 1195, 2015.

[31] Guofa Li, Yi Li, Xinge Zhang *, Chao Hou, Jialong He, Binbin Xu and Jinghao Chen, " Development of a Preventive Maintenance Strategy for an Automatic Production Line Based on Group Maintenance Method" Journal of Applied Science, Vol. 8, pp. 1-15, 2018.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 191 - 196

Jordan Journal of Mechanical and Industrial Engineering

Unconfined Laminar Nanofluid Flow and Heat Transfer Around a Square Cylinder with an Angle of Incidence

Rafik Bouakkaz[a*], Yacine Khelili[a], Faouzi Salhi[b]

[a] Military academy of Cherchell, Tipaza, Algeria.
[b] Département de Génie Mécanique, université Mouloud Mammeri Tizi ouzou

Received 5 November. 2018

Abstract

A finite-volume method simulation is used to investigate two-dimensional unsteady flow of nanofluids and heat transfer characteristics past a square cylinder inclined with respect to the main flow in the laminar regime. The computations are carried out of nanoparticle volume fractions varying from $0 \leq \ \leq 5\%$ for an inclination angle in the range $00 \leq \delta \leq 450$ at a Reynolds number of 100. The variation of stream line and isotherm patterns are presented for the above range of conditions. Also, it is noticed that the addition of nanoparticles enhances the heat transfer. Hence, the local Nusselt number is found to increase with increasing value of the concentration of nanoparticles for the fixed value of the inclination angle.

Keywords: Copper nanoparticles, heat transfer, square cylinder, inclination angle;

Nomenclature

L	side of the square cylinder	m
$\overline{Nu}$	average Nusselt number	
Nu_{local}	local Nusselt number	
Re	Reynolds number	
Pr	Prandtl number	
T_∞	free-surface temperature	k
U_∞	free-stream velocity	ms^{-1}
u	Stream-wise velocity	ms^{-1}
v	Cross-stream velocity	ms^{-1}

Greek symbols

ϕ	nanoparticle volume fractions	
δ	angle of incidence	^{0}
ϴ	non-dimensional temperature	
ν	kinematic viscosity	$m^2\ s^{-1}$

1. Introduction

In recent years, studies about free convective fluid flow and heat transfer around square cylinder has been a subject of enormous attention for scientists due to its high applicability in many environmental situations or engineering developments such as, textile, thermal insulation, buildings, electronic equipments and chemical processing industries, etc. Despite the configuration being simple, the flow around a square cylinder involves a complex transport phenomenon because of a lot of factors, such as the effect of angle of incidence on the creation of lift force, evolution of streamline and temperature field, etc. This work is concerned with the characteristics of flow of Cu–water nanofluid and heat transfer past a square cylinder with varying values of angle of incidence in the unsteady laminar regime.

In literature, enough information is available presenting convective heat transfer characteristics around a square cylinder [1-5]. The effect of blockage ration on the fluid flow from a square cylinder at angle incidence of 0–45 and Re = 45–200 has been investigated by Sohankar et al. [6]. Subsequently, Sohankar et al. [7] studied numerically the characteristics of unsteady flow for the same configuration for Reynolds number (Re = 150–500) and a blockage ratio of 5.6%. Their results show a two-dimensional laminar shedding flow at Re= 150, and at Re = 200 the flow becomes three-dimensional. A well-organized study was also published by Robichaux et al [8]. In that paper, the onset of three-dimensionality in the wake of a square cylinder was analyzed by Floquet stability to show the different modes of 3-D instabilities. They establish that the 3-D disturbance first becomes unstable at a Reynolds number of about 161. Further, a parametric study was carried out by [9] with various values of angle of incidence and Re. They established critical Reynolds number for periodic vortex shedding at each angle of incidence. Recently, a direct numerical simulation was applied by Rastan [10] et al. to investigate three-dimensional unsteady flow characteristics around a finite wall-mounted square cylinder with an aspect ratio of 7 at a Reynolds number

* Corresponding author e-mail: r.bouakkaz@gmail.com

(Re) of 40-250. They found that the wake flow becomes turbulent at Re > 200. In past studies, the fluids used have a low value of thermal conductivity, which limits heat transfer. For this reason, there are several methods to improve the heat transfer characteristics, which consist in adding high conducting solid particles in the base fluid. The resulting fluid is called ‘‘nanofluid’’ [11–12]. Furthermore, [13] have studied the momentum and forced convection heat transfer for a laminar and steady free stream flow of nanofluids past an isolated square. Various nanofluids consisting of Al_2O_3 and CuO with base fluids of water and a 60:40 (by mass) ethylene glycol and water mixture were selected to estimate their superiority over conventional fluids. They established that for any given particle diameter, there is an optimum value of particle concentration that results in the highest heat transfer coefficient. The fluid flow and heat transfer around a square cylinder utilizing Al_2O_3–H_2O nanofluid over low Reynolds numbers varied within the range of 1 to 40 and the volume fraction of nanoparticles (ϕ) is varied within the range of $0 < \phi < 0.05$ was also investigated by [14]. They found that increasing the nanoparticles volume fractions augments the drag coefficient. Moreover, Pressure coefficient increases by increasing the solid volume fraction for sides where pressure gradient is inverse, but for sides where the pressure gradient is favourable the pressure coefficient decreases. Recently, Rajendra et al. [15] have carried out a meticulous study on the forced convective heat transfer from an unconfined heated square cylinder utilizing nanofluids with multiphase modelling approach for different Reynolds number (10-40) and volume fractions of Al_2O_3 particles (0-5%) in water. The results indicated that the effect of Nano layer thickness and nanoparticle diameter on the overall heat transfer rate can be studied. So, the aim of the current study is to investigate numerically by using a finite volume method based on SIMPLE algorithm the laminar flow of nanofluid and heat transfer characteristics for an inclination angle in the range $0^0 \le \delta \le 45^0$ under the particle volumetric concentrations ranging from 0% to 5% at Reynolds numbers (Re =100).

2. Governing equations and boundary conditions

The governing partial differential equations here are the Navier-Stokes and energy equations in two dimensions the incompressible nanofluid flow around a square cylinder are given blow:

$$\frac{\partial U}{\partial X}+\frac{\partial V}{\partial Y}=0 \tag{1}$$

$$\frac{\partial U}{\partial \tau}+U\frac{\partial U}{\partial X}+V\frac{\partial U}{\partial Y}=-\frac{\partial P}{\partial X}+\frac{1}{\mathrm{Re}}\frac{\nu_{nf}}{\nu_f}\left(\frac{\partial^2 U}{\partial X^2}+\frac{\partial^2 U}{\partial Y^2}\right) \tag{2}$$

$$\frac{\partial V}{\partial \tau}+U\frac{\partial V}{\partial X}+V\frac{\partial V}{\partial Y}=-\frac{\partial P}{\partial Y}+\frac{1}{\mathrm{Re}}\frac{\nu_{nf}}{\nu_f}\left(\frac{\partial^2 V}{\partial X^2}+\frac{\partial^2 V}{\partial Y^2}\right) \tag{3}$$

$$\frac{\partial \theta}{\partial \tau}+U\frac{\partial \theta}{\partial X}+V\frac{\partial \theta}{\partial Y}=\frac{1}{\mathrm{Re}\,\mathrm{Pr}}\frac{\alpha_{nf}}{\alpha_f}\left(\frac{\partial^2 \theta}{\partial X^2}+\frac{\partial^2 \theta}{\partial Y^2}\right) \tag{4}$$

Where:

$$U=\frac{u}{U_\infty},V=\frac{v}{U_\infty},X=\frac{x}{L},Y=\frac{y}{L},\tau=\frac{tU_\infty}{L},$$

$$P=\frac{p}{\rho U_\infty^2},\mathrm{Pr}=\frac{\nu_f}{\alpha_f},\mathrm{Re}=\frac{U_\infty L}{\nu_f},\theta=\frac{T-T_\infty}{T_w-T_\infty}$$

Where U and V are the velocity components along X and Y axes, T denotes the temperature, P is the pressure, ρ is the density, μ the dynamic viscosity. The subscript *nf* stands for nanofluid, the subscript *f* stands for base fluid and the subscript s stands for solid nanoparticles. The thermophysical properties taken from [11], for the base fluid and copper oxide (at 300 K) are shown in Table 1.

Table 1. Thermo-physical properties of the base fluid and the Cu nanoparticles

Property	Water	Copper
$C_p\left(J Kg^{-1}K^{-1}\right)$	4179	385
$\rho\left(Kgm^{-3}\right)$	997.1	8.933
$k\left(Wm^{-1}K^{-1}\right)$	0.613	401

3. Boundary conditions

The dimensionless boundary conditions for the flow across a square cylinder surrounded by Cu-water nanofluid can be written as (Figure 1): The left-hand section, the Dirichlet-type boundary condition for the Cartesian velocity components is assumed.

$$U=1,\ V=0 \text{ and } \theta=0 \tag{5}$$

The right-hand, the diffusion flux in the direction normal to the exit surface is zero for all variables,

$$\partial U/\partial X=\partial V/\partial X=\partial \theta/\partial X=0 \tag{6}$$

On the straight horizontal segments (slip boundary), a zero normal velocity and a zero normal gradient of all variables are prescribed:

$$\partial U/\partial Y=\partial V/\partial Y=\partial \theta/\partial Y=0 \tag{7}$$

Finally, the dimensionless peripheral or tangential velocity is prescribed on the surface of the cylinder along with a no-slip boundary condition

$$U=0,\ V=0 \text{ and } \theta=1 \tag{8}$$

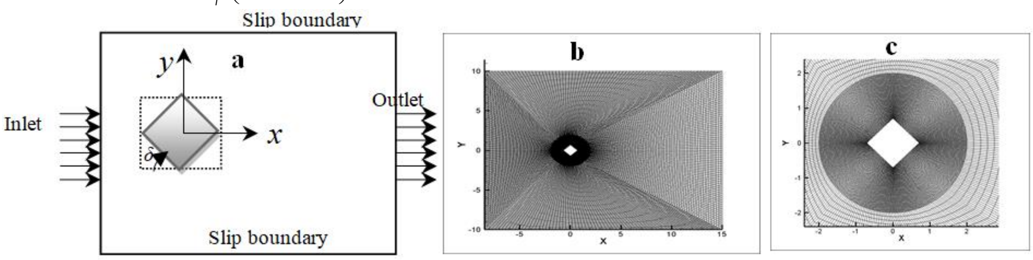


Figure 1. (a) Schematic of the unconfined flow and heat transfer around a square cylinder. (b) Grid structure. (c) Close-up view in the vicinity of the square cylinder.

The effective density, the thermal diffusivity, the heat capacitance, the effective dynamic viscosity and the effective thermal conductivity of the nanofluid are calculated using the following expressions:

$$\rho_{nf} = (1-\phi)\rho_{bf} + \phi\rho_s \quad (9)$$

$$(\rho C_p)_{nf} = (1-\phi)(\rho C_p)_{bf} + \phi(\rho C_p)_s \quad (10)$$

$$\alpha_{nf} = k_{nf} / (\rho C_p)_{nf} \quad (11)$$

$$\mu_{nf} = \frac{\mu_{bf}}{(1-\phi)^{2.5}} \quad (12)$$

$$k_{nf} = k_{bf}\left[\frac{(k_s + 2k_{bf}) - 2\phi(k_{bf} - k_s)}{(k_s + 2k_{bf}) + \phi(k_{bf} - k_s)}\right] \quad (13)$$

Where φ is the solid volume fraction and given as:

$$\phi = \frac{\text{Volume of nanopartides}}{\text{Total volume of solution}} \quad (14)$$

4. Force Coefficient

The relevant parameters computed from the velocity and pressure fields are the drag coefficient, which represent dimensionless expressions of the forces that the fluid produces on the circular cylinder. Defined as:

$$C_D = \frac{D}{\rho U_\infty^2 L} \quad (15)$$

where *D* is the drag force.

5. Results and discussion

5.1. Numerical details

The unsteady, laminar, segregated solver was employed here to solve the incompressible flow on the collocated grid arrangement. Semi implicit method for the pressure linked equations (SIMPLE) was used to solve Navier-Stokes and energy equations for above noted boundary conditions. Second order upwind scheme is used to discretize the convective terms of momentum equations, whereas the diffusive terms are discretized by central difference method. A convergence criterion of 10^{-8} is used for continuity, and x-y components of momentum equations. While, for energy equation the criteria of convergence were 10^{-10}. The numerical resolution was determined by a grid refinement study to ensure grid independency. The mesh used for all the two-dimensional computations consisted of 39200 quadrilateral cells and 39480 nodes. The grid is divided into two separate zones, and uniform as well as no uniform grid distributions are employed. The grid distribution was made uniform with a constant cell size, of 0.02, inside a region around the cylinder that extended 4 units to capture wake–wall interactions adequately. Then, a grid of much bigger size is clustered around the cylinder over the distance indicated above.

5.2. Comparison with other results

Table 2 compares the mean Nusselt number and drag coefficients computed here with results obtained from the literature. We have noted that the Nusselt number and drag coefficient values are in superb agreement with numerical data reported by other researchers.

Table 2. Comparison of Nu number and drag coefficient computed in the present study with literature data (Pr = 0.7)

	Re = 20		Re = 40		Re = 100	
	Nu	CD	Nu	CD	Nu	CD
Present study	2.07	2.43	2.71	1.81	4.07	1.51
Paliwali et al. [16]	2.07	-	2.71	1.98	-	-
Sharma and Eswaran [3]	2.05	2.35	2.71	1.75	-	-
Etminan-Farooji [13]	2.07	2.43	2.72	1.83	-	-
Sahu et al. [5]	-	-	-	-	4.03	1.49
Prasenjit [17]	-	-	-	-	3.84	1.53

5.3. Flow patter

Within the full range of possible incidences, this study indicated two different flow patternsnamely, Main separation, and Vortex creation, as illustrated in figure 2. These flow patterns are summarized below.

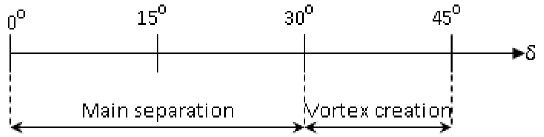


Figure 2. Flow patterns past a square cylinder with an angle of incidence at Re=100.

5.4. Main separation

Figure 3 shows the case of instantaneous streamlines past a square cylinder at δ =0°. In the Main separation pattern, main vortices of opposite sign roll up in an alternating manner as mentioned in Perry et al. [16] and instant ‘alleyways’ of fluid enter the cavity as well as an alleyway flow indicated by the arrows between. However, once the vortex-shedding process begins, this so-called ‘closed’ cavity becomes open, and instantaneous ‘alleyways’ of fluid are created which penetrate the cavity.

5.5. Vortex creation

The vortex merging pattern is characterized by creation and merging of two small vortices which are originally produced by separation and subsequent reattachment of the alleyway flow at two neigh boring edges of the square cylinder. Clearly, in figure 4, the vortices A and B are generated due to separation and reattachment of the alleyway flow. The two vortices are merged into a big one, C Figure 4b. The vortex C is shed into the wake H in Figure 4c, and small vortices D and E are produced by the opposite alleyway flow, Figure 4c. Vortices D and E are merged into F, Figure 4d. In most cases of vortex merging pattern, an alleyway flow exhibits the characteristic feature of the Vortex creation pattern, generation and subsequent merging of two small vortices, in an alternating manner with a downwash or an up-wash flow in the alleyway like in figure 4.

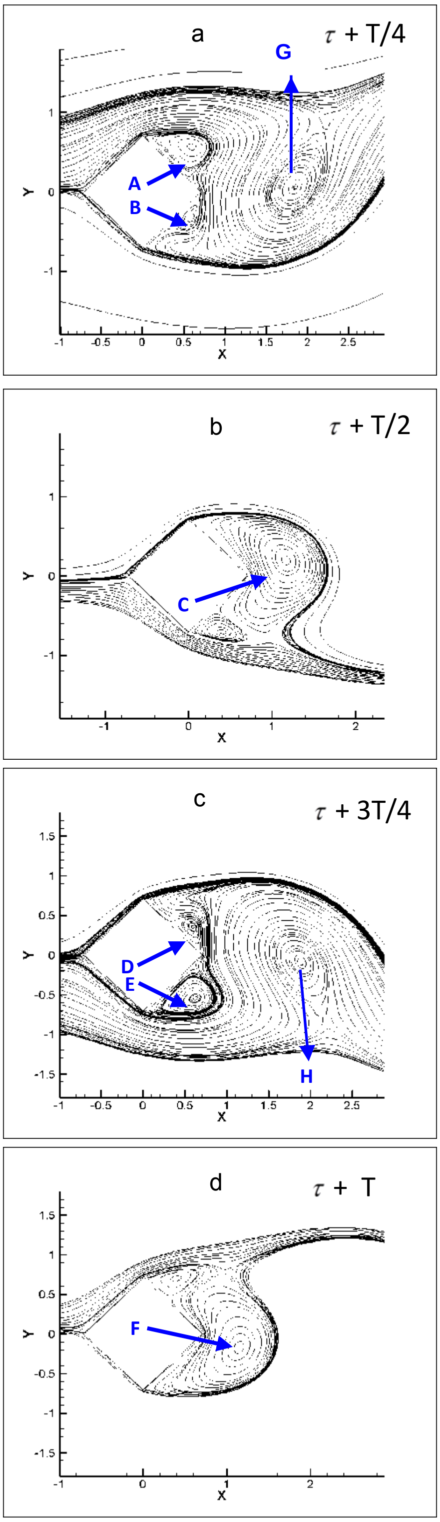


Figure 4. Instantaneous streamlines (base fluid) for Vortex creation, where T is one period of vortex shedding**.**

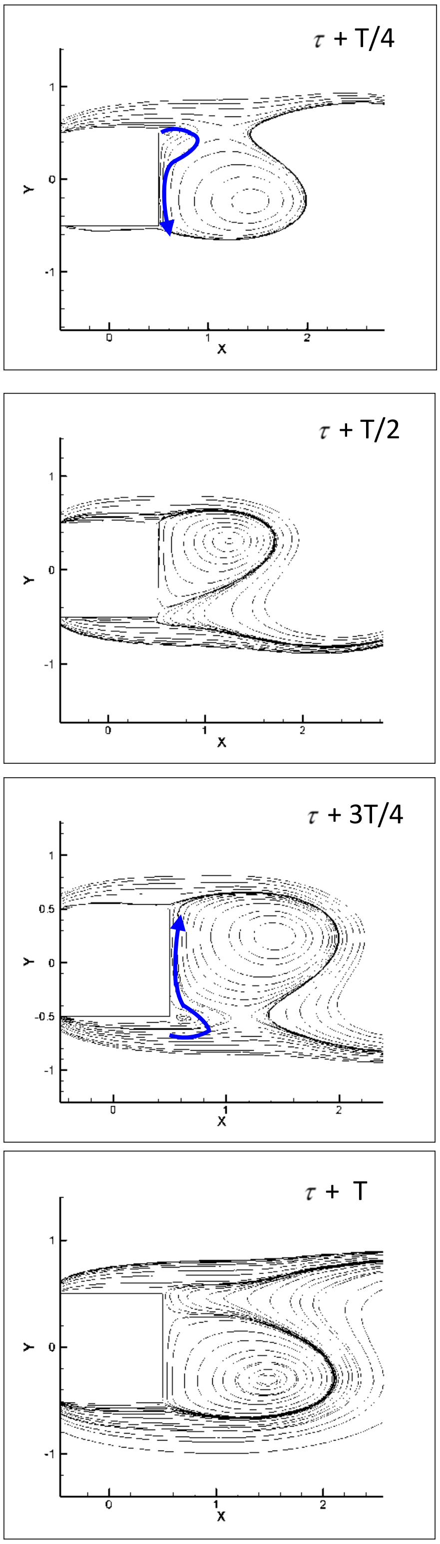


Figure 3. Color online Instantaneous streamlines (base fluid) for Main separation pattern, where T is one period of vortex shedding.

5.6. Isotherm patterns

The isotherms profiles around the cylinder for Reynolds number of 100 at δ=15^0 are compared between base fluid and nanofluid (ϕ =0.05) in figure 5. Clearly, the temperature distribution contours for base fluid are overlaid with that for nanofluid. This can be explained as the addition of solid particles to the base fluid increases the Reynolds number of nanofluid. Hence, a higher capacity of transferring the heat from the cylinder. It is obvious from Figure 4 that the isotherms have maximum density close to the front surface of the cylinder (AB); this indicates high values of the local Nusselt number near the front stagnation point on the front surface as compared to other points on the cylinder surface.

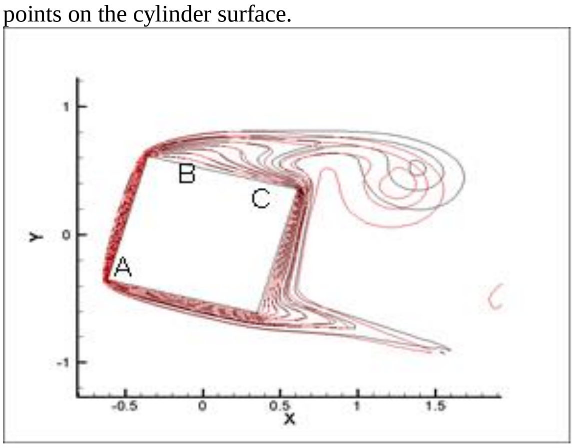


Figure 5. Temperature contours for the flow around the square cylinder (red line refers to base fluid and black line refers to nanofluid with solid volume fraction 0.05) at δ=15^0

5.7. Local Nusselt number

The local Nusselt number is calculated to evaluate the warmth transmission behaviour around the cylinder and given as:

$$Nu_{local} = \frac{h_{local}.D}{K_f} = -\left(\frac{k_{nf}}{k_{bf}}\right)\frac{\partial\theta}{\partial n} \quad (16)$$

Figure 6 shows the variation of local Nusselt number (Nu) on the surface of the square cylinder with increase for various volume fraction ϕ at δ=15^0. When the solid concentration increases, the thermal conductivity improves and consequently the local Nusselt number. Additionally, the thermal boundary layer is decreased by any increase in solid volume fraction. Therefore, the local Nusselt number is enhanced by any increasing in solid volume fraction (ϕ). Further, these plots also show that the Nusselt number along the left half of the face (AB) of the cylinder increases, it has a maximum at the corner of the square cylinder (B). On the other hand, a value for the inclination angle of δ=45^0 is found for which the local Nusselt number is highest along the left half of the face (BC) as seen in Figure 7.

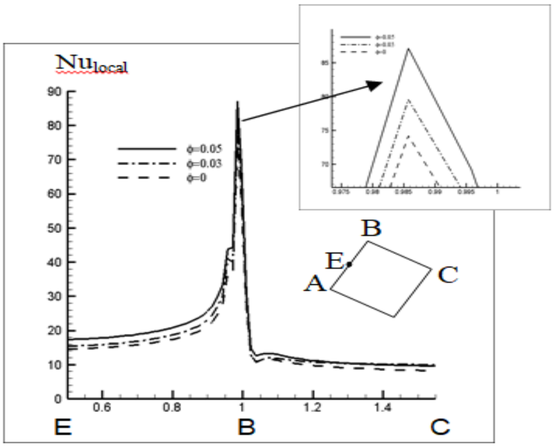


Figure 6. Local Nusselt number variation for δ=15^0 at various solid volume fractions.

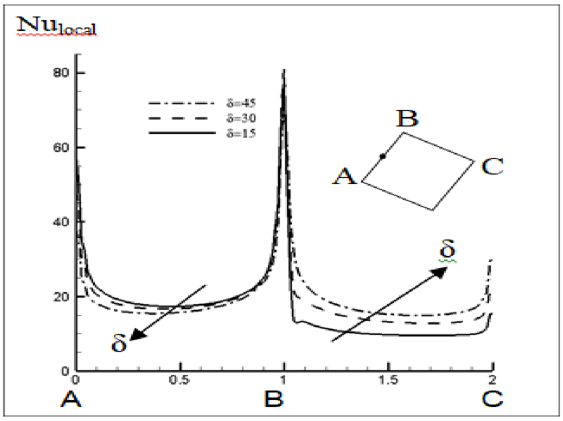


Figure 7. Local Nusselt number variation for ϕ =0.05 at various angle of incidence.

5.8. Averaged Nusselt number

Surface averaged Nusselt number of fully developed thermal boundary layer is defined as:

$$\overline{\mathrm{Nu}} = \frac{1}{S}\int_s Nuds \quad (17)$$

The average Nusselt number variation is presented in figure 8 for the solid volume fraction varying from 0 to 0.05 at different angle of incidence. This figure indicates that the averaged Nusselt number increases with increasing solid volume fraction of nanoparticles (ϕ) of the fixed value of the angle of incidence. This can be explained as when ϕ number increases the inertia of flow increases thus increasing the heat transfer. Also, the average Nusselt number is highest at δ =45°.

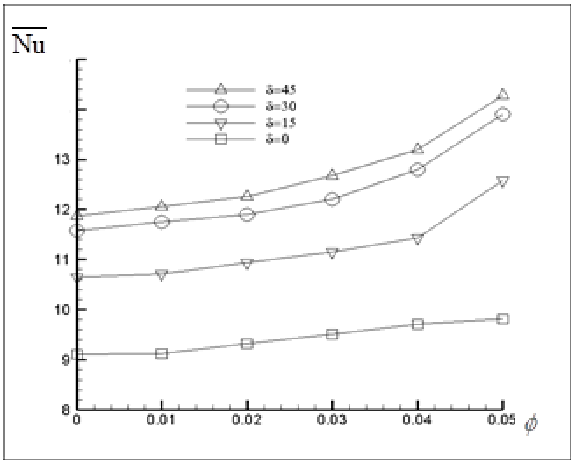


Figure 8. Variation of average Nusselt number at various solid volume fractions for varying values of angle of incidence.

6. Conclusions

The present study focuses on the unconfined laminar flow of nanofluid and heat transfer characteristics around a square cylinder under the influence of various angle of incidence in the unsteady regime. The illustrative streamline shows two different flow patterns, namely, Main separation, and Vortex creation. Further, the temperature distribution contours for base fluid are overlaid with that for nanofluid. This can be explained as the addition of solid particles to the base fluid increases the Reynolds number of nanofluid. Hence, a higher capacity of transferring the heat from the cylinder. On the other hand, the isotherms have maximum density close to the front surface of the cylinder; this indicates high values of the local Nusselt number near the front stagnation point on the front surface as compared to other points on the cylinder surface. Finally, it was showed that the average Nusselt numbers were enhanced by adding nanoparticles to base fluid for various angle of incidence. Moreover, the average Nusselt number is highest at δ =45°.

References

[1] M. Sarioglu, Y. E. Akansu, and T. Yavuz, "Control of flow around square cylinders at incidence by using a rod", AIAA J (2005) 43, 1419

[2] T. Igarashi, "Characteristics of the flow around a square prism," Bull. JSME 1984, 27, 1858

[3] A. Sharma, V. Eswaran, "Heat and fluid flow across a square cylinder in the two-dimensional laminar flow regime", Numer. Heat Transfer A (2004) 45, 3, 247– 269.

[4] A.K. Dhiman, R.P. Chhabra, V. Eswaran, "Flow and heat transfer across a confined square cylinder in the steady flow regime: effect of Peclet number", Int. J. Heat Mass Transfer (2006) Part A, 49: 717–731.

[5] A.K. Sahu, R.P. Chhabra, V. Eswaran, "Effects of Reynolds and Prandtl numbers on heat transfer from a square cylinder in the unsteady flow regime", Int. J. Heat Mass Transfer (2009) 52(3), 839–850.

[6] A. Sohankar, C. Norberg, L. Davidson, "Low-Reynolds-number flow around a square cylinder at incidence: study of blockage, onset of vortex shedding and outlet boundary condition", Int. J. Numer. Methods Fluids (1998) 26, 39

[7] A. Sohankar, C. Norberg, L. Davidson, " Numerical simulation of unsteady fow around a square two-dimensional cylinder",Proceedings of the 12th Australasian Fluid Mechanics Conference, Sydney,Australia (1995) pp. 517–520.

[8] J. Robichaux, S. Balachandar, and S. P. Vanka, "Three-dimensional
Floquet instability of the wake of square cylinder," Phys. Fluids (1999) 11, 560

[9] Dong-Hyeog Yoon, Kyung-Soo Yang,a and Choon-Bum Choi , "Flow past a square cylinder with an angle of incidence", PHYSICS OF FLUIDS (2010) 22, 043603

[10] M. R. Rastan, A. Sohankar, and Md. Mahbub Alam , "Low-Reynolds-number flow around a wall-mounted square cylinder: Flow structures and onset of vortex shedding ",Physics of Fluids (2017) 29, 103601.

[11] R.Bouakkaz et al., "Unconfined laminar nanofluid flow and heat transfer around a rotating circular cylinder in the steady regime ", archives of thermodynamics (2017) 38, No. 2, 3–20.

[12] H.C. Brinkman, "The viscosity of concentrated suspensions and solutions", The journal of Chemical Physics 20 (1952) 571–581.

[13] V.E. Farooji., E.E.Bajestan, H.Niazmand, S.Wongwises, "Unconfined laminar nanofluid flow and heat transfer around a square cylinder", Int. J. Heat Mass Tran. 2012,55, 5-6, 1475–1485.

[14] M.S. Valipour, R Masoodi, S. Rashidi, M. Bovand, M. Mirhosseini. "A numerical study on convection around a square cylinder using AL2O3-H2O nanofluid", Therm. Sci. (2014) 18, 4, 1305–1314.

[15] S. Rajpoot Rajendra, S. Dhinakaran, "Study of heat transfer from a square cylinder utilizing nanofluids with multiphase modeling approach Materials Today", Proceedings (2017) 4 ,10069–10073

[16] A. E. Perry, M. S. Chong, and T. T. Lim, "The vortex-shedding process behind two-dimensional bluff bodies," J. Fluid Mech (1982.)116, 77

[17] Dey Prasenjit and Das Ajoy kumar , "Analysis of Fluid Flow and Heat Transfer Characteristics Over a Square Cylinder: Effect of Corner Radius and Nanofluid Volume Fraction ," Arab J Sci Eng (2015) DOI 10.1007/s13369-016-2276-2

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 197 - 205

Jordan Journal of Mechanical and Industrial Engineering

Improvement of Operator Position Prediction in Teleoperation Systems with Time Delay: Simulation and Experimental Studies on Phantom Omni Devices

Behnam Yazdankhoo[a], Moein Nikpour[a], Borhan Beigzadeh[*a], Ali Meghdari[b]

[a] Biomechatronics and Cognitive Engineering Research Lab, School of Mechanical Engineering, Iran University of Science and Technology, Tehran, Iran.

[b] Social & Cognitive Robotics Laboratory, Center of Excellence in Design, Robotics and Automation (CEDRA), Sharif University of Technology, Tehran, Iran.

Received 7 February. 2019

Abstract

An online operator position prediction approach based on artificial neural network for teleoperation systems is proposed in this paper, which predicts future position of operator's hand based on current available data. The neural network gathers inputs for some time at the beginning of the operation, then is trained, and is finally exploited through the rest of the operation. Superiority of the proposed approach can be investigated from two aspects. Firstly, no limiting assumption is required in this approach in contrast with the proposed methods in the literature. Secondly, unknown operator intention can be dealt with in real time if it is not too sudden and unpredictable. Two different scenarios are considered in this paper: in scenario I a simple harmonic motion is both applied and predicted, whereas in scenario II not only the applied motion is more complicated, but also it is different from the motion which is supposed to be predicted. The results of the second scenario show that the designed architecture can be readily extended to a variety sort of situations in which little information exists regarding operator intention. Computer simulations and experiments using Phantom Omni haptic devices further validate the feasibility and performance of the proposed approach, i.e. master and slave robots can move simultaneously with no specific a priori knowledge about operator intention, despite large time delay in the system.

Keywords: operator position prediction, neural network, online prediction, teleoperation system, time delay;

1. Introduction

Novel approaches based on predictive control methods have been recently proposed in order to circumvent the time delay in the teleoperation system [1]. From a broader perspective, predictive control methods can be considered as a part of environment-, operator-, and task-adapted (EOT-adapted) controllers [2]. Here we stick to a simpler categorization, and we consider two main predictive approaches: predicting the environment model and predicting the operator motion.

To date, most of the works are concerned with the first approach. This method, which is called *model-mediated teleoperation* [3], has been widely used in designing teleoperation systems in recent years. Sample works regarding this approach can be found in [4, 5]. On the other hand, predicting the operator motion (Figure 1) can also play an important role in improving the quality of teleoperation systems, as Jarrassé et al. [6] proved the influence of human motion prediction on increasing transparency of teleoperation systems. This is, however, a more complicated task since the human operator might show unpredictable motions [7]. The works which have been done regarding this issue can be categorized into two main approaches: offline and online.

Offline motion prediction is fundamentally based on minimum principles [8], with the most widely used principle being the minimum jerk principle [9], which states that between two predefined points the human hand follows the trajectory whose jerk is the minimum amount. The basic concept of online motion prediction is, however, predicting the human behavior during the operation based on the online data generated by him/her. Linear prediction [10], polynomial or spline predictors [11], double exponential smoothing (DES) [12, 13], autoregressive (AR) method [14] and Markov-chain-based Kalman filter [15] are some samples of online methods. Recent methods based on state observer have also been proposed [16]. Unlike offline methods which are basically derived from physical principles, the common core of online methods is numerical data. This advantage makes online approaches more adaptive to various applications, but at the same time rather more intricate to design.

Each of the above-mentioned methods for online prediction has also some disadvantages. For instance, DES

* Corresponding author e-mail: b_beigzadeh@iust.ac.ir

can be utilized when the data shows a trend, but not seasonality. Triple exponential smoothing (TES) shall be used in this case instead [17]. As another example, AR method models the future state as a linear combination of the current and previous states, which may fail to correctly predict the future states for more complex motions where this dependency is nonlinear. To mitigate these restrictions, a prediction method based on artificial neural network (NN) is proposed in this paper.

Control of teleoperation systems based on NN has recently gained researchers' attentions [18, 19]. Due to the universal approximation property of NNs, any kind of linear or nonlinear relationship between the inputs and outputs can be modeled with a proper NN [20]. It might be interesting to say that it has been contended that even a quadratic regression model is computationally less efficient in comparison with a NN [21]. Also, from prediction task perspective, NNs do not require any assumption regarding the trend or seasonality of the motion beforehand. These properties have made NN a suitable choice for predicting the future states of a system. Nicolau et al., for instance, exploited a NN for prediction of the roll motion of a ship [22]. However, to the best of authors' knowledge, no previous work has utilized NN for predicting operator motion in teleoperation systems.

In this work, we make use of an artificial NN to predict the position of the operator's hand online. The NN is trained for a short time at the beginning of the operation, and then can predict the operator movements thereafter. Two main contributions can be mentioned for the proposed method: Firstly, using the proposed approach, the teleoperation system is able to predict not only the expected operator motions, but also the unexpected ones if they are not unpredictably sudden. This property enables human operator to carry out more than one single task at once, thus there is no need to reinitialize the operation if the task is decided to be altered slightly. Secondly, there is no restricting assumption in the proposed method from applicability point of view, i.e. no specific detailed information about the operator's intended motion is needed beforehand. This feature makes the applicability of the system much wider; from surgery and suturing in medical robotics, which is believed to have major research impact in the next decade [23], to ordinary pick and place tasks in industry. It should also be asserted that the defined motion is set just as a general case; the NN system can predict any other type of motion.

2. System architecture

In this work, both free motion (no environment) case and constant known time delay is assumed, shown in Figure 2. The structure of the NN is also illustrated in Figure 3, which is a feedforward network with three layers. The input layer is comprised of the current and previous positions of the master robot. If there are m neurons in the input layer and s is assumed to be the sample time, the input to the NN will be $\mathbf{X}(t) := [x(t) \quad x(t-s) \quad x(t-2s) \quad \cdots \quad x(t-(m-1)s)]^T$. The hyperbolic tangent function is considered as the activation function for the hidden layer.

The output layer always includes one neuron which is the predicted future position of the master robot, namely $\hat{x}(t+\tau)$, where τ denotes the time delay and the accent ^ refers to the predicted version of the respective parameter. The activation function for this neuron is the identity function, i.e. $f(x) = x$.

The NN training is carried out based on the backpropagation gradient descent method. According to this method, the weight update in each iteration or epoch follows the relation (2.1),

$$\Delta w_{ij}^{(n)} = -\eta \frac{\partial E}{\partial w_{ij}^{(n)}} \tag{2.1}$$

where $w_{ij}^{(n)}$ represents the connection weight between the ith neuron in the *(n-1)*th layer and the jth neuron in the nth layer, $n = 2, 3, \ldots, N$ with N denoting the total number of layers, and Δ shows the change in the respective weight. Also η is the learning rate, which is assumed to be constant, and E is the objective function to be minimized which is $E = \frac{1}{P}\sum_{p=1}^{P}(o_p - d_p)^2$, where o and d are the actual and desired outputs of the NN, respectively, and P represents the total number of presented training sets to the network. The update of the biases follows the same rule by simply replacing $w_{ij}^{(n)}$ with $b_j^{(n)}$ in the relation (2.1). More detailed explanation about the concepts and formulation of the gradient descent method can be found in [24].

Generally, there are two approaches to training a NN: batch-mode and incremental-mode [24]. In this work, both methods will be used for training the NN.

As mentioned before, the proposed teleoperation system in this paper is designed based on online prediction scheme. The system performs according to the following procedure:

In the first phase, the operator moves the master robot reciprocally for some time. Knowing a rough estimation of the period of the motion beforehand, which is denoted by T_h, we record the operator's motion for the first T_1 seconds, where T_1 is slightly bigger than T_h. The slave robot receives delayed signals in this phase.

In the second phase, the network is trained in T_2 seconds and the operator are asked to cease the operation. Due to small amount of T_2, this interruption is expected not to negatively affect the whole operation.

Finally, after $T_1 + T_2$ seconds, the NN is completely ready to be implemented in the system. Although the motion is expected not to be too far from the trained one in this phase, it is not indeed a restricting assumption, as we will show later in section 5 that the NN can predict even untrained motions and there is no need to keep the same motion during the whole operation.

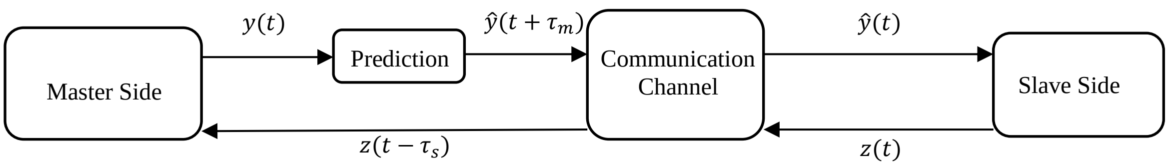


Figure 1. A teleoperation system with master state prediction. y and z are two arbitrary parameters

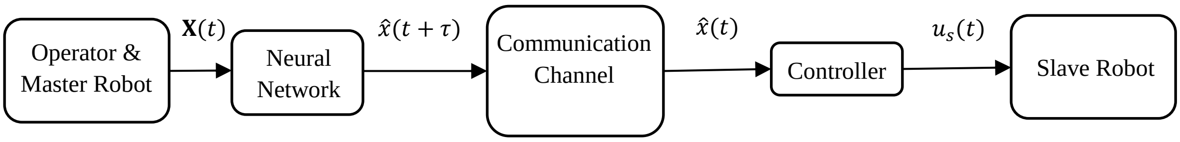


Figure 2. Architecture of a unilateral teleoperation system in which a neural network carries out the prediction task.

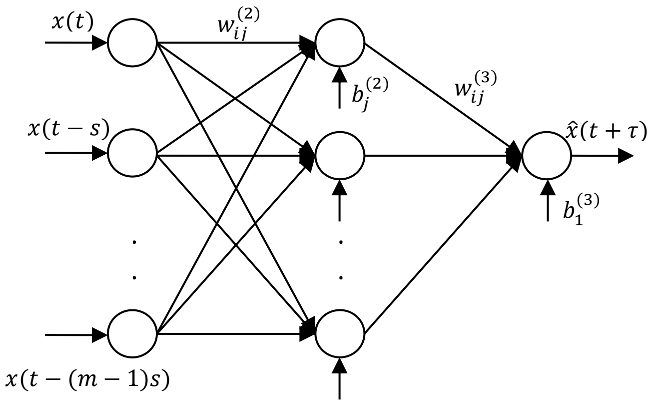


Figure 3. Structure of the proposed neural network.

3. Controller design for the slave side

3.1. Inverse kinematics of Phantom Omni

Phantom Omni haptic device provides three translational and three rotational degrees of freedom. The translational degrees of freedom are resulted from the three angles q_i $(i = 1, 2, 3)$ which are depicted in Figure 4. The inverse kinematics relations are expressed as follows [25]

$$q_1 = -\mathrm{atan2}(x_e, z_e + L_4); \quad q_2 = \gamma + \beta; \quad q_3 = q_2 + \alpha - \frac{\pi}{2} \tag{3.1}$$

where $\alpha := \cos^{-1}(\frac{{L_1}^2+{L_2}^2-r^2}{2L_1L_2})$, $\beta := \mathrm{atan2}(y_e - L_3, R)$, $\gamma := \cos^{-1}(\frac{{L_1}^2+r^2-{L_2}^2}{2L_1 r})$, $R := \sqrt{{x_e}^2 + (z_e + L_4)^2}$ and $r := \sqrt{{x_e}^2 + (z_e + L_4)^2 - (y_e - L_3)^2}$. Also $L_1 = L_2 = 133.35\, mm$, $L_3 = 23.35\, mm$, and $L_4 = 168.35\, mm$. For more detailed description and proof, please refer to the main reference [25].

3.2. Controller design

PID controller is adopted for the slave side in this paper. If the actuators are in Cartesian space themselves, the error can be defined straightforwardly as $\tilde{e}^{(x)}(t) := x_m(t-\tau) - x_s(t)$ before and during the NN training, and $e^{(x)}(t) := \hat{x}_m(t) - x_s(t)$ after the NN training. However, if the actuators function in joint space, the error should be defined as $\tilde{e}_i^{(q)}(t) := q_{i,m}(t-\tau) - q_{i,s}(t)$ before and during the NN training, and $e_i^{(q)}(t) := \hat{q}_{i,m}(t) - q_{i,s}(t)$ after the NN training for each joint (*i = 1, 2, 3*). Hence, the control input for the slave robot is obtained as

$$u_s(t) = k_p^{(x)} e^{(x)}(t) + k_d^{(x)} \dot{e}^{(x)}(t) + k_I^{(x)} \int e^{(x)}(t)dt \tag{3.2}$$

in Cartesian space, and

$$u_{s,i}(t) = k_{p,i}^{(q)} e_i^{(q)}(t) + k_{d,i}^{(q)} \dot{e}_i^{(q)}(t) + k_{I,i}^{(q)} \int e_i^{(q)}(t)dt \tag{3.3}$$

in joint space, where u_s is the slave control input, k_p, k_d and k_I are PID constant gains, superscripts *(x)* and *(q)* denote the Cartesian and joint space respectively, and subscript *i = 1, 2, 3* represents the joints of Phantom Omni. Note that before and during the NN training, $e(t)$ is replaced by $\tilde{e}(t)$ in (3.2) and (3.3).

We consider a one-degree-of-freedom motion in x-direction of Cartesian coordinates in this paper. So, for obtaining $\hat{q}_{i,m}(t)$ from the inverse kinematic relations we set $\hat{y}_m(t) = y_{des}$ and $\hat{z}_m(t) = z_{des}$, where y_{des} and z_{des} are constant desired positions in y- and z-directions for the end effector on which it is going to stay during the operation, while $\hat{x}_m(t)$ is attained from the NN output (section 2).

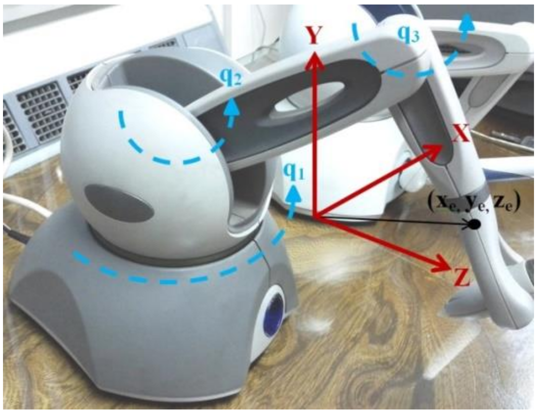


Figure 4. Phantom Omni reference XYZ coordinates, joints q_i and their resulting translational DoFs.

4. Stability analysis

There are well-known approaches for analyzing stability of teleoperation systems, such as absolute stability which considers the system as an input/output network [26], passivity which deals with energy generation/dissipation of the system [27], and methods such as Lyapunov analysis [28]. We use the absolute stability method in this paper. For this aim, the two-port network (Figure 5) should first be modeled as

$$\begin{bmatrix} F_h(s) \\ -V_s(s) \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} V_m(s) \\ F_e(s) \end{bmatrix} \tag{4.1}$$

where $V_s(s)$, $V_m(s)$,$F_h(s)$ and $F_e(s)$ are the velocity of the slave and master robots, and the human and environment forces, respectively, in Laplace space. The matrix containing the h-parameters is the hybrid matrix which is defined by the following relations if we assume linear second order dynamics for the desired behavior of the master and slave robots:

$$\begin{cases} h_{11} := \left.\dfrac{F_h(s)}{V_m(s)}\right|_{F_e=0} = \bar{m}_m s + \bar{b}_m + \dfrac{\bar{k}_m}{s} \\ h_{12} := \left.\dfrac{F_h(s)}{F_e(s)}\right|_{V_m=0} = k_1 e^{-\tau_s s} \\ h_{21} := \left.\dfrac{-V_s(s)}{V_m(s)}\right|_{F_e=0} = -k_2 e^{-\tau_m s} \\ h_{22} := \left.\dfrac{-V_s(s)}{F_e(s)}\right|_{V_m=0} = \dfrac{s}{\bar{m}_s s^2 + \bar{b}_s s + \bar{k}_s} \end{cases} \tag{4.2}$$

where $\bar{m}$, $\bar{b}$, $\bar{k}$ and τ represent desired mass, desired damping coefficient, desired stiffness and time delay, respectively, and the subscripts m and s denote the master and slave robots, respectively. $k_1 > 0$ and $k_2 > 0$ are also scaling factors for force and position, respectively.

Then, according to Llewellyn's stability criteria, the system is absolutely stable if the following conditions are met:

- No poles of h_{11} and h_{22} lie on the right half plane
- If any of poles of h_{11} and h_{22} lie on the imaginary axis, they are simple with real and positive residues
- For all real values of ω:
  (a): $\mathrm{Re}(h_{11}) \geq 0$
  (b): $\mathrm{Re}(h_{22}) \geq 0$
  (c): $2\mathrm{Re}(h_{11})\mathrm{Re}(h_{22}) - \mathrm{Re}(h_{12}h_{21}) - |h_{12}h_{21}| \geq 0$

The above conditions, provided that the operator and environment are passive (which is almost always the case [29]), result in a stable teleoperation system. The first and second conditions in addition to (a) and (b) of the third condition are fulfilled by choosing positive desired impedance parameters [30]. However, satisfying relation (c) of the third condition restricts the choices for the impedance parameters based on the time delay of the system.

Now we want to investigate the stability of the system if we add prediction blocks on master and slave sides. We have ideally $\hat{F}_e(s) = F_e(s)e^{\tau_s s}$ and $\hat{V}_m(s) = V_m(s)e^{\tau_m s}$ for the components which relate the transmitting signals together. By some mathematical manipulation, it is straightforward that $\hat{h}_{11} = h_{11}$ and $\hat{h}_{22} = h_{22}$. Also the other two components can be obtained as $\hat{h}_{12} := \left.\frac{F_h(s)}{\hat{F}_e(s)}\right|_{V_m=0} = k_1$ and $\hat{h}_{21} := \left.\frac{-V_s(s)}{\hat{V}_m(s)}\right|_{F_e=0} = -k_2$. The first and second conditions and (a) and (b) of the third one have obviously not been changed in comparison with (4.2), since the predictors do not affect the operator-master robot and the environment-slave robot interactions, and thus $\mathrm{Re}(h_{11})\mathrm{Re}(h_{22}) \geq 0$. Therefore, only (c) of the third condition should be investigated in order to prove the stability of the system. That condition is also met, because we have

$$\begin{aligned} 2\mathrm{Re}(\hat{h}_{11})\mathrm{Re}(\hat{h}_{22}) - \mathrm{Re}(\hat{h}_{12}\hat{h}_{21}) - |\hat{h}_{12}\hat{h}_{21}| \\ = 2\mathrm{Re}(h_{11})\mathrm{Re}(h_{22}) + 2k_1k_2 \\ \geq 0 \end{aligned} \tag{4.3}$$

where k_1k_2 is positive due to the definition presented previously.

Finally, provided that the outputs of the prediction block accurately converge to the actual values of their respective input parameters and the designed controller in section 3 is stable, we can conclude from the above-mentioned explanations that the whole proposed teleoperation system in this paper is stable, since it is clearly a special case of the system investigated above.

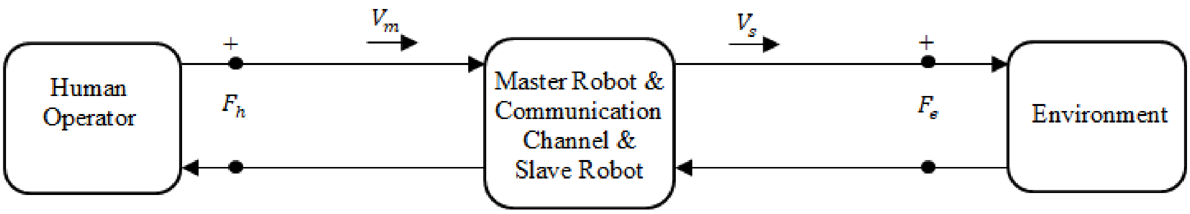


Figure 5. A two-port teleoperation system with classical architecture

5. Results

In the following subsections, the simulation and experimental results are presented. The parameters regarding the whole system and the NN are listed in Table 1. To find the proper sample time for training the NN, we ran the simulation test with different sample times, and we arrived at a conclusion that the best one is *0.01 s*, as mentioned in Table 1. To provide an example, we compared the sample times of *0.01 s* and *0.001 s*. It was found that with fixed structures, the NN training time in the latter case was approximately more than ten times the former case, and furthermore, the generality of the latter case results was highly poorer than the former case.

Two scenarios are designed in this section to validate the effectiveness of the proposed approach, the descriptions of which are given below:

Scenario I: In this scenario, a simple input force with an amplitude and frequency is applied by the operator and the same motion continues after the training. This scenario is to investigate whether or not the online trained NN is able to reproduce the results presented to it as input.

Scenario II: In this scenario, a more complex input force with amplitude and frequency is applied by the operator and the same motion continues after the training, but after some time, the amplitude and frequency are altered. This is done in order to investigate the prediction capability of the trained NN in the proposed teleoperation architecture when encountering untrained motions.

Note that in Figures 6 to 9 in this section, the end of $T_1$is marked by the first vertical dash-dot line, and the end of T_2 is marked by the second dash-dot line. It should also be stated that all inputs to the NNs are normalized in the interval $[-1,+1]$.

5.1. Simulation results

For simulation, a one-degree-of-freedom teleoperation system in x-direction of Cartesian XYZ system was simulated by means of Simmechanics library (second generation) of MATLAB's Simulink. The dynamics of the master and slave robots were considered as linear second order ones with $M_m = M_s = 0.223\,kg$, $B_m = B_s = 17.227\,N.s/m$ and $K_m = K_s = 6.286\,N/m$, where M, B and K denote mass, damping coefficient and stiffness, respectively. This dynamics can represent a Phantom Omni robot [31]. NN training is performed using batch-mode method. A constant time delay τ is also simulated for the system. The slave control input in this case is achieved from (3.2). The system and NN parameters for the two scenarios are summarized in Table 2.

Scenario I: The simulated human force which is applied to the master robot is illustrated in Figure 6(a). The results for this scenario are depicted in Figures 6(b) to 6(d).

As can be seen from Figure 6(b), the slave robot is able to predict the master robot motion and move ahead of it, which results in simultaneous motion of the master and slave robots, as can be seen in Figure 6(c). This leads to a highly transparent, yet stable, teleoperation system in which the time delay is not felt by the operator. A comparison between the pre-training and post-training portions of the operation in Figure 6(c) will further prove this fact. It is noticeable that a code written in MATLAB environment indeed needs more amount of time to train a NN in comparison with codes written in C or C++, as will be shown later in section 5-2. Therefore, the amount of time T_2 required for NN training is much larger in simulation study than in experimental test in this paper.

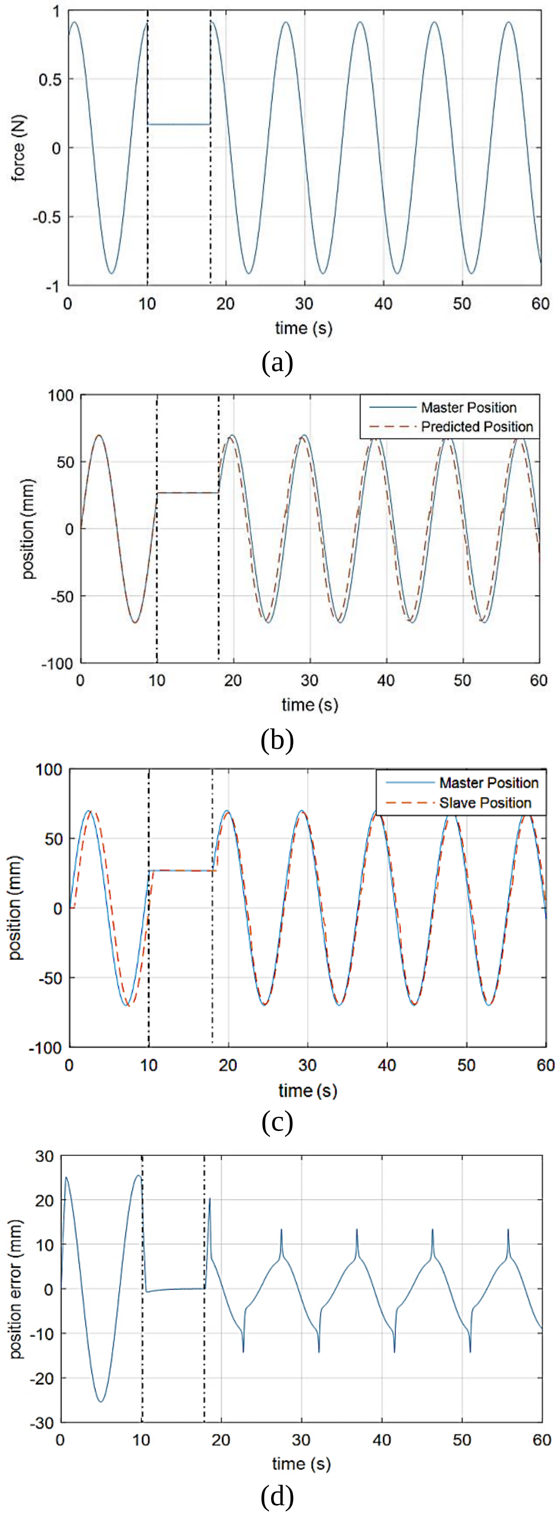


Figure 6. Simulation of Scenario I. (a) Human force. (b) Master position and predicted master position. Note that prediction occurs after T2 (second dash-dot line). (c) Master and slave positions. (d) Position error. Note that during the whole operation, including T1+T2, the definition for e^(x) is depicted.

Table 1. System and NN parameters which are in common among all simulation and experimental tests.

Parameter	Description	Value	Unit
m	Number of input layer neurons	10	-
s	System sample time	0.01	sec
τ	System time delay	0.5	sec
T_1	Time for gathering NN inputs	10	sec
-	Number of hidden layer neurons for Scenario I	10	-
-	Number of hidden layer neurons for Scenario II	18	-
-	Minimum value of objective function to stop NN training	0.001	-

Table 2. System and NN parameters for both scenarios in simulation study.

Parameter	Description	Value	Unit
η	NN learning rate	0.04	-
T_2	NN training time	8	sec
-	Maximum epochs to stop NN training	4000	-
$k_p^{(x)}$	Proportional PID gain	250	N/m
$k_d^{(x)}$	Derivative PID gain	0	N.s/m
$k_I^{(x)}$	Integral PID gain	155	N/(m.s)

Scenario II: The simulated human force which is applied to the master robot in this scenario is illustrated in Figure 7(a). The results for this scenario are shown in Figures 7(b) to 7(d).

It can be inferred from Figures 7(b) to 7(d) that not only the results of Scenario I are valid for more complex motions, but also using a NN in the system enables it to predict new motions if the amplitude and frequency are not too far from those of the trained motion. In other words, the NN can correctly predict newly intended motions of the operator which had been unknown to the system before their first appearance, particularly in the first phase. As can be seen in Figure 7(b), The NN has performed a good prediction of the master robot position during the whole operation. Therefore, in spite of large amount of time delay in the system, Figures 7(c) and 7(d) show that the slave robot has successfully tracked the master robot position in advance and the system is stable and highly transparent.

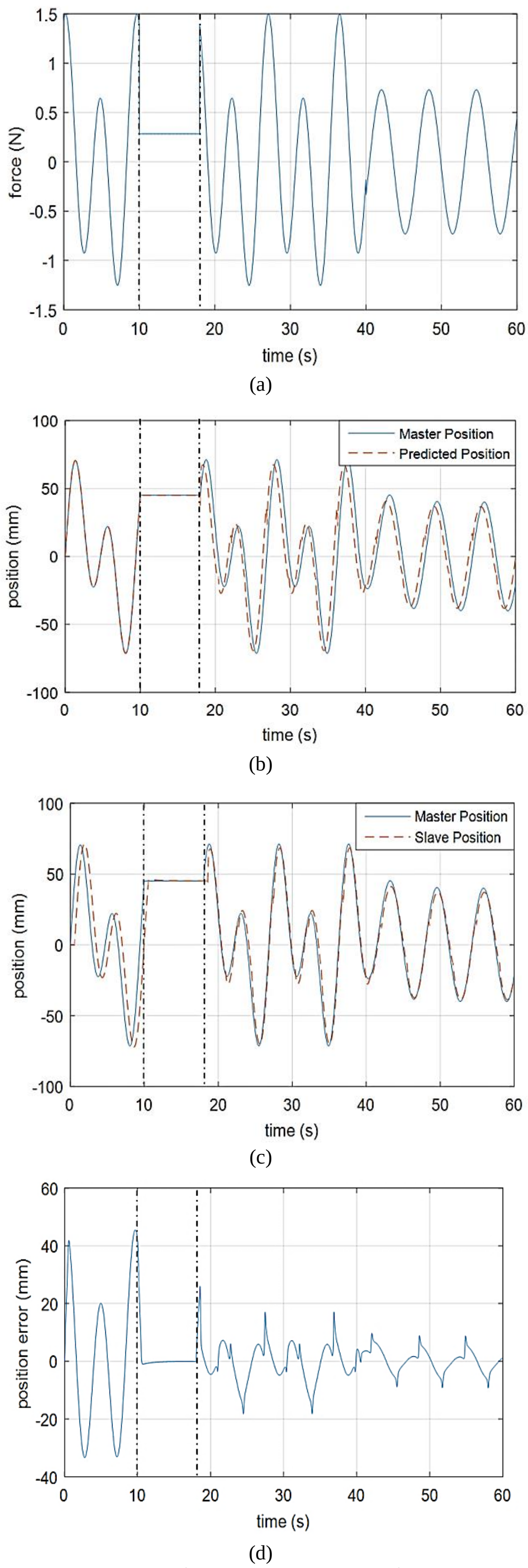


Figure 7. Simulation of Scenario II. (a) Human force. (b) Master position and predicted master position. Note that prediction occurs after T2 (second dash-dot line). (c) Master and slave positions. (d) Position error. Note that during the whole operation, including T1+T2, the definition for $e^{\wedge(x)}$ is depicted.

5.2. Experimental results

In this section, the end effector of Phantom Omni moves only in x-direction with respect to the reference Cartesian coordinates (Figure 4). Hence, three controllers should be designed for the joints q_1, q_2 and q_3.

The online NN training is carried out by incremental-mode method through Fast Artificial Neural Network (FANN) library for C++. A constant time delay τ is considered for the system. The system and NN parameters for these scenarios are summarized in Table 3.

Table 3. System and NN parameters for both scenarios in experimental test.

Parameter	Description	Value	Unit
η	NN learning rate	0.7	-
T_2	NN training time	4	sec
-	Maximum epochs to stop NN training	2000	-
$k_{p,1}^{(q)}$	Proportional PID gain for q_1	4	N.m/rad
$k_{d,1}^{(q)}$	Derivative PID gain for q_1	0.15	N.m.s/rad
$k_{I,1}^{(q)}$	Integral PID gain for q_1	0.4	N.m/(rad.s)
$k_{p,2}^{(q)}$	Proportional PID gain for q_2	2.2	N.m/rad
$k_{d,2}^{(q)}$	Derivative PID gain for q_2	0.01	N.m.s/rad
$k_{I,2}^{(q)}$	Integral PID gain for q_2	0.06	N.m/(rad.s)
$k_{p,3}^{(q)}$	Proportional PID gain for q_3	2.2	N.m/rad
$k_{d,3}^{(q)}$	Derivative PID gain for q_3	0.01	N.m.s/rad
$k_{I,3}^{(q)}$	Integral PID gain for q_3	0.06	N.m/(rad.s)

Scenario I: The results for this scenario are illustrated in Figures 8(a) to 8(c).

Consistent with the simulation results, Figure 8(a) shows that the motion of the master robot is well predicted by the NN and the slave robot has tracked the master position ahead, which is further demonstrated by Figure 8(b) in which the master and slave robots are shown to be moving simultaneously. High transparency of the system can also be deduced by comparing pre- and post-training intervals in Figures 8(b) and 8(c). Also, to quantitatively prove the stated facts, the root-mean-square error (RMSE) and the normalized root-mean-square error (NRMSE) during the prediction period for this scenario are obtained to be *7.42 mm* and *5.57%*, respectively. The NRMSE is defined by (5.1), where $max\{x_m\}$ and $min\{x_m\}$ denote the maximum and minimum position of the master robot, respectively, during the prediction period.

$$\text{NRMSE} := \frac{\text{RMSE}}{max\{x_m\} - min\{x_m\}} \tag{5.1}$$

Meanwhile, as mentioned earlier in section 5-1, the time T_2 in experimental test is seen to be smaller because NN is trained much faster by FANN than in MATLAB environment.

Scenario II: The results for this scenario are depicted in Figures 9(a) to 9(c).

Again, the results in Figures 9(a) to 9(c) show high transparency of the system. Consistent with the results obtained through simulation, Figure 9(a) indicates that the system is capable of predicting both trained and untrained motions of the master robot, despite having a large amount of time delay in the system. Furthermore, online tracking capability of the slave robot can be seen in Figures 9(b) and 9(c), especially by comparing pre- and post-training time intervals. To further prove these facts quantitatively, the RMSE and the NRMSE (as defined by (5.1)) during the prediction period for this scenario are obtained to be *11.71 mm* and *7.04%*, respectively.

Of course, looking more closely at Figure 9(b), the slave robot is observed not to precisely track the master robot position at some points, around $t = 28\ s$ for instance. This is because the NN is trained by a small amount of data in comparison with the applied motions after the training. In other words, unlike Scenario I, the slave robot does not have much information about what motion the human operator is going to apply next. Consequently, it may fail to accurately predict the operator motion at sometimes.

However, figure 9 (b) shows that the variable amplitude and frequency of the applied motion is well predicted despite almost merely one period of input training data to the NN. If more accurate

prediction is desired, obviously more input data should be presented to the NN which, in turn, takes more time and may seem a negative point to the operator.

Simulation and experimental results presented in this section also show that both batch-mode and incremental-mode trainings have yielded reliable results.

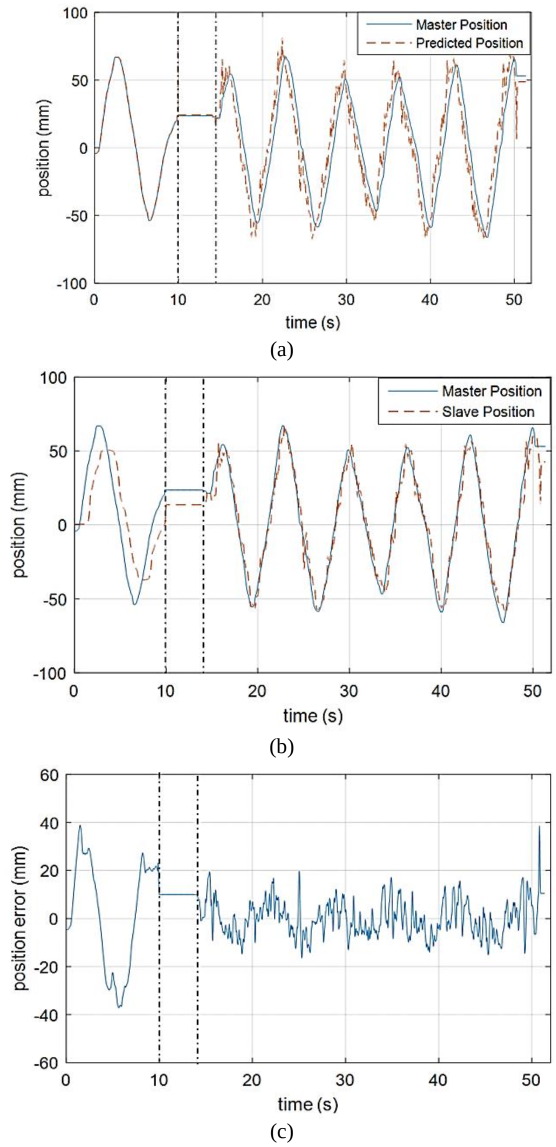
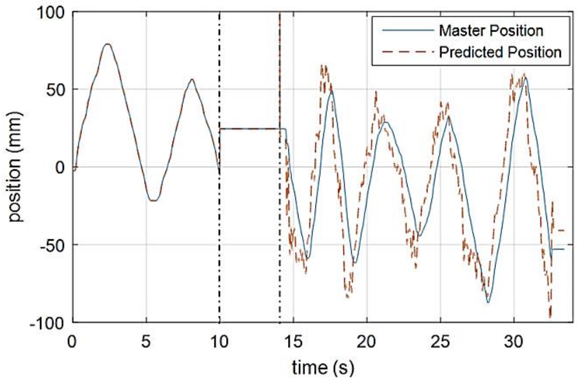
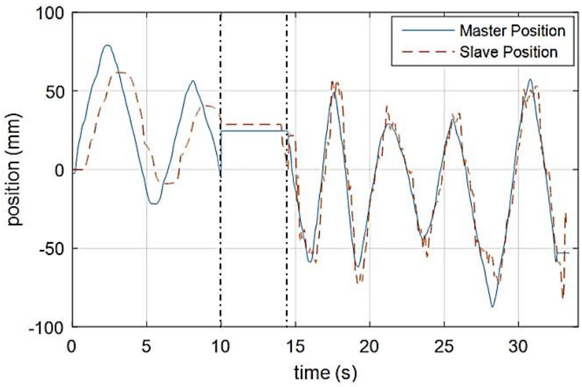


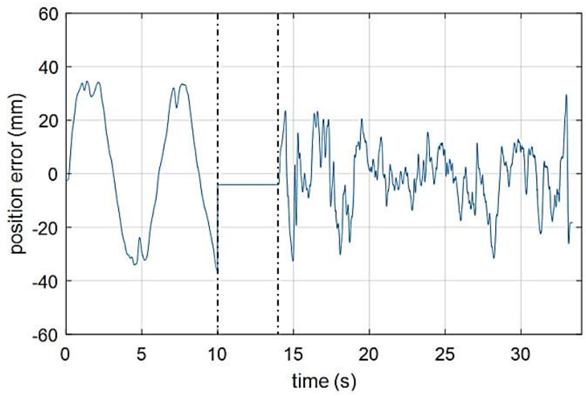
Figure 8. Experiment of Scenario I. (a) Master position and predicted master position. Note that prediction occurs after T2 (second dash-dot line). (b) Master and slave positions. (c) Position error. Note that during the whole operation, including T1+T2, the definition for $e^{\wedge(x)}$ is depicted.



(a)



(b)



(c)

Figure 9. Experiment of Scenario II. (a) Master position and predicted master position. Note that prediction occurs after T2 (second dash-dot line). (b) Master and slave positions. (c) Position error. Note that during the whole operation, including T1+T2, the definition for $e^{\wedge(x)}$ is depicted.

6. Conclusion and future works

In this paper, a neural network was incorporated on master side in a unilateral teleoperation system in order to predict the future position of the operator's hand online. The discussion on the superiority of the presented method in comparison with previously proposed methods in the literature was provided, too.

The simulation and experimental results indicated that the teleoperation system performs with high transparency, which means the slave robot can well track the master robot motions without delay. Meanwhile, the NN predictor can predict not only the trained motion, but also untrained motions.

Possible future directions include considering stochastic time delay in communication channel, taking environment force into account and reducing the time of motion halt during training phase.

References

[1] R. Uddin and J. Ryu, "Predictive control approaches for bilateral teleoperation". Annual Reviews in Control, Vol. 42, 2016, 82-99.

[2] C. Passenberg, A. Peer, and M. Buss, "A survey of environment-, operator-, and task-adapted controllers for teleoperation systems". Mechatronics, Vol. 20, No. 7, 2010, 787-801.

[3] X. Xu, B. Cizmeci, C. Schuwerk, and E. Steinbach, "Model-Mediated Teleoperation: Toward Stable and Transparent Teleoperation Systems". IEEE Access, Vol. 4, 2016, 425-449.

[4] B. Yazdankhoo and B. Beigzadeh, "Increasing stability in model-mediated teleoperation approach by reducing model jump effect". Scientia Iranica, Vol. 26, Special Issue on: Socio-Cognitive Engineering, 2019, 3-14.

[5] P. Mitra and G. Niemeyer, "Model-mediated telemanipulation". The International Journal of Robotics Research, Vol. 27, No. 2, 2008, 253-262.

[6] N. Jarrassé, J. Paik, V. Pasqui, and G. Morel, "How can human motion prediction increase transparency?". IEEE International Conference on Robotics and Automation (ICRA), Pasadena, CA, USA, 2008, 2134-2139.

[7] R. Uddin, S. Park, and J. Ryu, "A predictive energy-bounding approach for Haptic teleoperation". Mechatronics, Vol. 35, 2016, 148-161.

[8] S. E. Engelbrecht, "Minimum principles in motor control". Journal of Mathematical Psychology, Vol. 45, No. 3, 2001, 497-542.

[9] C. Smith and P. Jensfelt, "A predictor for operator input for time-delayed teleoperation". Mechatronics, Vol. 20, No. 7, 2010, 778-786.

[10] D. Feth, A. Peer, and M. Buss, "Enhancement of multi-user teleoperation systems by prediction of dyadic haptic interaction". Experimental Robotics, 2014, 855-869.

[11] P. Prekopiou, S. G. Tzafestas, and W. S. Harwin, "Towards variable-time-delays-robust telemanipulation through master state prediction". IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Atlanta, USA, 1999, 305-310.

[12] S. Clarke, G. Schillhuber, M. F. Zaeh, and H. Ulbrich, "Prediction-based methods for teleoperation across delayed networks". Multimedia Systems, Vol. 13, No. 4, 2008, 253-261.

[13] F. Stakem and G. AlRegib, "An adaptive approach to exponential smoothing for CVE state prediction". 2nd International Conference on Immersive Telecommunications, Berkley, USA, 2009, 141-146.

[14] N. Sakr, N. D. Georganas, J. Zhao, and X. Shen, "Motion and force prediction in haptic media". IEEE International Conference on Multimedia and Expo, Beijing, China, 2007, 2242-2245.

[15] A. Pentland and A. Liu, "Modeling and prediction of human behavior". Neural computation, Vol. 11, No. 1, 1999, 229-242.

[16] S. Shen, A. Song, T. Li, and H. Li, "Time delay compensation for nonlinear bilateral teleoperation: A motion prediction approach". Transactions of the Institute of Measurement and Control, 2019.

[17] E. S. Gardner Jr, "Exponential smoothing: The state of the art—Part II". International journal of forecasting, Vol. 22, No. 4, 2006, 637-666.

[18] Y. Ji, D. Liu, and Y. Guo, "Adaptive neural network based position tracking control for Dual-master/Single-slave teleoperation system under communication constant time delays". ISA transactions, Vol. 93, 2019, 80-92.

[19] P. M. Kebria, A. Khosravi, S. Nahavandi, D. Wu, and F. Bello, "Adaptive Type-2 Fuzzy Neural-Network Control for Teleoperation Systems with Delay and Uncertainties". IEEE Transactions on Fuzzy Systems, 2019.

[20] C.-C. Hua, Y. Yang, and X. Guan, "Neural network-based adaptive position tracking control for bilateral teleoperation under constant time delay". Neurocomputing, Vol. 113, 2013, 204-212.

[21] J. Nava and V. Kreinovich, "Why A Model Produced by Training a Neural Network Is Often More Computationally Efficient than a Nonlinear Regression Model: A Theoretical Explanation". Journal of Uncertain Systems, Vol. 8, No. 3, 2014, 193-204.

[22] V. Nicolau, V. Palade, D. Aiordachioaie, and C. Miholca, "Neural network prediction of the roll motion of a ship for intelligent course control". International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, Vietri sul Mare, Italy, 2007, 284-291.

[23] G.-Z. Yang et al., "The grand challenges of Science Robotics". Science Robotics, Vol. 3, No. 14, 2018, eaar7650.

[24] K. Mehrotra, C. K. Mohan, and S. Ranka, Elements of artificial neural networks. Cambridge: MIT press; 2000.

[25] T. Sansanayuth, I. Nilkhamhang, and K. Tungpimolrat, "Teleoperation with inverse dynamics control for phantom omni haptic device". SICE Annual Conference, Akita, Japan, 2012, 2121-2126.

[26] F. Llewellyn, "Some fundamental properties of transmission systems". Proceedings of the IRE, Vol. 40, No. 3, 1952, 271-283.

[27] D. Lee and M. W. Spong, "Passive bilateral teleoperation with constant time delay". IEEE transactions on robotics, Vol. 22, No. 2, 2006, 269-281.

[28] X. Liu and X. Dong, "4-Channel Sliding Mode Control of Teleoperation Systems with Disturbances". Asian Journal of Control, Vol. 17, No. 4, 2014, 1267-1273.

[29] H. C. Cho and J. H. Park, "Impedance control with variable damping for bilateral teleoperation under time delay". JSME International Journal Series C Mechanical Systems, Machine Elements and Manufacturing, Vol. 48, No. 4, 2005, 695-703.

[30] L.-G. García-Valdovinos, V. Parra-Vega, and M. A. Arteaga, "Observer-based sliding mode impedance control of bilateral teleoperation under constant unknown time delay". Robotics and Autonomous Systems, Vol. 55, No. 8, 2007, 609-617.

[31] T. Hilliard and Y.-J. Pan, "Stabilization of asymmetric bilateral teleoperation systems for haptic devices with time-varying delays". American Control Conference, Washington, DC, USA, 2013, 4538-4543.

JJMIE

Volume 13, Number 3, October. 2019
ISSN 1995-6665
Pages 207 - 220

Jordan Journal of Mechanical and Industrial Engineering

Renewable Resources of Energy for Electricity Generation – Development Trends and Necessities Within the Overall Energy System

Jelto Lange[a*], Jerrit Hilgedieck[a], Martin Kaltschmitt[a]

[a] Institute of Environmental Technology and Energy Economics, Hamburg University of Technology, Hamburg, Germany

Received 20 May. 2019

Abstract

Global primary energy demand will most likely increase in the coming decades as a result of a growing world population and generally increasing prosperity. Historically, increasing energy demand has usually been associated with increases in greenhouse gas (GHG) emissions as the additional energy needs are met by the utilization of fossil fuels. However, in order to meet the efforts to curb climate change, both existing and emerging energy demands (and therefore energy supply systems) must be sustainably covered using renewable energy resources. In this context, two different development scenarios are presented in this paper: First, based on historical developments in recent years, the given expansion pathway of renewable energies is extrapolated until the year 2050 (bottom-up perspective). However, as this expansion does not meet the climate goals set to meet the 1.5 °C target, additionally a development pathway, which meets the GHG reduction goals in 2050 is developed (top-down perspective). Given the very high levels of energy delivered from (fluctuating) renewable energy sources needed for such a sustainable development path, comprehensive decarbonization (concerning fossil carbon) requires a much better interconnection of the different energy sectors (electricity, heat, mobility) compared to today; some of these necessities are also presented because they influence the further development of the energy system and thus the role of renewables considerably. All over, the paper shows that the already ongoing efforts to achieve the 1.5 °C target must be intensified dramatically; the already significant developments related to the installation of new systems to use renewable sources of energy are not enough to achieve the defined targets.

Keywords: Renewable energies, primary energy demand, greenhouse gas emissions, system-transformation;

1. Introduction

Global primary energy demand amounted to roughly 547 EJ (152 PWh) in 2017 ([1, 2] and own calculations). This energy demand is expected to grow significantly in the coming decades summing up to an overall energy consumption of roughly 670 EJ/a (187 PWh/a) [3] until the year 2040. Around 88 % of this demand is covered by using fossil fuel energy (namely coal, natural gas and crude oil), whose energy-related use leads to a release of additional (anthropogenic) carbon dioxide as well as other greenhouse gases (GHG), which together in turn affect global climate [3]. Especially the combustion of such fossil fuels in a variety of consumption sectors constitutes the main contribution to manmade greenhouse gas emissions and therefore the so-called anthropogenic greenhouse effect. Thus, Fig. 1 shows the specific GHG emissions in relation to the specific energy demand of different countries related to the year 2014. A clear interrelationship between high-income and high GHG emissions is apparent. Higher income is related to an increased consumption of energy and thus (based on current energy supply structures) higher GHG emissions. To address the increasing riskiness of global climate change, the overarching goal should be to drastically reduce such emissions. Stopping or hindering global development towards increasing energy use (due to prosperity reasons) cannot be a widely acceptable solution. Therefore – in order to limit global GHG-emissions – the establishment of global low-GHG emission or GHG emission-free energy supply structures should be the overarching goal of energy and environmental politics within the various governments throughout the world. Consequently, energy supply systems in countries with already high or even very high GHG emissions must be transformed in a way, that a high energy demand can increasingly be met without releasing (significant amounts of) GHG emissions (dark grey arrow in Fig. 1). Likewise, countries with low GHG-emissions – and this usually means so far low-income countries (light grey circles) – must be able to increase the amount of supplied energy – and thus welfare and standard of living – while also slightly reducing (or at least without increasing) GHG emissions (light grey arrow).

To achieve such a development goal contributing towards a more sustainable development, energy supply

* Corresponding author e-mail: j.lange@tuhh.de

must become increasingly less dependent on fossil carbon-based energy carriers (i.e. coal, crude oil, natural gas). The only way to do this in a sustainable manner (without producing nuclear waste) and based on acceptable costs lies in a strongly expanded use of renewable energies – and this means for the time being a clearly increased utilization of wind energy and solar radiation.

Energy conversion technologies and systems using renewable energies mostly provide electrical energy. Due to the almost arbitrarily realizable transformation of electricity to other forms of energy (e.g. heat, mechanical energy), electrical energy from renewables can be used to substitute carbon-based fossil fuels in basically all sectors of the overall energy system, enabling (on the longer term) a basically climate-neutral energy supply. Nevertheless, such fundamental transformations will imply additional changes within the overall energy system. An appropriate and improved coupling of different energy sectors (e.g., electricity sector, heat sector, mobility sector) is needed to integrate lager shares of electricity in formerly un-electrified energy sectors (e.g., mobility). Also, some additional applications are needed to sustain a secure system operation especially of the electricity supply system even with high shares of wind and solar power production characterized by strong fluctuations and clear regional supply differences.

As a result of the Kyoto process, the need for the significant reduction of GHG emissions has been recognized by more and more countries (90 % of the warmest years since the start of temperature records in 1880 occurred within the past decade [6], which is becoming more and more evident globally). The consequences of this slow process of cognition have been a multitude of multilateral and international agreements aiming to reduce GHG emissions in the coming years. Even though global GHG emissions continue to rise significantly due to over-proportional rising welfare, such agreements – even if certain states have temporarily stepped out for primarily domestic political reasons – support a rethinking process within the global energy industry. Thus, on a global scale, in recent years, most capacity additions within the electricity sector have been in systems using renewable energies – and here in particular photovoltaic and wind powered plants. For example, in 2017 157 GW of renewable power production capacity was installed, by far exceeding the 70 GW for conventional power plants [7].

Against this background, first the most important drivers for the development of global energy demand are elucidated. Then, based on the current energy supply system structure, the global energy demand for the year 2050 is estimated (the forecast for 2040 from [3] is extrapolated till 2050 by assuming an overall growth rate that is 50 % of the growth rate between today and 2040). This global energy demand implicitly states the population's demand for usable energy (considering conversion efficiencies). Based on this demand, two development pathways until the year 2050 are shown – on the one hand based on currently visible and already ongoing worldwide developments especially for renewable energies (bottom-up) and on the other hand based on the necessary reductions of GHG emissions to achieve the so called 1.5 °C-goal (top-down), assuming the same demand for usable energy in all cases. Finally, some conclusions are drawn and discussed. The approach is schematically depicted in Fig. 2.

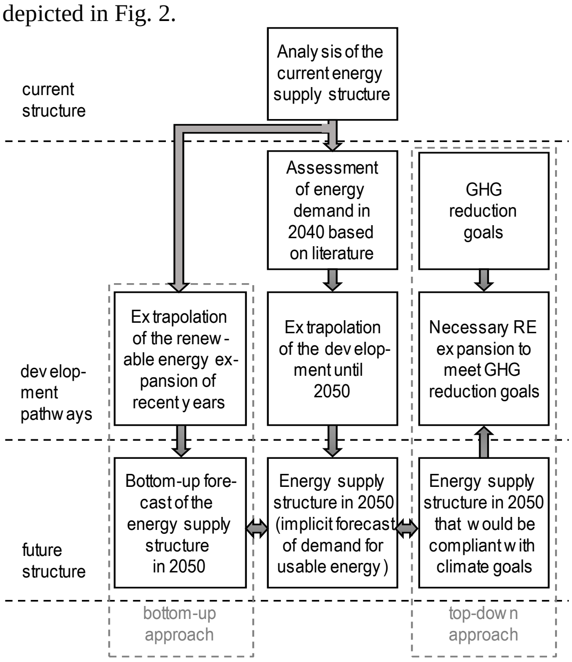


Figure 2. Schematic illustration of the approach

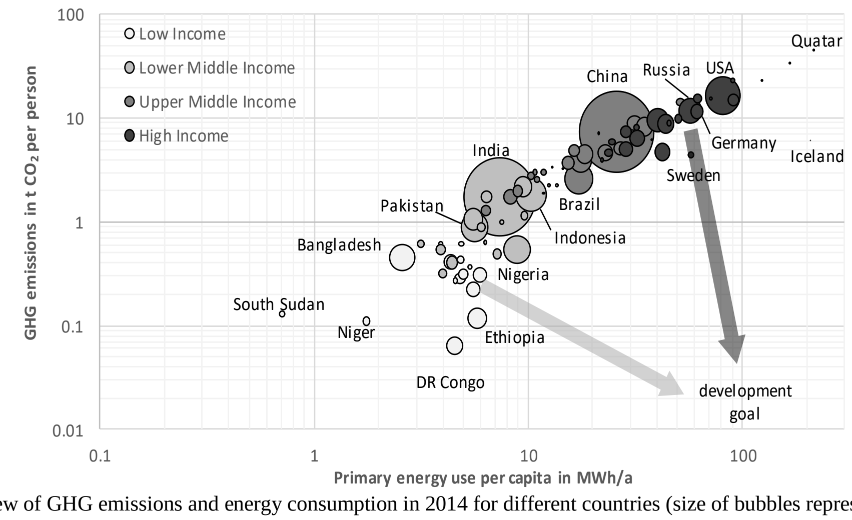


Figure 1. Overview of GHG emissions and energy consumption in 2014 for different countries (size of bubbles represents the population of each country) [4, 5]

There are multiple different forecasts for the development of the global energy supply for the year 2040 / 2050 (e.g. [3, 8–10]). Additionally, there are different elaborations on development pathways for reaching an energy supply structure that is compliant with climate goals (e.g. [11, 12]). The following elaboration serves as a connection between both types of forecasts by explicitly pointing out the gap between the system development that is currently taking place and one that would be needed in order to evade the threat of a global climate crisis.

2. Global Trends and key drivers

The further development of the energy demand and thus of the global energy supply system are fundamentally determined and driven by only very few main key drivers. These globally visible trends have significant and profound implications on the worldwide development of the various energy systems; they are described hereinafter.

2.1. Growing energy demand

Certain ongoing and irreversible trends inevitably lead to a strong increase in global energy demand. The main driving developments (world population and prosperity) are discussed below.

World population. The current world population of around 7.5 billion people demands on average around 20.1 MWh/a or 72.5 GJ/a of primary energy per person ([13, 14]). By the year 2100, about 10 to 11 billion people will populate our planet (Fig. 3) [13]. Due to the still increasing life expectancy in connection with birth rates, which are decreasing but still amount to roughly 2.5 babies per woman on average [15], there will be a significant increase in world population in the coming decades. Especially in highly populated countries such as China or India, the sharply increasing life expectancy in recent decades is particularly apparent. Under the assumption, that the global energy demand (without consideration of the individual energy demand) is at least proportional to the population, this alone will inevitably lead to a (strong) increase in energy demand in the future.

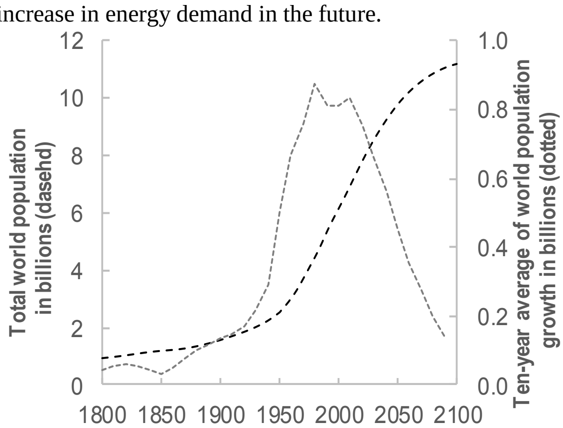


Figure 3. Development of the world population (dashed line: left axis; dotted line; right axis) [13]

Prosperity. In addition to population growth, the average prosperity of the overall society respectively of each average citizen is increasing on a global scale. In recent decades, there has been a clear trend towards significantly higher incomes of larger and larger population groups in an increasing number of countries with a former low-income level; i.e., the so-called middle class is growing very fast worldwide in absolute and relative terms. Consequently, growing prosperity leads to an over-proportionally rising energy demand (Fig. 4). The average global income increased by a factor of 2.5 between 1965 and 2014 with an even higher growth in energy consumption by a factor of 3.5. The decade with the highest growth was from 2000 to 2010, where emerging countries like China entered a phase of rapid economic development paired with a significant growth in energy demand. The growing incomes and the associated additional financial opportunities within more and more emerging countries lead to more and new needs, typically associated with (increasing) energy consumption (e.g., mobility, consumption of high-quality and energy-intensive food and consumer goods, air conditioning or communication).

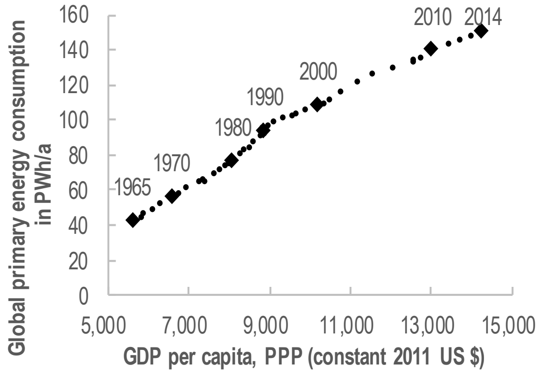


Figure 4. Yearly global average of energy consumption in relation to the yearly average global income (purchasing power parity, PPP) [5, 16]

In the medium to longer term, it is likely that the rise in energy demand due to increasing prosperity in many developing and emerging countries will approach the level of the industrialized OECD countries characterized today by a significantly higher demand of energy (Fig. 1). Even if the energy efficiency efforts should be more successful in the future than in the past in the industrialized countries (and in many developing and emerging countries), this development will most likely lead to a significant increase in global energy demand in the future.

The historical increase of global energy demand caused a steady increase in global GHG emissions over the past decades and centuries. As shown in Fig. 5, there is a strong correlation between increasing energy consumption and global GHG emissions due to the strong utilization of mostly carbon-based fossil fuels. This is due the fact that apart from hydropower (and nuclear power based on public subsidies) typically the use of fossil fuels has been the economically most promising (and most easily realizable) option for energy provision in the last century. For this reason, existing energy systems are largely based on fossil fuels, since these were (and are) available in large quantities and at low cost.

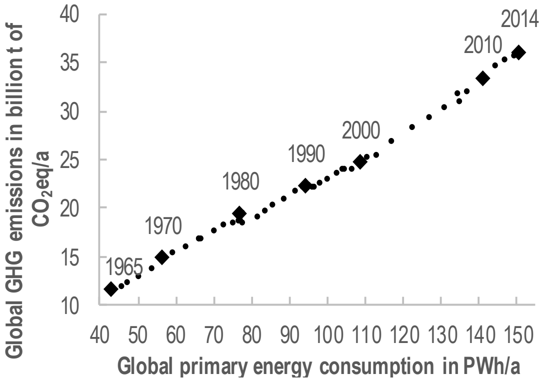


Figure 5. Yearly global average of GHG emissions in relation to energy consumption [15, 17]

2.2. Environmentally friendly energy

The strong economic growth in recent years and the corresponding sharp rise in energy consumption (due to rising population as well as growing prosperity) have led to massive consequences for the global energy supply system. Most of the power plants, which were built in the past, are based on the conversion of fossil fuels and therefore are responsible for emitting carbon dioxide and other substances like Sulphur dioxide and nitrogen oxide influencing the local as well as the global environment. Furthermore, the exploration, extraction and use of fossil fuels as well as the use of nuclear power is strongly linked with the risk of significant anthropogenic environmental catastrophes with a global impact (e.g., the Chernobyl and Fukushima Daiichi nuclear disaster, the Deepwater Horizon explosion and the resulting oil spill in the Gulf of Mexico or the disaster of the crude oil super tanker Exxon Valdez). Such significant impacts on local and global environment resulting in the loss of single species and even complete operating eco systems lead to a growing awareness related to environmental aspects within people of various countries on this globe. Additionally, an increasing sensitivity for a (more) sustainable development emerges globally from such environmental disasters, which are presented by media globally.

Consequently, research and technological development of the recent decades has focused primarily on a more efficient use of wind and solar power, as these represent options to address the described challenges and can enable an emission-free and sustainable provision of electricity respectively energy. Additionally, they can contribute to a rural electricity supply, which does only exist rudimentary (or not at all) in some developing and emerging countries. These technological trends towards more efficient and less expensive technologies for energy supply from renewable energy resources characterized by a modular installation have contributed significantly to the recent unexpected expansion of renewable energies on a global scale. And this development will continue in the years to come due to the reasons outlined above.

2.3. Interim conclusion

As explained in Chapter 2.1, global energy demand will most likely continue to rise sharply. In order to secure a climate and environmentally friendly energy provision, this mandatory rising energy demand has to be met mostly based on renewable energies. This necessity demands for the two different transformation pathways for energy systems in countries around the globe outlined below.

- Already developed energy systems, which predominantly exist in high-income and highly industrializes countries, have to be transformed by replacing conventional (fossil) energy provision with an increasing amount of renewable power plants (dark grey arrow in Fig. 1). These countries would have to follow a path of growth-neutral (or growth-unaffected) decarbonization concerning the carbon of fossil origin.
- Emerging energy supply systems, which predominantly exist in lower income countries (i.e., developing and emerging countries), will undoubtedly be characterized by growing energy systems coupled with growing prosperity. Nevertheless, these countries would have to manage a path of (fossil) carbon-neutral growth to allow compliance with global climate goals (light grey arrow in Fig. 1).

Hybrid paths (combinations of both paths) are also possible for countries in the medium income region (i.e., emerging and partly industrialized countries). Nonetheless, all these transformation paths are connected to a massive expansion of the utilization of renewable energies globally. Therefore, the following chapters analyze global expansion trends for renewable energies. Additionally, based on a bottom-up perspective, global shares of renewable energies are forecasted based on current development trends. Beside this, from a top-down perspective, based on the globally agreed GHG-emission limitations the necessary development of a use of renewable energies to be compliant with climate goals in the year 2050 is assessed.

3. Global energy supply – status and developments

Analyzing the current global energy supply system, a clear dominance of fossil fuels in terms of total primary energy supply is apparent (Fig. 6). In 2016, fossil fuels accounted for roughly 88 % of the world's primary energy supply (hydro and nuclear are treated directly as electrical energy); that means coal (hard coal and lignite), crude oil and natural gas are currently the most commonly used energy carriers worldwide. Renewable energies play only a minor role with a total contribution of roughly 9.1 % ([1, 2] and own calculation). The share of primary energy used for the provision of electricity is roughly 30 % globally (accounting for transformation losses for fossil fuels); i.e., about 70 % of worlds primary energy demand is used for heating, mobility and other purposes (e.g., used as a raw material for the chemical industry) [1]. The largest share of electricity generation is provided based on coal-fired power plants; crude oil respectively its derivate are mostly used for mobility and natural gas is flexibly divided between different sectors.

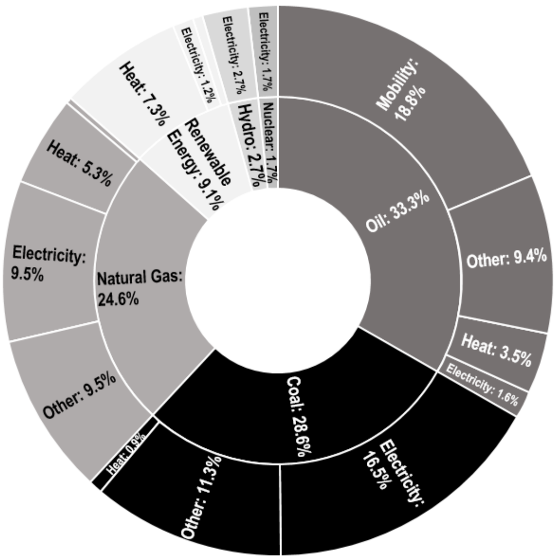


Figure 6. Share of the total global primary energy consumption (151 PWh) by fuel in 2016 and sector of use [1, 2].

The shares of the energy sources by sector indicate, which parts of energy supply will be easily substituted by electricity from power plants based on the use of renewable sources of energy (e.g. electricity from all other sources) and which parts ask for additional technology. For example, roughly 20 % of global primary energy demand is attributable to mobility (almost completely met using oil). To substitute this primary energy component through renewable energies (which predominantly provide electrical energy), formerly separated energy sectors existing within the overall energy system must be increasingly interconnected (e.g. meeting mobility demand through e-mobility). Such sector coupling normally requires further integration technologies like new charging infrastructures or battery storage in cars.

Fig. 7 shows the development of global primary energy demand until the year 2050 (based on [17] and own calculations). Following this analysis, the total energy demand will rise considerably (driven by the factors described above). Most likely, this clearly increasing energy demand will be covered by electricity to a much higher share compared to today (this development tendency is already apparent).

The shown demand for 2050 is treated as a basis for the global population's demand for usable energy (considering conversion efficiencies), that must be met by the energy supply system for each development pathway. While electricity from renewable energies (as well as from nuclear power) is directly treated as primary energy in this paper, this increasing share of electricity consumption within the overall global energy system reduces overall conversion losses from primary to usable energy which in turn reduces overall primary energy demand. This is a reason why the prospected energy demand does not increase as strong as expected based on the currently visible developments also discussed in chapter 2. Another reason for this rather moderate increase of energy demand is the fact that improvements of energy efficiency measures are implied to become more effective. To be compliant with climate goals, overall GHG emissions must be reduced drastically. In 2017 they amounted to roughly 36 Gt/a [18], which would have to be reduced as close to zero as possible until 2050.

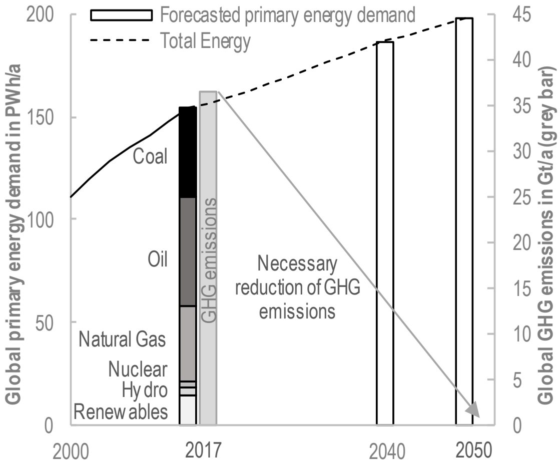


Figure 7. Development of primary energy demand until 2050 (based on [17] and own calculations) and necessary reductions of GHG emissions

4. Energy from renewables – status and developments

The described drivers (chapter 2.1 and 2.2) will significantly impact the expansion of renewable power production capacities in the coming decades. In addition, there will most likely be some further technological development as it has been realized in the past. Hereinafter, a bottom-up discussion about the advancement of the most relevant options to use renewable sources of energy for electricity provision is presented from today's point of view. Already emerging tendencies are carried on and developed for the following years under the assumption that the expansion of renewable power plants is mostly based on economic and not predominantly on environmental and/or political respectively legally controlled decisions.

4.1. Photovoltaics

Economic and technological development of photovoltaics (concerning solar cells / modules as well as system components) has been unexpectedly rapid in the past decade. Photovoltaics systems are becoming increasingly durable, robust, efficient and especially cheaper. Additionally, it is most likely that this development will continue in the coming years.

By the end of 2017, around 385 GW of photovoltaic capacity had been installed worldwide, with more than 90 GW newly added in 2017 alone. For the past four years, photovoltaic has been the world's fastest-growing technology for power generation from renewable energies with a potential electricity generation of around 435 to 635 TWh in 2017 [19, 20]. This development is shown in Fig. 8.

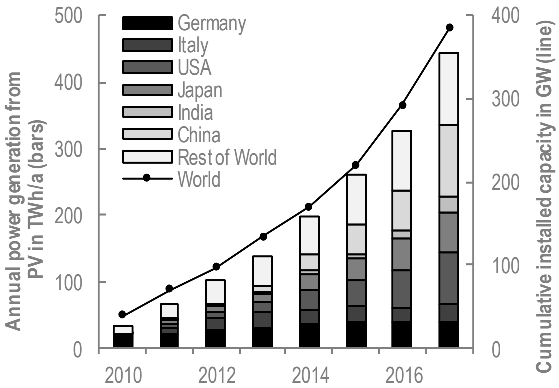


Figure 8. Annual power generation and cumulative installed photovoltaic capacity worldwide [17, 19]

This overall development has led to a significant global expansion of the installed photovoltaic capacity as well as power production. Additionally, the globally available potential for further photovoltaic development does not seem to be limited for the years to come; i.e., a further fast development of a photovoltaic electricity generation is expected with a high probability.

Under consideration of the known development goals for the main active markets till 2022 (adding remaining nations with average past growth rates) a global installed capacity of at least 835 GW is likely. Electricity production would than amount to 900 till 1,300 TWh/a. With an ongoing decline of cost and thus accelerated expansion, installed capacities would rise to roughly 1,600 GW until the year 2030 with a power production between 1,700 and 2,800 TWh/a (the upper value assumes that a rise in global mean capacity factor takes place). In 2050 an installed capacity of over 8,000 GW is quite likely, which would provide roughly 10,800 to 14,800 TWh/a. From today's perspective, it is rather probable, that photovoltaic will become the most important power generation technology based on renewable energies in the future [20].

4.2. Wind Energy

Strong cost reduction for wind power plants has driven a global increase of on- and offshore wind power production in the past years paired with increasingly more robust and reliable wind mills. This is true for onshore application anyway and in recent years more and more valid also for offshore wind parks.

By the end of 2017, around 539 GW of installed capacity of wind turbines were operating worldwide (new capacity in 2017 of around 52 GW) [21], which generated around 1,060 TWh of electricity (Fig. 9). The majority is installed onshore, only about 4 % are installed offshore [20].

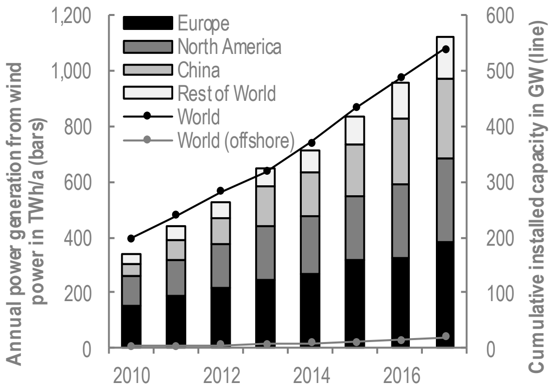


Figure 9. Annual power generation and cumulative installed wind turbines capacity worldwide [17, 20]

Globally there are still many undeveloped sites (onshore as well as offshore) with high average wind speeds, allowing the assumption, that wind power will be further developed in der future. This development goes hand in hand with reduced specific prices. On a global scale there is no real limitation to wind power expansion from today's point of view. Existing national expansion goals add up to around 50 to 60 GW of additional capacity in the next years mostly in Asia, followed by Europe and North America. Until the year 2022, globally installed capacity could potentially amount to 840 GW, providing roughly 1,700 TWh/a of electricity. Offshore wind will only make up a small share of this (about 50 GW). Under the assumption of a continuously progressive expansion, the worldwide installed capacity will amount to 1,400 GW (electricity provision of 3,000 TWh/a) in 2030. Until the year 2050 this capacity can amount to roughly 5,000 GW – implying an ongoing stable expansion. The power production will than sum up to 10,000 to 12,500 TWh/a. In the long run, wind power together with photovoltaic will most likely show the largest market shares of renewable energies [20].

4.3. Hydro power

Hydroelectric power plants are an established technology. Thus, there has not been any major technological development in the recent years and is not foreseeable that there will be any major improvement in the future (the overall efficiency lies between 80 and 90 % and is therefore already quite high). Main disadvantages of hydro power are its dependency on suitable locations and high specific investments. Many suitable locations are already exploited; therefore, the focus will increasingly lie on small hydroelectric power stations – and even there, the potential is foreseeably limited.

In 2017, approximately 1,267 GW of electrical power was installed in hydropower plants (including an addition of 22 GW in 2017), including approximately 153 GW of capacity installed in pumped storage power plants. In total, about 4,185 TWh of electricity were generated in 2017 by hydro power (Fig. 10).

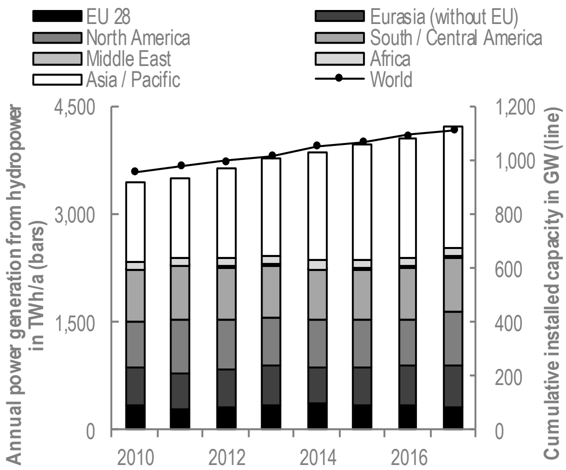


Figure 10. Annual power generation and cumulative installed hydro power capacity worldwide [20]

Nonetheless, multiple hydro power plants are currently planned or constructed, which will lead to an increase of the overall installed capacity in the following years, resulting in a global installed capacity of around 1,400 GW providing roughly 4,600 TWh/a of electricity in 2022. By the year 2030 this capacity will potentially increase to between 1,500 to 1,600 GW (roughly 5,000 TWh/a) [20] and until 2050 global installed hydroelectric power capacity might amount to 2,200 GW with an electricity production of roughly 7,300 TWh/a. The high cost of exploitation of further, more inaccessible sites will be the main obstruction for a stronger advancement of hydro power.

4.4. Biomass

Another renewable energy option is the utilization of biomass. Nowadays, biomass constitutes the largest renewable energy source, mainly due to the use of biomass for cooking and heating in less developed energy systems as well as in industrialized countries especially in the country side. Additionally, modern biomass utilization has increased globally for heat and electricity production in the last decades, as a result of improved and more reliable conversion technologies. An increased use of organic waste materials (e.g. organic waste fraction, organic waste from the food processing industry) adds up to this development. The sustainably usable amount of biomass will nevertheless be a limiting factor for biomass related energy provision.

At the end of 2017, around 112 GW of power plants using solid biofuels were operating, generating a total of 555 TWh of electricity; the largest existing capacities are in the USA (17 GW) [22]. In 2017, about 330 million t of organic waste have been thermally utilized providing between 82 and 110 TWh of electricity. For 2017, it can be assumed that the installed capacity for biogas lies between 19 to 23 GW. With this capacity, a potential power generation of 114 to 137 TWh is possible globally ([20, 23]).

Concerning solid biomass for electricity generation with a continuously increasing global capacity until 2022, roughly 112 GW will be installed with an energy provision of around 460 TWh. Global capacity might rise to 130 GW in 2030 (600 TWh/a) and to 190 GW in 2050 (850 to 950 TWh/a). Biogas production is also increasing, leading to roughly 25 GW of installed capacity in 2022, providing 162 TWh/a of electricity. This capacity will potentially rise to 30 GW till 2030 [20] and 40 GW until 2050 with an electricity production of 230 TWh/a and 300 TWh/a respectively. A possibly larger potential for biomass energy lies in the use for thermal energy supply, which is already partly exploited today and will most likely be an integrative part of future energy systems.

4.5. Other renewables

Geothermal energy and concentrated solar power (CSP) will most likely only play a minor role in renewable power production in the future.

In the year 2017 the global electric capacity of geothermal energy amounted to roughly 13 GW, providing around 73 TWh/a of electricity. Strongly limited high-enthalpy resources make the electricity provision from geothermal energy a niche technology with no major influence on global electricity production in the future. Nevertheless, geothermal energy might be of larger importance concerning heat energy supply (in district heating systems as well as small applications like single family household heat pumps). CSP is even less developed globally. Merely 5 GW of power are installed today. Due to the high cost compared to photovoltaics, the expansion will be drastically limited.

4.6. Summary – current development trends

A bottom-up approach from today's point of view (assuming a commercially driven expansion of renewable energies based on the current frame conditions and price relations) forecasts a quite fast expansion of power production from renewable sources of energy. Fig. 11 shows this expected development of the electricity provision namely for wind power, photovoltaics, hydro power and others for the coming decades (Fig. 11 does only include electricity from renewables).

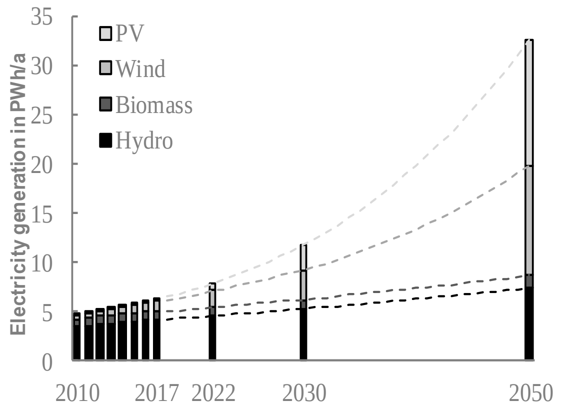


Figure 11. Forecast of global renewable electricity production [17]

4.7. Comparison

Even though this development constitutes a strong increase compared to the development in recent years, renewable energies for electricity provision will only contribute with a rather small part of overall electricity provision according to the development scenario outlined above.

As another consequence of a raising share of electricity generation especially from renewable sources of energy, also other energy sectors (e.g., heat sector, mobility sector) will be increasingly electrified. In this context electricity will most likely also be used in combination with other energy sources (like low enthalpy heat for heat pumps) to over-proportionately increase the share of renewables within the overall energy supply system; but for technological reasons this option is limited to certain markets. This development is very much powered and pushed by GHG reduction efforts in some countries and/or regions.

Based on the forecasted energy demand by fuel for the year 2040 according to [3] including the bottom-up forecast of renewable energies realized above, the overall composition of global energy supply for the year 2050 can be roughly estimated. For this assessment, the following assumptions are made.

- Growth rates by fuel between 2040 and 2050 are roughly half of those between 2016 and 2040 as assumed in [3].
- Often in such assessments, electricity from nuclear, hydro and other renewable sources of energy are converted back to primary energy by calculating the primary energy equivalent of an average thermal power plant (with an efficiency of 38 %) needed to supply the same amount of electricity; this is for example also realized in [3]. In contrast to that, here, electricity from those technologies using e.g. renewable sources of energy is directly treated as primary energy (which does in fact lead to an underrepresentation of renewable energies compared to the approach realized in [3]).
- Differences between the assumed development for hydro power and renewables based on [3] and those described in this chapter (i.e., the development shown in Fig. 11) are included in the overall energy supply forecast for the year 2050 (i.e. here, the amount of energy from renewable sources is roughly 25 % higher in comparison to [3]). The amount of energy by which the forecasts of Fig. 11 exceeds the forecasts in [3] is accounted for by reducing the share of fossil fuel (compared to [3]). The reduction is determined by calculating the primary energy a thermal power plant with an efficiency of 38 % would need to supply the same amount of electricity. This amount is than subtracted from the fossil primary energy components.
- An increasing electricity provision from renewable sources will primarily substitute electricity production formerly accomplished by conventional power plants (i.e., driven by fossil fuel energy). In this context primarily oil and coal fired power plants are replaced. Power production from hydroelectric power plants is assumed to directly substitute electricity provided by conventional power plants. For other renewable power provision, which supposedly mostly relies on fluctuating production from wind energy and solar power, an increasing need for energy storage will reduce its capability to replace demand-driven conventional power production. To account for this reduction, renewable electricity is assumed to replace fossil-based electricity production by a factor of 0.75 (e.g., to replace 0.75 MWh of electricity from demand-driven power plants 1 MWh of electricity from fluctuating renewable energies must be provided).

This calculation results in the shares shown in Fig. 12. Overall energy demand amounts to roughly 189 PWh/a (considering an overall tendency towards increasing electrification). Out of these around 42.2 PWh/a are provided based on renewable sources of energy including hydro power. The main source of primary energy is now natural gas with 59 PWh/a, because of the massive build-up of the use of renewable energy, both less oil and coal are used; thus, their shares within the overall electricity provision system are declining.

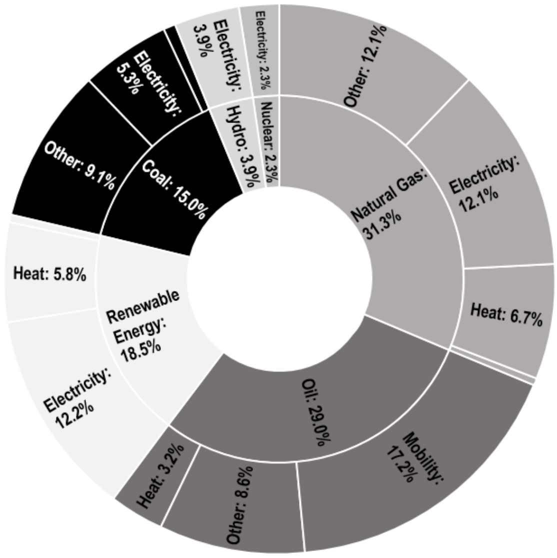


Figure 12. Forecasted shares of total global primary energy consumption (189 PWh) in 2050 (bottom up)

Against this background, based on such a bottom-up perspective the total primary energy demand estimated by 2050 (chapter 3) can be covered by roughly 22 % by the described utilization of renewable sources of energy.

Even though this development assumes a quite strong increase of renewable power production, there is still a large deficit between the resulting energy supply and an energy provision system that is conform with the globally accepted 1.5 °C-goal. If the use of energy sources is structured as shown in Fig. 12, there will be energy related GHG-emissions of roughly 36 Gt/a in 2050 emitted, which is roughly the amount of today's global emissions [18]. This is again shown in Fig. 13. The increased energy supply from renewable energy sources is almost compensated by growth, leading to mostly constant GHG emissions. Overall, primary energy consumption is slightly lower than the forecasted amount, due to a higher degree of electrification.

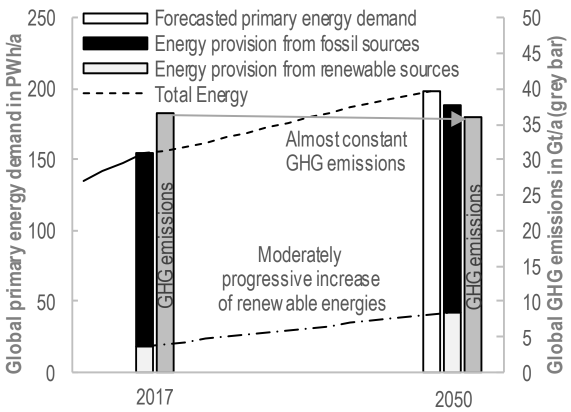


Figure 13. Bottom-up development of renewable energies and the overall energy supply system

The amount of GHG-emissions in that scenario will quite clearly exceed the emissions for the year 2050 according to the latest GHG abatement agreement (e.g. Paris climate agreement); i.e., this scenario is not compliant with global climate goals. Despite the assumed strong development of the use of renewable energies, countries will clearly fail to follow the necessary reduction paths depicted in Fig. 1. To achieve those goals, a much more rapid expansion would be needed.

5. Necessary global expansion of renewable energies

In order to be compliant with the Paris climate agreement, it is necessary, to significantly reduce GHG emissions in the years to come. The Intergovernmental Panel on Climate Change (IPCC) stated in 2018, that according to limit the rise of mean global temperature to 1.5 °C with a probability of 66 %, the overall global GHG emissions yet to be made should not exceed 420 billion t [24]. If the goal is to limit the increase to 2 °C, the amount of allowed GHG emissions slightly increases. But nevertheless, a net zero emission status has to be reached rather early than late. Thus, numerous GHG-reduction pathways have been developed. One of these pathways assumes halving the CO_2 emissions every decade from 2020 to 2050 [25]. For 2050 this results in an allowed amount of 5 Gt/a of GHG to be emitted. Assuming that all emissions not related to fossil fuels are either reduced through a change of process or captured (carbon capture and storage, CSS), these 5 Gt/a would be the absolute limit for emissions from the use of fossil fuels in the year 2050.

5.1. Necessary development

The trends described in the previous chapter assume a mostly cost-driven (commercially viable) expansion of an increased use of renewable energies based on current frame conditions and the relative cost differences valid now. In opposite to this approach in the following, a top-down discussion about the necessary global expansion of the use of renewable sources of energy is given based on the climate goals described above. Therefore, this chapter presents a development of the increased use of renewable energies under the precondition that the development fulfils the valid reduction goals and thus the advancements are mostly driven by environmental aspects. Additionally, it is assumed that such a development is possible without additionally reducing the overall energy demand (i.e., the economic development is assumed to be the same as already defined in chapters 3 and 4).

To reach an energy supply compliant with global climate goals, the use of fossil fuels as an energy source would have to be drastically reduced. To estimate a necessary global energy supply structure, that would fulfil this strong reduction of fossil fuels, a broad variety of assumptions must be made (the following elaboration, therefore, must be explicitly understood as a rough estimation).

- The still allowed GHG-emissions of 5 Gt/a in 2050 are divided equally between natural gas (equal shares for electricity, heat and other) and crude oil (equal shares for mobility and other).
- The use of coal is out-phased and substituted by electricity provided by renewables sources of energy.
- Nuclear power production stays constant in comparison to the previous scenario.
- Hydroelectric power generation additionally increases by 25 % compared to the forecast in Fig. 11. Additional hydroelectric power generation is used to directly substitute electricity from fossil fuel-based power plants.
- Including higher shares of renewables demands for more frequent use of energy storage. This use again is assumed to have an efficiency of 75 % (estimated average across all storage technologies available and under operation in 2050). Electricity from renewables other than hydro power therefore substitutes electricity from fossil-based power plants by a factor of 0.75.
- Using renewables to meet mobility demand relates to a shift of mobility structures. The still usable oil can be utilized in aviation, shipping and heavy-duty vehicles. The demand (for these types of transport) which goes beyond that, is met through the generation of synthetic fuels. It is assumed, that 20 % of the remaining mobility demand is met by such fuels with an efficiency for the conversion of electrical energy to chemical energy of 35 %. The additional 80 % are met by electro-mobility. Direct electric mobility needs less primary energy as conventional engine driven mobility, because in combustion engines large shares of the used primary energy (mostly oil) are lost due to the engine's efficiency (which is assumed to be 33 %).
- Including renewable sources of energy within the heating sector leads to additional implications. It is assumed that 60 % of the heat demand will be met by using heat pump technology (with an overall coefficient of power of 4), 25 % is met by generating and using synthetic gas (with a process efficiency for gas production of 50 %) and 15 % of the overall heat demand is met by using electricity directly.
- Meeting the demand for "Other" (Fig. 6 and Fig. 12) through electricity from renewable power plants is assumed to be one-to-one. Due to the lack of a definable efficiency for conversion, this constitutes the best guess.

These assumptions result in an overall renewable power production (mostly wind and photovoltaics) of nearly 126 PWh/a (including renewable heat and mobility shares) related to the year 2050. Due to the fact, that renewable power would than mostly rely on producing electricity, the global share of electrical energy as an energy carrier would increase to roughly 82 % of which nearly 88 % would be based on renewable energies (other than hydroelectric power generation). Fig. 14 shows the overall shares of primary energy by fuel and sector.

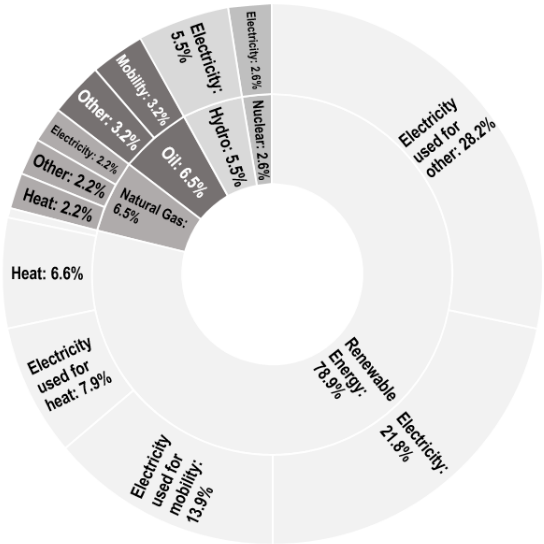


Figure 14. Shares of total global primary energy consumption (167 PWh) in 2050 needed to be compliant with global climate goals (top-down)

To supply such a massive amount of energy from renewable power plants, global installed capacities would have to amount to roughly 33 TW of photovoltaic systems (providing 66 PWh/a), 19.2 TW of wind power plants (providing 48 PWh/a) and 1.2 TW of other renewable electricity generation facilities providing 6 PWh/a (assuming global mean full load hours of 2 000 h/a for photovoltaics, 2 500 h/a for wind power and 5 000 h/a for power provision from other renewable sources of energy). These capacity calculations assume that photovoltaic power production exceeds that of wind power plants so that renewable electricity generation is divided upon photovoltaics, wind power and others in shares of 55 %, 40 % and 5 %.

The necessary development is shown in Fig. 15. Electricity production from power plants based on renewable energies must be drastically increased to reach the described shares of roughly 81 %. Consequential, the GHG emissions drop substantially, and reach 5 Gt/a.

Such a massive expansion of the use of renewable sources of energy would only be possible, if in the near future a kind of collective global realization of climate change threats leads to far-reaching political decisions, additionally boosting the expansion of the use of renewable energies and also the expansion of technologies that would enable the amount of sector coupling necessary to allow renewable based electricity to replace fossil fuels in different sectors (the additional expansion of nuclear power is not considered here, due to its high cost, high riskiness and unresolved issues concerning nuclear waste repositories; additionally, possible technological leaps – e.g. getting nuclear fusion to work – are also not considered as a realistic option for those time horizons). Additionally, the legal frame condition must support such a development.

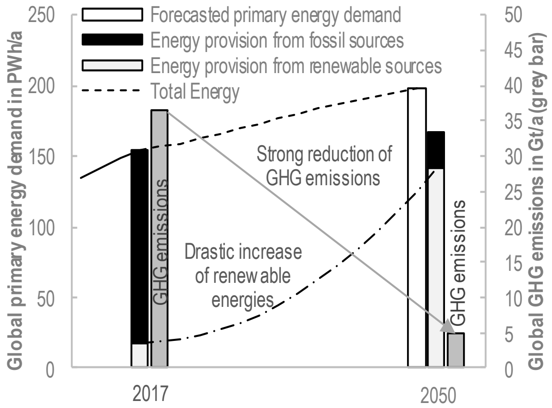


Figure 15. Top-down development of renewable energies and the overall energy supply system

5.2. Consequences

Even though, it is not the scope of this paper to decide, how realistic such a development is, such an advanced expansion would lead to extensive implications for global energy supply system structures. Therefore, below an overview to different challenges arising from such strongly increasing shares of renewable power production are discussed.

Electricity system. The expansion of renewable power plants will increasingly challenge electricity supply systems globally. Thereby it is mostly irrelevant, whether the specific energy system is developing (i.e. energy demand is rising and there is a net increase of power production) or transforming (i.e. there is no considerable change in overall energy demand, but the used energy sources are switched from fossil to renewable). Anyway, the resulting system will have to include several aspects, to cope with increasing shares of (fluctuating) renewable energies.

The power production of photovoltaic and wind power plants is characterized by strong fluctuations and additionally relies strongly on the availability of solar radiation and high wind speeds, respectively. Therefore, it will become increasingly important for electricity systems to flexibly react upon those increasingly prevailing fluctuations and availability uncertainties. This leads to several implications, that will make future electricity systems differ from today's. Some of them are discussed below.

- The overall power supply system needs to become increasingly flexible. For example, the thermal power plants must operate with higher load gradients and a wider range of operational states. Such an improvement of the flexibility of the overall power production can be reached by different approaches. Concerning renewable power production, among others it is very important to increase forecast quality even further to secure the highest possible plannability of renewable power output. Additionally, power plants providing the residual load (i.e. the load not covered by fluctuation sources like wind and solar) must be upgraded and new-built plants have to be prepared for flexible power production implicitly. This will most likely result in operation with higher load gradients and fewer full load hours [26]. Due the fact that these thermal power plants are not operated by fossil fuel energy any more this is true for conversion units using e.g. demolition and residual wood, organic waste fractions from the food processing industry, domestic waste and other similar fuels.

- Fluctuations must be spatially and temporally balanced by either increased transmission capability (i.e., expansion of electricity grids related to capacity and covering) and/or large-scale storage systems (i.e., expansion of system wide storage capacity). Spatial balancing / compensation can be achieved through strong electricity grids with high transmission capabilities and a high degree of crosslinking. Those structures allow for energy harnessing in high potential regions (e.g., coastal areas for high wind power production) independent of the location of high demand regions; but the installation of such an infrastructure is very costly. For energy supply systems in countries with developing demand structures, it might also be possible, to locate high consumption facilities close to high yield areas, but all in all, enough transmission capabilities seem indispensable; this is basically a precondition to close the gap between areas most promising for energy provision and the huge demand centers. Temporal balancing is another important issue. Diurnal patterns of solar power production as well as the stochastic behavior of wind energy supply demands for the possibility to store energy for a timely shifted use. Several storage technologies are already state of the art like pumped-storage power plants and electrochemical battery storages; beside this, these concepts are under ongoing improvement and additional approaches are researched in depth (e.g., compressed air storages). The main focus is to find cost efficient concepts with the ability to store sufficient amounts of energy for a certain period of time. Thus, storage systems – either "real" storage like pumped hydropower stations or battery systems or "virtual" storage concepts like virtual power plants or sector coupling with e.g. the heat sector – are needed necessarily. Such storage systems can be realized on the small scale in clearly decentralized structures as well as within large scale systems helping to stabilize the overall supply system; most likely both approaches are needed within a fully integrated system expected in the years to come. Strong grid structures reduce the need for energy storage, because weather effects partly compensate over larger areas. Therefore, spatial and temporal compensation of fluctuation should be considered in combination.
- The demand side must become increasingly flexible; a key measure to achieve this goal is that the customer can benefit from the strongly volatile prices to be expected due to this development. The overarching goal is to make the demand increasingly more flexible and controllable. There are numerous options. For example, the operation of selected time-flexible industrial processes could follow the power production from renewable sources (i.e. increasing electricity demand while electricity provision from renewable energies is high and reducing it, while production is low). Additionally, some energy consuming applications show intrinsic storage effects (e.g., electro thermal applications with large thermal masses like houses with large surface area heating or cooling devices), which also enable demand flexibility. Thus, to completely reduce GHG emissions using mainly wind and photovoltaic systems, the rest of the energy system will have to be considered too.

Overall energy system. In addition to the implications for the electricity system, there will be some major implications for the overall energy system. Because most conversion systems using renewable sources of energy provide electrical energy as an output, it will be increasingly important to couple the power production system to other energy systems respectively other energy sectors (i.e. sector coupling). If for GHG-reduction reasons large amounts of oil and gas are substituted by electrical energy from renewable sources in the years to come, the demand that was formerly met by those fossil fuels must be met by electricity in the future (this was the basis of the system adaptation for high renewable shares in Fig. 14). One example for this kind of sector coupling is power-to-heat – the transformation of electrical energy to thermal energy in order to fulfil heat demand especially during winter time. This can be achieved by utilizing heat-pumps, which use the energy of low temperature sources (e.g., air, soil, groundwater) and increase its temperature to a technically useable level with the help of electrical energy. Heat pumps can supply thermal energy several factors larger than the input electricity amount and thus constitute a highly efficient way of sector coupling. Due to thermal storage effects of buildings such heat pump systems additionally help to flexibly react to renewable power production fluctuations. Power-to-heat can additionally be implemented as highly flexible direct electric heating, to solely use electricity for heating purposes, when a surplus of renewable energy would otherwise be prevailing.

Another way of sector coupling, which directly substitutes fossil fuels, lies in the power-to-gas and power-to-liquid technology. Both processes start with using electricity and water to produce hydrogen by electrolysis. This hydrogen can either be used directly or transformed to other forms of energy carriers. A direct use (instantaneously or after storage) would constitute the most efficient way, but there are practical issues (storability, existing non hydrogen-based infrastructure in many parts of the world) that hinder the implementation of a pure hydrogen infrastructure (regions with still developing energy demand might be able to directly implement a more hydrogen-based energy supply, provided reasonable costs of implementation and therefore reduced overall system cost). To evade these problems, additional processes are used to turn hydrogen into e.g. methane (power-to-gas) or a liquid energy carrier (power-to-liquid) by adding CO_2. After such a further treatment they can be used in already existing infrastructure and with existing technology. The liquid fuels can fulfil all specifications of currently used fuels and are therefore able to directly substitute those fossil-based fuels and be used in e.g. aviation. Both – power-to-gas and the power-to-liquid – processes show rather low overall efficiencies. Therefore, a direct use of electricity should be the preferred solution. E-mobility offers that kind of solution.

By either being directly grid connected (rail-based transport) or by using mobile batteries (electricity-based individual mobility) that feed an electric motor, electricity can directly power the engine of vehicles and therefore help to reduce the demand for fossil fuels within the transport sector. For global scale implementation, a reliable and sustainable battery technology would therefore be needed. Nevertheless, there might always be some cases, in which liquid fuels with a high energy density are needed, which might not be possible for battery systems (e.g. aviation). Therefore, it is most likely that no single technology will help to master the stated challenges but rather a mixture of different approaches and technologies.

6. Conclusion

Trends in global population and prosperity development will most likely lead to a significant further increase of global primary energy demand in the years to come. Having in mind the increasingly drastic effects of global climate change, it is necessary, that this rising primary energy demand will be met by the utilization of basically GHG-emission free renewable energies. In this context, two different scenarios for the development of the global primary energy supply system – especially focusing on renewable energies – have been elaborated and discussed.

As a basis for these development scenarios, the current status of primary energy demand (2016) was discussed in chapter 3 and a forecast for the year 2050 has been developed. Based on this, first, a bottom-up approach was presented in chapter 4, progressively extrapolating the current trends and developments on providing useable energy based on renewable sources of energy until 2050. Due to falling prices, photovoltaics and wind power will be the dominant technologies of renewable energies in the decades to come. However, the main energy carriers will be natural gas and crude oil in this scenario (providing roughly 60 % of the global primary energy); i.e. most of the the world's energy demand will be served by fossil fuels (roughly 75 % in total). Thus, as a matter of fact, based on the currently ongoing developments the challenging GHG reduction goals cannot be met by the year 2050.

In contrast to the bottom-up approach presented in chapter 4, a top-down scenario has been developed and discussed in chapter 5. To be compliant with the overall goals of the Paris climate agreement a total of roughly 5 Gt/a of GHG emissions are still allowed in the year 2050; this overarching goal must be achieved within this scenario. This goal in view, a new share of renewable energy carriers has been calculated, summing up to roughly 84 % of total primary energy by 2050. Although, the global population's demand for usable energy was kept constant, this scenario implied a reduction of the global primary energy demand due to the direct use of electricity from renewable energy sources (and thus a reduction of overall energy conversion losses; i.e., electricity from e.g. wind power and solar radiation is treated directly as primary energy).

Both pathways (bottom-up as well as the top-down approach) are again shown in Fig 16. Clearly recognizable are the different developments for the expansion of energy provision from renewable sources (dashed/dotted lines). There it becomes obvious that the top-down approach enforced basically by the Paris agreement would require a very strong expansion of power plants based on renewable energies within the next decades. And to achieve such a development much stronger political measures need to be implemented as already in force today – and such a development needs to be implemented by basically all countries globally.

Even though, these development and expansion paths are merely estimations of possible (and necessary) developments for the future and do not necessarily have to turn out exactly as described, they show that on the one hand it is quite obvious that some major transformations are about to occur in global energy systems; on the other hand, the estimations suggest, that a substantial increase in efforts from politics and industry – as well as from the customers – concerning the expansion of renewable energies would be needed globally, to be able to meet global climate goals.

High income countries with already high energy demand will have to successively replace power plants operating with fossil fuels by power supply technologies based on renewable energies. Simultaneously, transmission capabilities must be strengthened, storage systems must be installed, demand must become more flexible and energy sectors must become increasingly intertwined. Low income countries will have to align growth of their energy system with being climate-neutral and must incorporate mostly (or only) technology based on renewable energies in their growing energy supply structures.

With an increasing share of fluctuating electricity related to the overall electricity provision more and more technical and organizational measures need to be implemented to allow for a stable electricity and thus also energy supply system. Fig. 17 illustrates the necessary and partly already ongoing transformation processes. The historical structure of electricity supply from the large power plants down to consumers will increasingly develop into a more and more integrated energy provision structure. Industry and private end consumers are increasingly interacting with the distribution structures and the conventional power plants based on fossil fuels are being successively reduced while more and more power plants based on renewable energies are installed. In addition to that, due to more and more volatile renewable energy infeed, the historical load following operation is not further suitable, a flexibilization of electricity demand and integration of further flexibility measures to respond to increasingly fluctuating power generation is needed.

The historically independent view of electricity, heating and mobility markets must transform into an integrative optimization of all (sub-) energy markets in the entire energy system, taking advantage of possible cross-sector synergy effects. If the perspective of optimizing the energy system, sets in the short to medium term – and a practical alternative to this is not recognizable – solutions for the resulting challenges need to be found as part of this paradigm shift; first ideas and approaches are already visible. Thus, decisions to be taken today about the further development of the electricity system should take care about this development and the resulting necessities to maximize the exploitation of possible synergy effects in the future.

From a global perspective, about 80 % of the primary energy demand must be met by renewable energies in 2050 (which will primarily be supplied as electricity) to be compliant with the climate goals agreed on within the Paris agreement. In addition to being used as such, this electricity must be integrated across the various energy sectors for heating, mobility and other uses, to allow for widespread decarbonization (concerning fossil carbon) in order to achieve the set climate goals. For these paths to be followed accordingly a coordinated global effort from both politics and industry would be necessary; this has to be supported by the customers. Mastering these overall tasks is truly challenging but nevertheless not impossible.

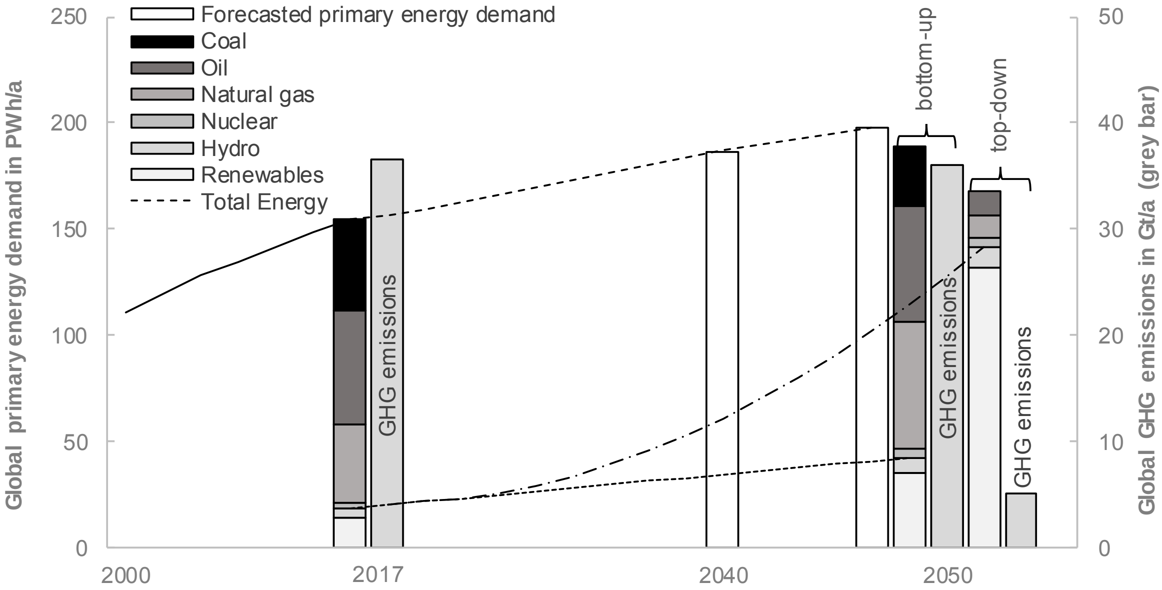


Figure 16. Possible development of primary energy demand until 2050 for bottom-up and top-down scenarios for renewable energy expansion ([17] and own calculation)

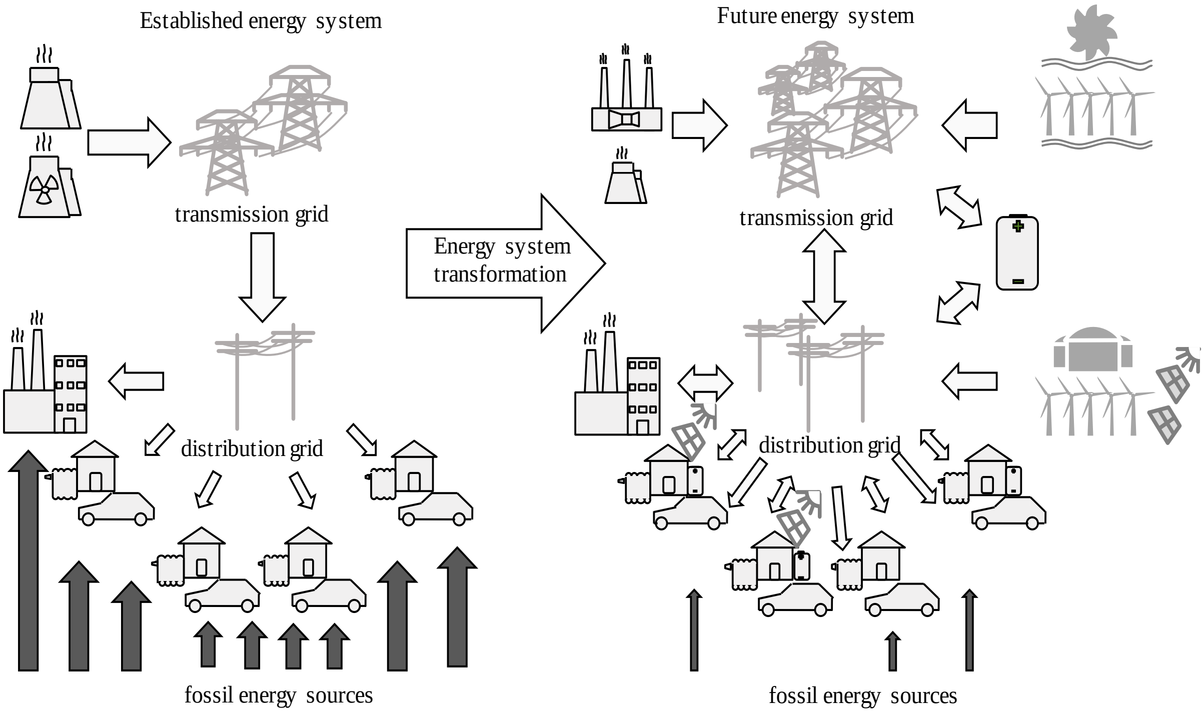


Figure 17. Energy system transformation (light arrows: flow of electricity, dark arrows: flow of fossil fuels, based on [27])

References

[1] BP p.l.c: BP Statistical Review of World Energy 2018. London, 2018.

[2] REN21: Renewables 2018 Global Status Report. Paris, 2018.

[3] BP p.l.c: BP Energy Outlook. 2018 edition. London, 2018.

[4] Carbon Dioxide Information Analysis Center: Fossil-Fuel CO2 Emissions by Nation. http://cdiac.ess-dive.lbl.gov/ftp/ndp030/CSV-FILES/nation.1751_2014.csv (Abruf: 16.01.2018).

[5] Gapminder Foundation: Based on free material from GAPMINDER.ORG, CC-BY LICENSE. www.gapminder.org (Abruf: 01.10.2018).

[6] National Centers for Environmental Information: Global Climate Report - Annual 2017. Ten Warmest Years (1880–2017). https://www.ncdc.noaa.gov/sotc/global/201713 (Abruf: 20.09.2018).

[7] Frankfurt School-UNEP Centre: Global Trends in Renewable Energy Investment 2018. Frankfurt am Main, 2018.

[8] IEA - International Energy Agency: World Energy Outlook 2018. Executive Summary, 2018.

[9] McKinsey Energy Insight: Global Energy Perspective 2019:. Reference Case Summary, 2019.

[10] EIA: International Energy Outlook 2019. With projections to 2050, U.S. Energy Information Administration, 2019.

[11] IRENA: Global Energy Transformation:. A Roadmap to 2050, International Renewable Energy Agency. Abu Dhabi, 2018.

[12] Edenhofer, O.: Renewable energy sources and climate change mitigation. Summary for policymakers and technical summary : special report of the intergovernmental panel on climate change. Cambridge University Press, New York, 2011.

[13] United Nations DESA/Population Devision: World Population Prospects 2017. https://population.un.org/wpp/Download/Standard/Population/ (Abruf: 18.09.2018).

[14] Bundesministerium für Wirtschaft und Energie: Primärenergieverbrauch nach Ländern und Regionen. https://www.bmwi.de/Redaktion/DE/Binaer/Energiedaten/Internationaler-Energiemarkt/energiedaten-int-energiemarkt-01-xls.xlsx?__blob=publicationFile&v=19 (Abruf: 18.09.2018).

[15] The World Bank: Fertility rate, total (births per woman) | Data. https://data.worldbank.org/indicator/SP.DYN.TFRT.IN (Abruf: 24.01.2019).

[16] The World Bank: World Bank national accounts data, and OECD National Accounts data files. data.worldbank.org (Abruf: 04.10.2018).

[17] BP p.l.c: Statistical Review of World Energy. All data, 1965-2017. https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy/downloads.html (Abruf: 17.01.2019).

[18] Global Carbon Project: Global Carbon Atlas. http://www.globalcarbonatlas.org/en/CO2-emissions (Abruf: 21.01.2019).

[19] International Energy Agency (IEA): Solar Energy. Solar Energy Data. https://www.irena.org/solar (Abruf: 01.02.2019).

[20] Witt, J.; Magdowski, A.; Janczik, S. et al.: Erneuerbare Energien Weltweit. Globaler Stand 2017. BWK 70 (201), 7/8, S. 15–35.

[21] Global Wind Energy Council: Global Wind Report. Annual Market Update 2017. Brussels, 2018.

[22] Renewable Energy Policy Network: Renewables 2018. Global Status Report. Paris, 2018.

[23] ecoprog GmbH: Waste to Energy 2017/2018. Technologies, plants, projects, players and backgrounds of the global thermal waste treatment business. Cologne, 2017.

[24] J. Rogelj, D. Shindell, K. Jiang, S. Fifita, P. Forster, V. Ginzburg, C. Handa, H. Kheshgi, S. Kobayashi, E. Kriegler, L. Mundaca, R. Séférian, M. V. Vilariño: Mitigation pathways compatible with 1.5°C in the context of sustainable development. In: IPCC (Hrsg.): Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty, S. 93–174.

[25] Rockström, J.; Gaffney, O.; Rogelj, J. et al.: A roadmap for rapid decarbonization. Science (New York, N.Y.) 355 (2017), Heft 6331, S. 1269–1271.

[26] Enervis energy advisors GmbH: Der „ideale Kraftwerkspark" der Zukunft. Flexibel, klimafreundlich, kosteneffizient - Maßstab für einen optimierten Entwicklungspfad der Energieversorgung bis 2040, 2014.

[27] Böttcher, J. (Hrsg.): Green Banking. Realizing Renewable Energy Projects. De Gruyter Oldenbourg, Berlin, in press.

المملكة الأردنية الهاشمية

الجامعة الهاشمية

المجلة الأردنية

للهندسة الميكانيكية والصناعية

JJMIE

مجلة علمية عالمية محكمة

تصدر بدعم من صندوق البحث العلمي

http://jjmie.hu.edu.jo/

ISSN 1995-6665

المجلة الأردنية للهندسة الميكانيكية والصناعية

المجلة الأردنية للهندسة الميكانيكية والصناعية: مجلة علمية عالمية محكمة تصدرعن الجامعة الهاشمية بالتعاون مع صندوق دعم البحث العلمي والابتكار- وزارة التعليم العالي والبحث العلمي في الأردن.

ترسل البحوث إلى العنوان التالي

رئيس تحرير المجلة الأردنية للهندسة الميكانيكية و الصناعية

الجامعة الهاشمية
كلية الهندسة
قسم الهندسة الميكانيكية
الزرقاء - الأردن

هاتف : 3903333 5 00962 فرعي 4147

Email: jjmie@hu.edu.jo

Website: www.jjmie.hu.edu.jo