THE 7TH INTERNATIONAL WORKSHOP ON SIMULATION FOR ENERGY, SUSTAINABLE DEVELOPMENT & ENVIRONMENT

SEPTEMBER 18 – 20, 2019
LISBON, PORTUGAL

EDITED BY
AGOSTINO BRUZZONE
JANOS SEBESTYEN JANOSY
LETIZIA NICOLETTI
GREGORY ZACHAREWICZ

PRINTED IN RENDE (CS), ITALY, SEPTEMBER 2019

ISBN 978-88-85741-38-6 (Paperback)
ISBN 978-88-85741-37-9 (PDF)

ISBN 978-88-85741-38-6 (Paperback)
ISBN 978-88-85741-37-9 (PDF)

THE 7TH INTERNATIONAL WORKSHOP ON SIMULATION FOR ENERGY, SUSTAINABLE DEVELOPMENT & ENVIROMENT, SESDE 2019

SEPTEMBER 18-20, 2019

LISBON, PORTUGAL

ORGANIZED BY

DIME – UNIVERSITY OF GENOA

LIOPHANT SIMULATION

SIMULATION TEAM

IMCS – INTERNATIONAL MEDITERRANEAN & LATIN AMERICAN COUNCIL OF SIMULATION

DIMEG, UNIVERSITY OF CALABRIA

MSC-LES, MODELING & SIMULATION CENTER, LABORATORY OF ENTERPRISE SOLUTIONS

HUNGARIAN ACADEMY OF SCIENCES CENTRE FOR ENERGY RESEARCH

AUTONOMOUS UNIVERSITY OF BARCELONA

MODELING AND SIMULATION CENTER OF EXCELLENCE (MSCOE)

LATVIAN SIMULATION CENTER – RIGA TECHNICAL UNIVERSITY

LOGISIM

LSIS – LABORATOIRE DES SCIENCES DE L'INFORMATION ET DES SYSTEMES

MIMOS – MOVIMENTO ITALIANO MODELLAZIONE E SIMULAZIONE

MITIM PERUGIA CENTER – UNIVERSITY OF PERUGIA

BRASILIAN SIMULATION CENTER, LAMCE-COPPE-UFRJ

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

MITIM - McLeod Institute of Technology and Interoperable Modeling and Simulation – Genoa Center

M&SNet - McLeod Modeling and Simulation Network

Latvian Simulation Society

Ecole Superieure d'Ingenierie en Sciences Appliquees

Facultad de Ciencias Exactas. Ingegneria y Agrimensura

University of La Laguna

CIFASIS: CONICET-UNR-UPCAM

INSTICC - Institute for Systems and Technologies of Information, Control and Communication

National Russian Simulation Society

CEA - IFAC

University of Bordeaux

University of Cyprus

Dutch Benelux Simulation Society

University of Minho

I3M 2019 Industrial Sponsors

CAL-TEK SRL

LIOTECH LTD

MAST SRL

SIM-4-FUTURE

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

I3M 2019 MEDIA PARTNERS

INDERSCIENCE PUBLISHERS – INTERNATIONAL JOURNAL OF SIMULATION AND PROCESS MODELING

INDERSCIENCE PUBLISHERS – INTERNATIONAL JOURNAL OF SERVICE AND COMPUTING ORIENTED MANUFACTURING

IGI GLOBAL – INTERNATIONAL JOURNAL OF PRIVACY AND HEALTH INFORMATION MANAGEMENT (IJPHIM)

HALLDALE MEDIA GROUP: THE MILITARY SIMULATION AND TRAINING MAGAZINE

HALLDALE MEDIA GROUP: THE JOURNAL FOR HEALTHCARE EDUCATION, SIMULATION AND TRAINING

SAGE
SIMULATION TRANSACTION OF SCS

DE GRUYTER
INTERNATIONAL JOURNAL OF FOOD ENGINEERING

MDPI - SUSTAINABILITY

EUROMERCI: THE ITALIAN MONTHLY LOGISTICS JOURNAL

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

EDITORS

AGOSTINO BRUZZONE
MITIM-DIME, UNIVERSITY OF GENOA, ITALY
agostino@itim.unige.it

JANOS SEBESTYEN JANOSY
MTA EK CENTRE FOR ENERGY RESEARCH HUNGARIAN ACADEMY OF SCIENCES, HUNGARY
Janos.S.Janosy@energia.mta.hu

LETIZIA NICOLETTI
CAL-TEK SRL
l.nicoletti@cal-tek.eu

GREGORY ZACHAREWICZ
IMS UNIVERSITÉ BORDEAUX 1, FRANCE
Gregory.zacharewicz@ims-bordeaux.fr

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

THE INTERNATIONAL MULTIDISCIPLINARY MODELING AND SIMULATION MULTICONFERENCE, I3M 2019

GENERAL CO-CHAIRS

AGOSTINO BRUZZONE, *MITIM DIME, UNIVERSITY OF GENOA, ITALY*
MIQUEL ANGEL PIERA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*

PROGRAM CO-CHAIRS

FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
YURY MERKURYEV, *RIGA TECHNICAL UNIVERSITY, LATVIA*

THE 7TH INTERNATIONAL WORKSHOP ON SIMULATION FOR ENERGY, SUSTAINABLE DEVELOPMENT & ENVIRONMENT, SESDE 2019

GENERAL CO- CHAIRS

JANOS SEBESTYEN JANOSY, *MTA EK CENTRE FOR ENERGY RESEARCH HUNGARIAN ACADEMY OF SCIENCES, HUNGARY*
GREGORY ZACHAREWICZ, *IMS UNIVERSITÉ BORDEAUX 1, FRANCE*

PROGRAM CHAIR

LETIZIA NICOLETTI, CAL-TEK SRL, ITALY

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

SESDE 2019 International Program Committee

Michael Affenzeller, *Upper Austria UAS, Austria*
Matteo Agresta, *University of Genoa, Italy*
Emre Artun, *Middle East Technical University, Cyprus*
Agostino Bruzzone, *University of Genoa, Italy*
Daniel Buchbinder, *Alterna, Guatemala*
Stephan Hutterer, *Upper Austria University of Applied Sciences, Austria*
Janos Sebestyen Janosy, *MTA EK Centre for Energy Research Hungarian Academy of Sciences, Hungary*
Massimo La Scala, Polytechnic University of Bari, Italy
Luis Le Moyne, *University of Bourgogne, France*
Arnis Lektauers, *Riga Technical University, Latvia*
Francesco Longo, *University of Calabria, Italy*
Marina Massei, *University of Genoa, Italy*
Letizia Nicoletti, *Cal-Tek Srl, Italy*
Antonio Padovano, *University of Calabria, Italy*
M.R. Riazi, *University of Kuwait , Kuwait*
Gregory Zacharewicz, *IMS Université Bordeaux 1, France*

Tracks and Workshop Chairs

Transportations
Co-Chairs: Arnis Lektauers, Riga Technical University. Latvia; Luis Le Moyne, University of Bourgogne, France

Smart Grid Planning, Optimization and Control - SPOC
Chairs: Michael Affenzeller, Upper Austria University of Applied Sciences, Austria; Stephan Hutterer, Upper Austria University of Applied Sciences, Austria

Sustainable Land Use and Transport Adopting M&S
Chair: Daniel Buchbinder, Alterna, Guatemala

Consumerism
Chair: Daniel Buchbinder, Alterna, Guatemala

Sustainability
Chair: Daniel Buchbinder, Alterna, Guatemala

Simulation and Modeling in the Hydrocarbon Industry
Chairs: M. R. Riazi, Kuwait University (Kuwait); Emre Artun, Middle East Technical University, Northern Cyprus Campus, Cyprus

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

CHAIRS' MESSAGE

Nowadays the greatest challenge of the mankind: the climate change. The energy industry must be re-structured in order to accommodate the new renewable but unpredictable energy sources. It needs a lot of huge investments well before the actual ones paid back the finances. Things are proceeding too slowly. We have to change our lifestyle, our consumption habits to be more environmental-friendly. Actual transportation methods are not suitable any more. Our habitation methods should be improved too. Smart houses, smart electrical grids, smart energy storage.
All this costs too much - that's why the feasibility studies should be carried out with extensive simulation to show the effectiveness and financial consequences of the planned changes. A lot of work to do - hopefully it is still not too few, not too late.

Janos Sebestyen Janosy

MTA EK Centre for Energy Research Hungarian Academy of Sciences, Hungary

Gregory Zacharewicz

Ims Université Bordeaux 1 France

Letizia Nicoletti

Cal-Tek Srl, Italy

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

ACKNOWLEDGEMENTS

The SESDE 2019 International Program Committee (IPC) has selected the papers for the Conference among many submissions; therefore, based on this effort, a very successful event is expected. The SESDE 2019 IPC would like to thank all the authors as well as the reviewers for their invaluable work.

I3M 2019 INTERNAL STAFF

AGOSTINO G. BRUZZONE, *DIME, UNIVERSITY OF GENOA, ITALY*
ALESSANDRO CHIURCO, *CAL-TEK SRL, ITALY*
JESSICA FRANGELLA, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
CATERINA FUSTO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
LUCIA GAZZANEO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
FRANCESCO LONGO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
MARINA MASSEI, *DIME, UNIVERSITY OF GENOA, ITALY*
LETIZIA NICOLETTI, *CAL-TEK SRL, ITALY*
ANTONIO PADOVANO, *DIMEG, UNIVERSITY OF CALABRIA, ITALY*
CATALDO RUSSO, *CAL-TEK SRL, ITALY*
KIRILL SINELSHCHIKOV, *SIMULATION TEAM, ITALY*
MARCO VETRANO, *CAL-TEK SRL, ITALY*

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

This International Workshop is part of the I3M Multiconference: the Congress leading **Simulation around the World and Along the Years**

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

Index

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.

MODELING AND ANALYSIS OF THE IMPACT OF TEXTURING ANGLES ON DOPING PROFILES IN ION IMPLANTED N-TYPE SOLAR CELLS

Abderrazzak El Boukili

Al Akhawayn University in Ifrane, Morocco

Email: a.elboukili@aui.ma

ABSTRACT

The aim of this paper is to develop an accurate model to study the impact of texturing angles on doping profiles in ion implanted solar cells. This study will help designers and manufacturers choose an optimal angle in texturing the surfaces of innovative solar cells. Using an optimal texturing angle will improve the performance of solar cells. Randomly chosen texturing angles may decrease the absorption of the sun light or introduce excessive defects as clustering or channeling. These defects will represent recombination centers for active electrons and holes. This will contribute seriously to the loss of the active carriers and then to the loss of solar cell efficiency. This loss is known as a recombination loss. This loss alone may reduce the efficiency of a solar cell by 20%. Numerical results showing the effects of texturing angles on doping profiles will be presented, analyzed, and validated.

Keywords: Mathematical models, simulation, texturing angles, solar cells.

1. INTRODUCTION

Numerical modeling and simulation represent a powerful approach that could be used by researchers and manufacturers to understand and investigate the loss mechanisms that are today limiting the efficiency of industrial solar cells. In this paper, we are developing and using an accurate numerical model to investigate, understand, and optimize the effects of texturing angles on the doping profiles and efficiency of modern textured and ion implanted silicon solar cells.

For the past few decades, the workhorse of the terrestrial solar cell industry has undoubtedly been the P-type crystalline silicon solar cells. However, this technology is reaching the limits of its efficiency potential (Bothe 2005), (Ho 2011), (Fabian 2014). Today, there is a need for innovative solar cell designs in order to reach higher efficiencies at lower cost.

N-type substrates and ion implantation are among the candidates of choice for the new silicon-based solar cells in recent photovoltaic market (Bothe 2005), (Ohrdese 2011), (Zimbardi 2012).

Phosphorus doped N-type solar cells present significant advantages over boron doped P-type solar cells. Mainly, N-type solar cells have higher carrier lifetime than P-type cells (Benick 2008, 2009), (Rohatgi 2012). This is primarily caused by the absence of the well-known boron-oxygen defects that exist in P-type cells (Glunz 2001). N-type solar cells do not suffer from light-induced degradation and are insensitive to common contaminants in the fabrication process (Rohatgi 2012); (Geerligs 2004).

Light-induced degradation is a well-known phenomenon in boron doped P-type silicon substrates. During the fabrication of a solar cell, oxygen impurities can be introduced into the bulk. Therefore, under sun light illumination, the minority carrier lifetime of the solar cell is seen to decrease over time. This issue has been caused by the formation of the metastable boron-oxygen complexes (Glunz 2001).

The use of phosphorus doped N-type silicon as substrate material of the new solar cells has the advantage of bypassing this light-induced degradation (Coletti 2012).

Ion implanted solar cells have recently emerged as the best candidates to make low cost and high efficient industrial solar cells (Rohatgi 2012); (Pawlak 2012); (Benick 2009); (Ohrdes 2011); (Zimbardi 2012); (Coletti 2012); (Meier 2010).

Different doping methods are used by different companies to doping industrial solar cells. Some companies use a blanket doping process from the gas phase by using phosphorous-ox-chloride (POCl) or from the paste attached by screen printing (Bateman 2011); (Ryu 2012); (Spitzer 1984). But, blanket doping is very imprecise and does not allow selective doping.

Recently, most of the solar cell manufacturers are using ion implantation technique to doping high efficiency and low cost solar cells. Ion implantation process has many advantages over the blanket doping process. Ion

implantation allows better control of the doping doses, doping depths, doping profiles, and doping-induced defects by adjusting the doping energy of the ion beam and other doping parameters as desired.

Advanced Texturing is one of the best candidates used today in photovoltaic industry to enhance the optical absorption of crystalline silicon solar cells and make lower cost solar cells. Efficient texturing could reduce the sun light reflection losses at the surface of the solar cells by more than 10% (Netsor 2010).
Randomly textured Silicon (001) wafers are used in modern solar cells (Xing 2011); (Netsor 2010). The texturing is fabricated using wet-chemical KOH etch including some special additives. After the application of etching, the textured solar cell surfaces will consist of square pyramids with {111} surfaces of different heights and angles.

Ion implantation on solar cells with textured surfaces is significantly different from ion implantation on solar cells with planar surfaces. This is due to the fact that texturing will increase the surface area of the solar cell and introduce extra defects. The increase of the surface area will lead to a reduction in the effective dose of the implanted ions. This reduction of the effective dose depends strongly on the texturing angle . The texturing angle may also create new types of doping-induced defects.

The goal of this paper is to develop and use an accurate mathematical model to investigate, understand, and optimize the effects of texturing angle on the doping profiles in ion implanted N-type solar cells with textured surfaces. The measurement of doping profiles in textured solar cells is still a challenge. Therefore, calculating the doping profiles in textured surfaces using mathematical models is strongly needed.

This model has been implemented in the process simulation software Suprem-IV and has been used to simulate different doping profiles under different texturing angles.
To the best of the authors' knowledge, no theoretical studies exist in literature in two dimensions (2D) about the effects of the texturing angle on the doping profiles in 2D textured solar cells.

This paper is organized as follows. Section 2 presents the theoretical models we have developed. Section 3 presents the numerical results and validation.
Section 4 presents the concluding thoughts and future work.

2. MATHEMATICAL MODEL

To improve efficiency of the novel solar cells, the reflection of the sunlight should be reduced. Surface texturing is used to reduce light reflections and enhance light trapping. When the photons of the sunlight hit the surface of the solar cell, some lights are reflected, some are passed through, and some are absorbed by the p-n junction region. Absorbing the light produces current by creating electron-hole pairs. Roughening the surface is called texturing. Texturing can be done by different ways.

Moncrystalline silicon solar cells are generally processed in an alkaline texturing solution to produce a random set of upright pyramids on the etched surface (Singh 2001); (Vazsonyi 1999); (Nishimoto 2000); (Marrero 2007) as shown in Figure 1 from (Jan 2014). It is the anisotropic nature of such an etchant that drives the formation of a repeated pyramidal geometries with different sizes and shapes as in Figure 1. This anisotropy is due mainly to the fact that the etch rate in the <100> direction is several times greater than that in the <111> direction. Many different etchants have been investigated, including aqueous solutions of NaOH (Singh 2001); Na2CO3 (Nishimoto 2000); KOH (Marrero 2007) and TMAH (Papet 2006).

The final geometry of the obtained pyramids depends strongly on the texturing techniques, texturing temperatures, composition of the solution, and on the pre-etch properties of the silicon surface. The geometry and the size of the obtained pyramids depends on the texturing angle and width (w) and height (h) of the pyramids as shown in the Figure 1.

To help designers and manufacturers of industrial solar cells find and use the values of , w, and h that will lower cost and improve efficiency, we should use mathematical models and numerical simulations.

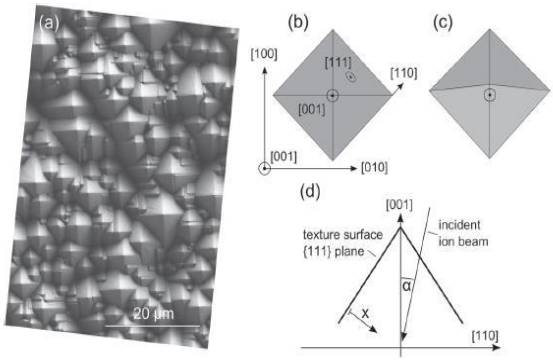


Figure 1: SEMS images of random pyramids.

Let ϕ be the dose of the incident ions from the ion implantation process. Let ϕ_{eff} be the effective dose of the ions that will penetrate the silicon substrate. Because of texturing ϕ_{eff} is smaller than ϕ. Let's define the texturing angle to be the angle between the incident ions and the normal to the substrate as

shown in Figure 1. Using simple vector analysis, we get the relation between ϕ, ϕ_{eff}, and as follows:

$$\phi_{eff} = \phi.\cos(\alpha) \quad (1)$$

When no texturing is applied, $\alpha = 0$, and $\phi_{eff} = \phi$.

Ion implantation is a random process. In literature (Betemann 2011); (Ryu 2012) the distribution of ions in the silicon substrate has been represented by Gauss probability density function C(x) for symmetric distributions and by Pearson IV probability density function f(x) for asymmetric distributions. Here, x represent the ion penetration depth. The Gaussian density function C(x) is given by:

$$C(x) = \frac{\phi}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(x-R)^2}{2\sigma}\right], \quad (2)$$

where, R is the projected range, σ is the standard deviation. Let υ_i represents the ith moment of C(x). Then, we have:

$$\upsilon_i = \frac{\int_0^\infty (x-R)^i C(x)dx}{\int_0^\infty C(x)dx} \quad (3)$$

Let γ, β represent the skewness and the kurtosis of C(x) respectively. Then, R, σ, γ, and β are calculated from the four moments of C(x) as follows:

$$R = \upsilon_1; \sigma = \sqrt{\upsilon_2}; \gamma = \frac{\upsilon_3}{\sigma^3}; \beta = \frac{\upsilon_4}{\sigma^4} \quad (4)$$

The Pearson IV probability density function f(x) is the 4th solution of the following first-order linear differential equation:

$$\frac{f'(s)}{f(s)} - \frac{s-a}{b+cs+ds^2} = 0; s = x - R \quad (5)$$

This differential equation has 7 solutions. The 4th solution is given by:

$$f(x) = \phi E \exp\left[F \times G\right]. \quad (6)$$

Where,

$$E = \left|b + cs + ds^2\right|^{\frac{1}{2d}} \quad (7)$$

$$F = -(\frac{c}{2d} + a)\frac{2}{\sqrt{4bd - c^2}} \quad (8)$$

$$G = \arctan(\frac{2ds + c}{\sqrt{4bd - c^2}}) \quad (9)$$

The coefficients a, b, c, d are calculated from the four moments of C(x) as follows:

$$a = c \quad (10)$$

$$b = \frac{4\beta - 3\gamma^2}{10\beta - 12\gamma^2 - 18}\sigma^2 \quad (11)$$

$$c = -\frac{\beta + 3}{10\beta - 12\gamma^2 - 18}\gamma\sigma \quad (12)$$

$$d = -\frac{2\beta - 3\gamma^2 - 6}{10\beta - 12\gamma^2 - 18} \quad (14)$$

The originality in this paper is that the dose ϕ in the equations (2) and (6) used to calculate the Gaussian and the Pearson IV densities C(x) and f(x) is replaced by the effective dose ϕ_{eff} which includes the texturing angle . In the new models we are proposing, the ion densities C(x), and f(x) are be replaced by the ion densities $Cn(x,\alpha)$ and $fn(x,\alpha)$ given by:

$$\begin{aligned} Cn(x,\alpha) &= \phi\cos(\alpha)C(x) \\ fn(x,\alpha) &= \phi\cos(\alpha)f(x) \end{aligned} \quad (11)$$

We have implemented these novel densities in the open source process simulator Suprem-IV (Hansen, 1993). We have used them to investigate the effects of the angle on the doping profile of Boron in ion implanted and textured silicon (001) solar cell shown in Figure 2.

3. NUMERICAL RESULTS AND VALIDATION

We use an N-type silicon (001) solar cell textured with identical pyramids of faces (111) as shown in Figure 2. The texturing angle was varied from 20° to 64° degrees. The height of pyramids is 0.12um. The width of the pyramids is 0.05um. The dose of Boron is $2\times10^{15} cm^{-2}$, implant energy is 10 KeV, the implant model is $fn(y,\alpha)$. The depth, y, of the cell is 0.37 um and the length, x, is 0.3 um. Figure 2 also shows the 2D Boron doping profile on textured solar cell for =36°. Figure 3 shows the cut of the Boron doping profile at x=0.15 and =36° and y between 0 um and 0.15 um. Figure 4 shows the cut of the Boron doping profile at x=0.15 and =64° and y between 0 um and 0.15 um. Figure 5 shows the cut of the Boron doping profile at x=0.15 and y between 0 um and 0.15 um when the surface of the solar cell is not textured. By comparing Figure 5 with Figures 3 and 4, we see that the texturing of the solar cell changes significantly the doping profile. This will affect significantly the efficiency of the textured solar cell. From Figures 3 and

4 we see the significant effects of the texturing angle on the doping profile. Two effects are visible: (i) reduction of the doping values with increasing . (ii) reduction of the implantation depth with increasing . The first effect could be explained by the reduction of the effective dose due to the angle as shown by the Equation (1). The other effect (ii) could also be explained by the reduction of the effective dose due to the angle . Small ion doses will produce shallow doping profiles seen in ii). For =36°, the maximum value of the doping is 1.2e+20 (atom/cm-2) (see Figure 3). For =64°, the maximum value of the doping is 7e+19 (atom/cm-2) (see Figure 4). These 2 values are smaller than the maximum value of the doping on non-textured surface which is 4e+20 (atom/cm-2) (see Figure 5). The implantation depth on textured surface for =36° is about 0.06 um and on non-textured surface for =0° is about 0.08 um. From these simulation results, we could suggest to use smaller texturing angles for ion implantation on textured surfaces if shallow profiles are not needed. On the other hand, it has been shown in (Zahi, 2018), that with smaller angles around 36°, the measurements of doping profiles in textured solar cells using the secondary ion mass spectrometry (SIMS) technique are more accurate than with larger angles. We should note that the measurements of doping profiles using SIMS technique are strongly depending on the texturing angles as well. The doping profiles obtained in this simulation are in a good agreement with SIMS measured profiles obtained in (Zahi 2018).

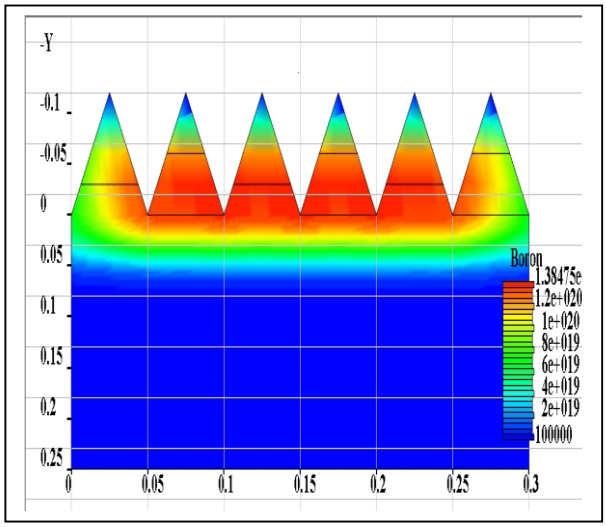


Figure 2: 2D sample textured solar cell.

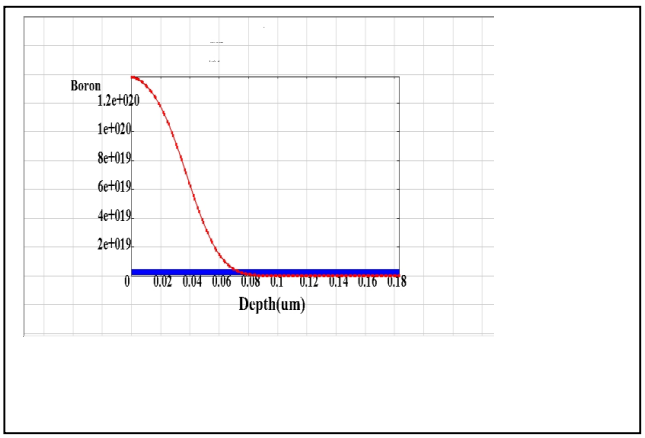


Figure 3: Boron doping profile on textured surface for α =36°.

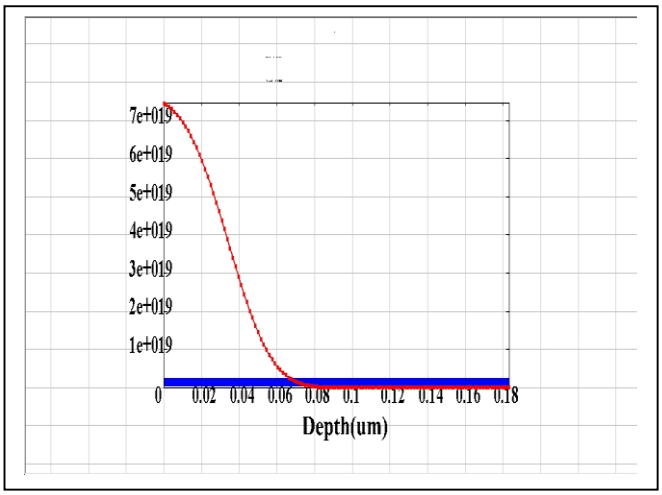


Figure 4: Boron doping profile on textured surface for α =64°.

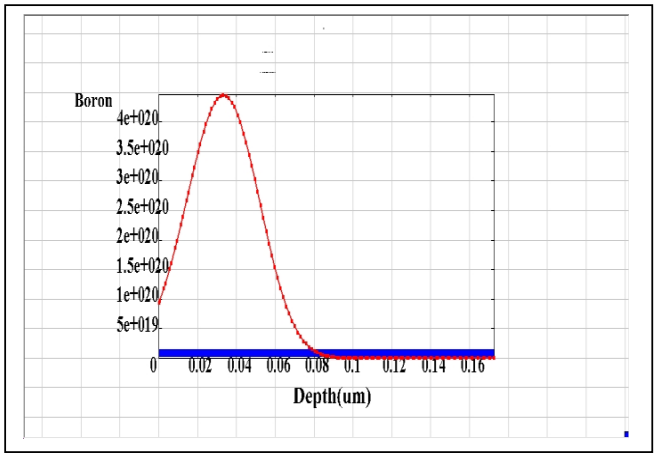


Figure 5: Boron doping profile on non-textured solar cell (for α =0°).

4. CONCLUSIONS

We have shown that the texturing angle affects significantly the doping profiles in textured solar cells. The values of the doping and the implantation depth are decreasing with increasing . By comparing the doping profiles on planar surface shown in Figure 5 with doping profiles on textured surfaces shown in Figures 3 and 4, we see that the shape of the doping profiles are also affected by the texturing angle . In future work, we will investigate the effects of the size of the pyramids on the doping profiles.

REFERENCES

Bothe, K., Sinton, R., Schmidt, J., 2005. Fundamental boron-oxygen-related carrier lifetime limit in mono- and multicrystalline silicon. Progress in Photovoltaics Research and Applications 13(4):287 – 296.

Ho, W., Huang, Y., Hsu, W., Chen, Y., Liu, C., 2011. Ion implanted boron emitter N-silicon solar cells with wet oxide passivation. 37th IEEE Photovoltaic Specialists Conference, Seattle, WDC, USA.

Fabian, K., Tobias, O., Robby, P., Rolf, B., 2011. Analyzing the recombination current densities in industrial like n-type PERT solar cells exceeding 20% efficiency. 23rd IEEE Photovoltaic Specialists Conference, Seattle, WDC, USA.

Kveder, V., Kittler, M., Schroter, W., 2001. Recombination activity of contaminated dislocations in silicon: A model describing electron-beam-induced contrast behavior. Physics Review, B 63, 115208.

Benick, J., Hoew, B., Van, M., Kessels, O., Schultz, O., Glunz, W., 2008. High efficiency n-type si solar cells on alo-passivated boron emitters. Applied Physics Letters, Vol. 92, pp. 253504.

Benick, J., Hoew, B., Dingemans, G., Richter, A., Hermle, M., Glunz, W., 2009. High efficiency n-type si solar cells with front side boron emitter. Proceedings of the 24th European Photovoltaic Solar Energy Conference, pp. 863-870.

Ohrdes, T., Steingrube, S., Wagner, H., Zechner, C., Letay, G., Chen, R., Dunham, S., Altermatt, P., 2011. Solar cell emitter design with pv-tailored implantation. Energy Procedia, Vol., 8, pp. 167-173.

Zimbardi, F., Upadhyaya, D., Tao, Y., OK, Y., Ning, S., Rohatgi, A., 2012. Ion implanted and screen-printed large area 19.6% efficiency n-type bifacial si solar cell. Photovoltaic Specialits Conference, pp. 002240-002243.

Rohatgi, A., Meier, B., McPherson, B., OK, Y., Upadhyaya, D., Lai, H., Zimbardi, F., 2012. High-throughput ion-implantation for low-cost high efficiency silicon solar cells. Energy Procedia, Vol., 15, pp.10-19.

Glunz, S., Rein, S., Lee, J., Warta, W., 2001. Minority carrier lifetime degradation in boron doped czochralski silicon. Journal of Applied Physics, Vol., 90, no. 5, pp.2397-2404.

Geerligs, L., Machdonald, L., 2004. Recombination activity of interstitial iron and other transition metal point defects in p-type crystalline silicon. Applied Physics Letters, Vol., 85, no. 18, pp. 4061-4063.

Coletti, G., Mihailetchi, V., Komatsu, Y., Geerligs, L., Kvande, R., Arnberg, L. Wambach, K., Knopf, C., Kopecek, R., Weeber, A., 2012. Large area screen printed n-type base silicon solar cells with efficiency exceeding 18%. Solar energy, Vol. 2011, pp. 2010.

Pawlak, B., Janssens, T., Singh, S., Kuzma, I., Robbelei, J., Posthuma, N., Poortmans, J., Cristioan, F., Bazizi, E., 2012. Studies of implanted boron emitters for solar cell applications. Progress in Photovoltaic: Research and Applications, Vol. 20, no. 1, pp. 106-110.

Benick, J., Hoex, B., Dingemans, G., Richter, A., Hermle, M., Glunz, S., 2009. High efficiency n-type silicon solar cells with front side boron emitter. Proceedings of the 24th European Photovoltaic Solar nergy Conference, pp. 863-870.

Meier, D., Rohatgi, A., 2010. Developing novel low-cost, high-throughput processing techniques for 20% efficient monocrystalline silicon solar cells. Photovoltaics Internation, Vol. 10, pp.87-93.

Bateman, N., Sullivan, P., Reichel, C., Benick, J., Hermle, Rohatgi, D., A., 2011. High quality ion implanted boron emitters in an inter-digitated back contact solar cell with 20% efficiency. Energy Procedia, Vol. 8, pp. 509-514.

Ryu, K., Upadhyaya, A., Ok, W., Xu, H., Metin, L., Rohatgi, N., 2012. High efficiency n-type solar cells with screen-printed boron emitters and ion-implanted back surface field. Photovoltaic Specialists Conference, 38th IEEE, pp. 002247-002249.

Xing, S., Steven, G., Jurgen, M., Lionel, C., 2011. Optimization-based design of surface textures for thin-films Si solar cells. Optics Express, Vol. 19, no. S4, pp. 841-850.

Netsor, X., Amir, D., Sascha, S., Giso, H., , S., Steven, G., Jurgen, M., Lionel, C., 2010. Influence of pyramid size of chemically textured silicon wafers on the characterization of industrial solar cells. Proceedings of the 25th EU PVSEC, 6-10 September, Valencia.

Liopis, F., Tobias, I., 2005. Influence of texture size on the optical performance of silicon solar cells. Progress on Photovoltaic: Research and Applications, Vol. 13, no.27.

Jan, K., Robby, P., Jorg, O., 2014. Structural analysis of textured silicon surfaces after ion implantation under tilted angle. Semiconductor Science and Technology, Vol. 29, pp. 1-7.

Singh, P., Kumar, R., Lal, M., Das, B., 2001. Effectiveness of anisotropic etching of silicon in aqueous alkaline solutions. Solar Energy Materials and Solar Cells, Vol. 70, no. 1, pp. 103-113.

Vazsonyi, E., Declercq, K., Einhaus, R., Van, E., Said, K., Poortmans, J., Szlufcik, J., Nijs, J., , P., Kumar, R., Lal, M., Das, B., 1999. Improved anisotropic etching process for industrial texturing of silicon solar cells. Solar Energy Materials and Solar Cells, Vol. 57, no. 2, pp. 179-188.

Nishimoto, Y., Namba, K., 2000. Investigation of texturisation for crystalline silicon solar cells with sodium carbonate solutions. Solar Energy Materials and Solar Cells, Vol. 61, no. 4, pp. 393-402.

Marrero, N., Gonzalez, D., Guerrero, R., Borhcert, D., Hernandez, R., 2007. Optimization of sodium carbonate texturisation on large-area crystalline silicon solar cells. Solar Energy Materials and Solar Cells, Vol. 91, no. 20, pp. 1943-1947.

Hansen, S., Deal, M., 1993. Two dimensional process simulation for silicon and Gallium arsenide. Suprem IV User's Manual, Stanford University, USA.

Zahi, E., Nadjib, T., Bernards, S., Cedreic, B., Barbara, B., Maurice, Q., 2018. Doping profile measurement on textured silicon surface. EPJ Photovoltaics Vol. 9, No. 5, pp. 1-8.

ELECTRICITY SUPPLIER SELECTION BY A HOUSEHOLD IN THE CZECH REPUBLIC IN 2017 AND 2018 – MONTE CARLO SIMULATION APPROACH

Martina Kuncova[(a)]

[(a)]University of Economics Prague, Czech Republic, Faculty of Informatics and Statistics, Dpt. Of Econometrics

[(a)]martina.kuncova@vse.cz

ABSTRACT

The situation on the electricity retail market in the Czech Republic is not clear because of the number of suppliers and its products. Although the information about the prices for the electricity consumption for households is available on the web and each household can change the supplier nearly with no extra effort and cost, households are still often not familiar with the individual price items of the products. In this article the analysis of the Czech electricity market from the distribution rate D25d point of view is made for the years 2017-2018 when the household annual consumption is simulated via Monte Carlo simulation model. The aim of this paper is to select such a supplier and product that minimizes the total costs of the electricity for a household for the selected distribution rate and compare it with the results from the previous years.

Keywords: Electricity consumption, distribution rate D25d, Monte Carlo simulation, cost minimization

1. INTRODUCTION

Electricity is nowadays a commodity that is essential for nearly all activities we are used to doing. As still new and new electricity-based equipment arises we might expect the rise of electricity consumption. But on the other hand, there is the necessity of renewable resource usage and the decrease in the electricity consumption of the new machines. That is why the final trend in the consumption is not so clear.

Electricity is used in all sectors of the economy. Usually, the biggest part of the electricity is consumed by industry followed by services or households, energy transfer, transport, etc. The development of the consumption in European Union countries in the years 1990-2016 is shown in Figure 1 – the highest line is industry followed by services and residential sector (households). The structure of the consumption in 2016 in the Czech Republic is seen in Figure 2. The situation is similar: the highest share of electricity consumption is associated with the industry (30%) followed by households (26%).

Based on the data from Eurostat (2019) we see in Figure 3 that nearly in the half of the European Union countries the electricity consumption by households was higher in 2016 compared to 2006, but in the other half, we see the decrease of the electricity consumption.

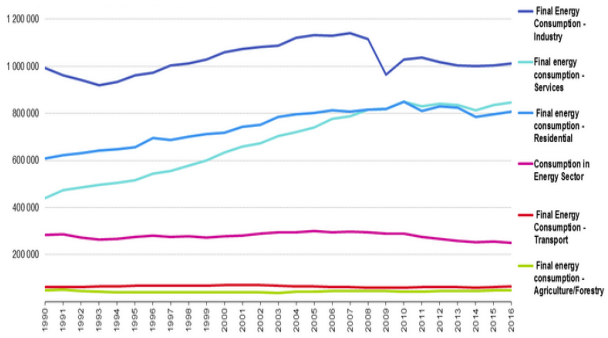


Figure 1: Consumption of electricity by sector in EU 28. Source: Eurostat (2019a)

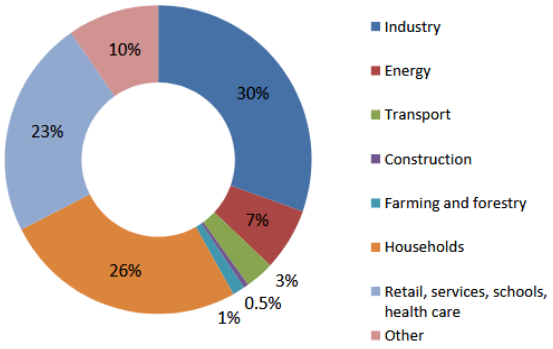


Figure 2: National economy sectors' shares of total electricity consumption in the Czech Republic. Source: ERU (2016)

The highest increase we see in Romania (20.7%), the highest decrease was in Belgium (about 17%, so the rate in 2016 was 83% compared to 2006). In the Czech Republic, we see the small decrease (1.7%) that is close to the EU average (decrease about 0.7%). As Karanfil and Li (2015) explain, the electricity-growth nexus is highly sensitive to regional differences, countries' income levels, urbanization rates and supply risks. The increase of the consumption of households might be caused by the development of ICT and the growing number of various machines and other electrical devices, the decrease by the lower demand of new

appliances for the electricity consumption. For any household it is probably the most used commodity although it does not belong to the common consumer goods - it is invisible, untouchable and it is used only via other equipment.

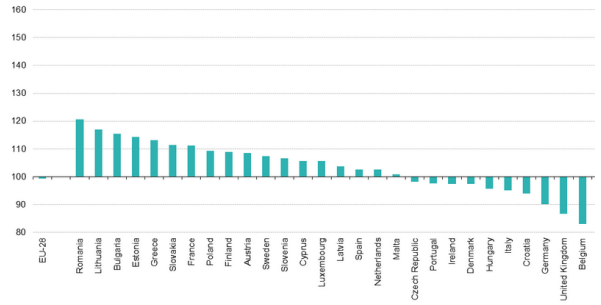


Figure 3: Electricity consumption by households in 2016 compared to 2006 in EU28. Source: Eurostat (2019b)

The liberalization process which started in Europe in the period 2003-2009 initiated the possibility to choose the electricity supplier by households (Newbery, 2013). Each household cannot choose the distributor but only the supplier that sells the electricity. The complete list of the suppliers and their products and prices is changing every year and sometimes it is hard to follow the rules nor the conditions for the households to understand the situation and to choose the appropriate product. The selection of the product depends on the contract conditions given by the distributor but mainly on the prices set by distributors and suppliers. Various techniques and methods can be used to model the situation on the market – multi-criteria analysis, simulation modelling, optimization (Ventosa et al. 2005; Kuncova 2015). In this paper the simulation of the electricity consumption of one household (tariff rate D25d) via Monte Carlo model is applied on the data from the years 2017 and 2018 for all suppliers and distributors in the Czech Republic. Afterwards, the comparison with the previous research (Kuncova, Seknickova 2014; Kuncova, Seknickova 2015; Kuncova 2015; Kuncova 2016) is added.

2. CZECH ELECTRICITY MARKET

The development of the energy sector in the Czech Republic after the Velvet revolution in 1989 started with the restructuring of the centrally run energy system and its division into a number of smaller entities with one state-owned (Vlcek, Cernoch 2013). Afterwards, the privatization of the key energy companies followed. Based on the EU demands, the liberalization of the Czech electricity market started in 2002, later from 2006 also the households could choose its electricity and gas supplier – but as the situation was hardly understandable for a household, the higher number of suppliers' switches has started around the year 2010 (Vlcek, Cernoch 2013) – see Figure 5.

Except suppliers, there are other subjects on the electricity market, especially 3 distributors (PRE, CEZ, E.ON), the Energy Regulatory Office (ERU) and the Operator of the market (OTE). OTE predicates the whole market consumption and analyses the differences, ERU regulates the prices of the transfer and distribution of the electricity. The high number of the suppliers and their products on the retail market (Figure 4) embarrasses the position of the households. Based on the ERU (2019) data, it is evident that only 9 companies (including 3 distributors that also serves as suppliers) have been on the market offering D25d rates products since 2011.

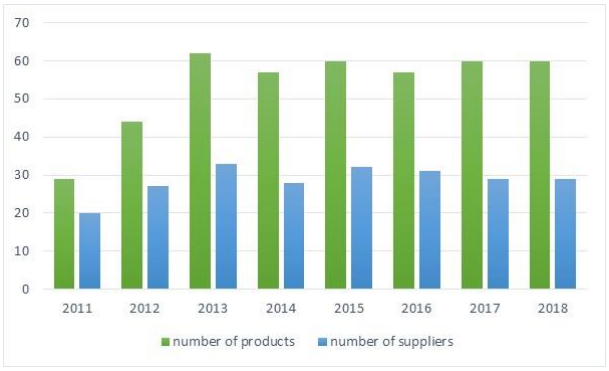


Figure 4: Number of suppliers and its products. Source: ERU (2019)

As the ERU offers the online calculator (ERU, 2019) and the conditions of the supplier switch are more and more explained in media, the number of supplier changes increased, mainly last year (Figure 5). The reason might be connected with the increase of electricity consumption prices – the development of the average prices for high and low tariff rates in each distribution region for the selected distribution rate D25d are shown on Figure 6. This tariff rate is given to household when the electricity is used also for the accumulative heating and hot water heating for lower and middle yearly offtake with operative management of the validity period of the low tariff for 8 hours. It is the so-called dual tariff rate as it has 2 periods (high tariff, low tariff) during the day (ERU, 2019).

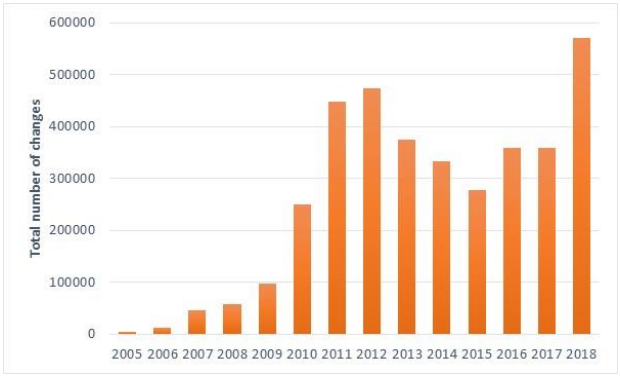


Figure 5: Number of changes of electricity suppliers in the Czech Republic. Source: OTE (2019)

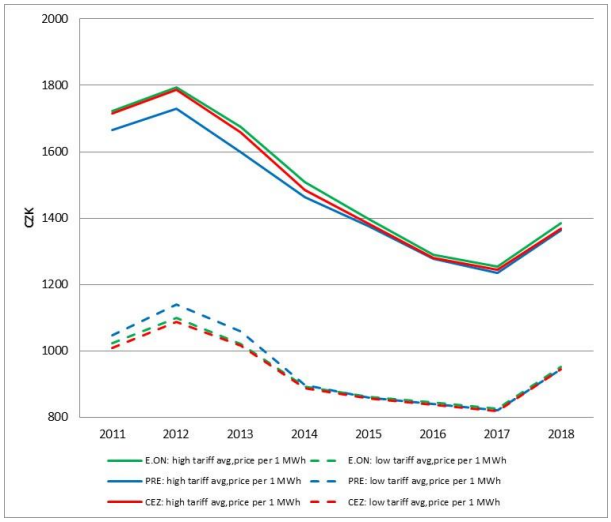


Figure 6: The development of the average prices in high and low tariffs per 1MWh for D25d rate in 3 distribution regions. Source: ERU (2019), own calculations

The final costs for the electricity consumption of the household are influenced by consumption, fixed fees or taxes. Generally, the prices are divided into two components. The first one is the controlled charge for services related to electricity transport from the generator to the final customer. This charge is annually given by ERU (2019). It covers:

- monthly lease for the circuit breaker,
- price per megawatt hour (MWh) in high tariff (HT),
- price per megawatt hour in low tariff (LT),
- price per system services,
- price for the support of the renewable energy purchase,
- charges for the electricity market operator,
- electricity ecological tax (28,30 CZK per 1 MWh).

The second part of the total price is given by the electricity supplier. It covers:

- fixed monthly fee for the selected product,
- price per megawatt hour (MWh) in high tariff (HT),
- price per megawatt hour in low tariff (LT).

The final price is increased by VAT that is 21% from 2013.

3. DATA AND METHODS

3.1. Data for the Analysis

Several previous analyses were aimed at the comparison of the products offered for the households in the distribution rate D25d: Kuncova, Seknickova (2014) where multi-criteria evaluation methods were used; in Kuncova, Seknickova (2015) the Monte Carlo simulation was applied on the data from the year 2015; in Kuncova (2015) the multi-criteria comparison, Monte Carlo simulation and the linear optimization model were used and the results of the product choice compared; in Kuncova (2016) the changes in period 2011-2016 were described and analyzed. To be able to compare the results from the years 2017 and 2018 with the previous ones the same assumptions are used in this article.

The ranges for the electricity consumption were taken from the real data and real household with the electricity consumption about 10 MWh annually, 45% energy in high tariff and 55% in low tariff and with the circuit breaker from 3x20A to 3x25A.

Data for the 60 products offered by 29 companies for the years 2017 and 2018 were taken to find the best product for each year with the lowest electricity consumption cost. As there are differences in prices in the distribution regions, all 3 regions were analyzed separately. The distributors' prices and the average prices of the suppliers are in Table 1 and Table 2. As it was mentioned above, the prices raised in 2018 compared to 2017 but the increase was not equal in all distribution regions. The most expensive (for the households) is still the CEZ region (north part of the Czech Republic) which is visible from Figure 6. In this region the highest increase of the average suppliers' monthly fee in 2018 compared with 2017 is visible – about 34%, while in other two regions it was about 12% (PRE region) or 13% (E.ON region).

Table 1: Prices for each distribution area

year	distrib. region	circuit-breaker monthly fee	Distrib. HT price per 1 MWh	Distrib. LT price per 1 MWh	Distrib. Other services price per 1 MWh
2017	E.ON	113	1607.38	72.22	593.84
	PRE	116	1503.96	70.58	
	CEZ	120	1624.71	64.01	
2018	E.ON	119	1687.78	77.28	594.03
	PRE	123	1590.53	75.52	
	CEZ	127	1719.17	71.669	

Table 2: Average prices of the suppliers for each distribution area

year	distrib. region	Suppliers monthly fee avg,	high tariff avg,price per 1 MWh	low tariff avg,price per 1 MWh
2017	E.ON	43.838	1253.853	826.139
	PRE	46.824	1234.639	820.437
	CEZ	44.990	1244.705	818.420
2018	E.ON	49.646	1384.163	951.274
	PRE	52.630	1363.746	945.424
	CEZ	60.446	1369.246	945.424

3.2. Monte Carlo Simulation

Simulation models can be applied in a situation when some variables of the model are uncertain. The simulation itself is a technique for imitation of some real situations, processes or activities that already exist or that are in preparation – just to create a computer model (Banks 1998). Monte Carlo simulation (or technique) is closed to statistics as it is a repeated

process of random sampling from the selected probability distributions that represent the real-life processes (Turban, Meredith 1994). This method is based on running many times and for each sample random variates are generated on each input variable (Thomopoulos, 2013). Based on the existed information the type of probability distribution (that corresponds with the expectations about the values of the variable) must be selected. The most typical and frequent distribution types are normal, triangular, uniform (discrete uniform), Poisson, lognormal and exponential ones. Mathematical specification of these variables and the calculations derived from them might be complicated (especially when a non-trivial distribution is chosen). But via the simulation Monte Carlo and via MS Excel and its add-ins (for example Crystal Ball or @RISK) it is possible to analyze the problem and find a recommendation for each specified situation.

Crystal Ball is one of the MS Excel add-in applications for the Monte Carlo simulation models. "Oracle Crystal Ball solutions begin with Oracle Crystal Ball, the base package for predictive modeling, Monte Carlo simulation and forecasting. Oracle Crystal Ball Enterprise Performance Management builds on that set of tools by adding the power of Oracle Crystal Ball to your Oracle Enterprise Performance Management (EPM) and Oracle Business Intelligence (BI) applications" (www.oracle.com). The advantage of this software is the usage of Excel tables. It is possible to use models created before but change the distribution for random inputs generation. Then usually 1000 trials are run and afterwards the programme gives all statistics (and histogram) of the selected decision cell. Figure 7 shows all the possible statistical distributions that can be chosen.

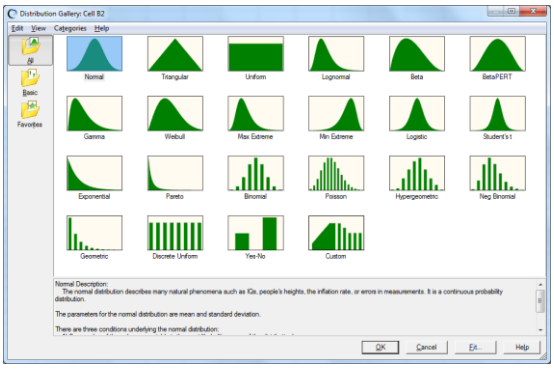


Figure 7: Crystal Ball – Distribution Gallery (www.oracle.com)

Monte Carlo simulation is associated with the systems affected by randomness when several different scenarios are randomly generated to obtain the probability description of the selected results (Brandimarte, 2014). The method repeats a lot of random experiments to find out the possible outcomes. This is a typical situation for various decision-making processes in finance (Razgaitis 2003), banking (Kuncova, Lizalova 2012) and also in energetics to generate the whole demand for the distributed units (Hegazy et al. 2003) or to generate the annual electricity consumption (Kuncova, Seknickova 2014).

To find the best product the online ERU calculator (2019) can be used – the only problem is to set the annual electricity consumption. But as the consumption is not fixed for more years, the Monte Carlo simulation could be another possibility to find the best product minimizing the annual electricity consumption costs for the given household.

For the calculations we use the same model as in Kuncova, Seknickova (2014; 2015). The ranges for the electricity consumption in each month were set as values from the normal distribution with the standard deviation equal to 20% of the average and the average (at about 900 kWh per month – Table 3) was taken from the real data of the household.

Table 3: Parameters of the normal distribution used in the simulation model

Month	Mean (kWh)	St.deviation (kWh)
January	933	186.6
February	973	194.6
March	900	180
April	819	163.8
May	771	154.2
June	730	146
July	689	137.8
August	665	133
September	730	146
October	795	159
November	835	167
December	892	178.4

The consumption has been generated for each month from the normal distribution with 45% in the higher tariff rate. In all Monte Carlo simulations, 1000 experiments have been tried using MS add-in application Crystal Ball to randomly select consumption for each month and afterwards the annual costs are calculated. The formula for the annual cost calculation for each supplier's product is following (ERU, 2019):

$$COST_{ij} = (1 + VAT) \cdot \begin{bmatrix} 12 \cdot (mf_{ij} + mf_j + ot) + \\ 0.45 \cdot gc \cdot (ph_{ij} + ph_j) + \\ 0.55 \cdot gc \cdot (pl_{ij} + pl_j) + \\ gc \cdot (os + t) \end{bmatrix} \quad (1)$$

where
i … product, i =1, …, 60,
j … distributor, j = 1, …, 3,
VAT … value added tax (VAT = 0.21 in 2017 and 2018),
mf … fix monthly fee,
gc … yearly generated consumption in MWh,
ph … price in high tariff per 1 MWh,
pl … price in low tariff per 1 MWh,
ot … price for other services per 1 month,
os … price for other services per 1 MWh,
t … electricity tax per 1 MWh (t = 28.3 CZK).

The only difference in the formula (1) in the comparison with the year 2015 is in the *ot* part which is now calculated for 1 month while in 2015 it was a part of *os*.

4. RESULTS

The comparison of suppliers is based on the 1000 simulation made in the MS Excel add-in application Crystal Ball. The results of the years 2017 (Table 4) and 2018 (Table 5) are completely different not only in the average annual costs but also from the suppliers' point of view. The best and worst products (out of 60 for PRE and CEZ, and 59 for E.ON distribution area) are the same for all regions in the given year, but the average annual electricity consumption costs are different. The cheapest region when the same product is selected seems to be PRE in 2017 and E.ON in 2018. The CEZ region is the most expensive one.

The best (cheapest) products in 2018 are about 5-7% (based on the region) more expensive than the winners in 2017 (Figure 8). The difference between the cheapest and the most expensive products in 2017 are about 17.5-23.5%, in 2018 about 25.5-28%, so in 2018 there is not only a price increase but also a widening of the gap between the cheapest and the most expensive product is visible.

The difference in average costs does not necessarily mean that the product will always be more expensive - but even the histograms (Figure 9, Figure 10) show quite a significant difference in the resulting costs.

When we compare the lowest average annual cost in 2017 and the highest annual cost in 2018, the difference is about 36%.

Table 4: Order in 2017 for all distributors and the average annual costs

distr. area / order	Product	E.ON avg. annual cost CZK	PRE avg. annual cost CZK	CEZ avg. annual cost CZK
1	CARBOUNION BOHEMIA,spol.s.r.o. , STANDARD	28742.32	27704.17	28770.51
2	CARBOUNION BOHEMIA,spol.s.r.o. , STANDARD 12	28742.32	27704.17	28770.51
3	Nano Energies Trade s.r.o., Dobrý skutek	28778.59	27733.56	28805.33
4	Fonergy, PREMIUM Aku 8	28983.79	27929.08	29008.5
5	Amper Market, HOME_AKU	29042.79	27994.20	29068.79
57	LAMA energy, STANDARD Akumulace 8 region PRE	32799.81	32421.80	32674.85
58 (56)	ČEZ Prodej, Elektřina na dobu neurčitou	32094.76	32430.87	32783.20
59 (58)	E. ON. Energie, Elektřina/Aku	33221.91	33591.60	33931.83
60 (59)	E. ON. Energie, Elektřina TrendAku_prosinec	33796.54	34182.91	34517.13

Table 5: Order in 2018 for all distributors and the average annual costs

distr. area / order	Product	E.ON avg. annual cost CZK	PRE avg. annual cost CZK	CEZ avg. annual cost CZK
1	Eneka s.r.o., Jednička	29791.17	29842.49	30268.39
2	Europe Easy Energy, eDOMÁCNOST	29939.88	29999.12	30420.60
3	Eneka s.r.o., STANDARD	29991.84	30047.04	30470.78
4	BOHEMIA ENERGY entity s.r.o., Benefit Plus	30230.28	30289.52	30711.00
5	X Energie, PREMIUM	30230.28	30289.52	30711.00
57 (56)	CARBOUNION BOHEMIA,spol.s.r.o. , STANDARD 24	38133.93	36780.26	38109.91
58 (57)	E. ON. Energie, Komplet Elektřina 36	38180.19	36827.18	38156.37
59 (58)	CARBOUNION BOHEMIA,spol.s.r.o. , STANDARD 12	38804.26	37433.77	38775.11
60 (59)	CARBOUNION BOHEMIA,spol.s.r.o. , STANDARD	39043.67	37667.17	39012.69

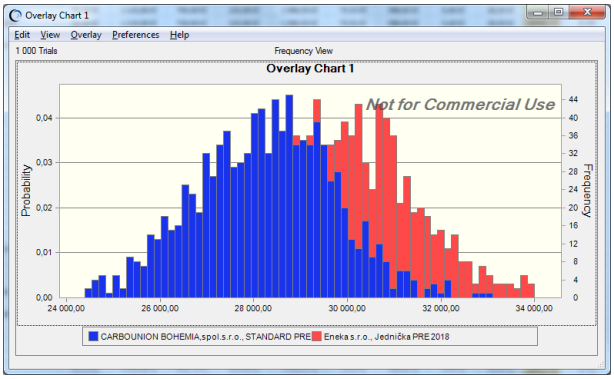


Figure 8: Histograms for the annual costs for the best products in 2017 and 2018

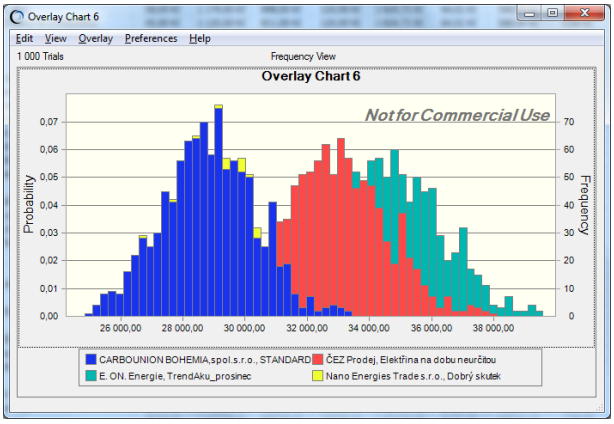


Figure 9: Histograms for the annual costs for the best and worst products in 2017

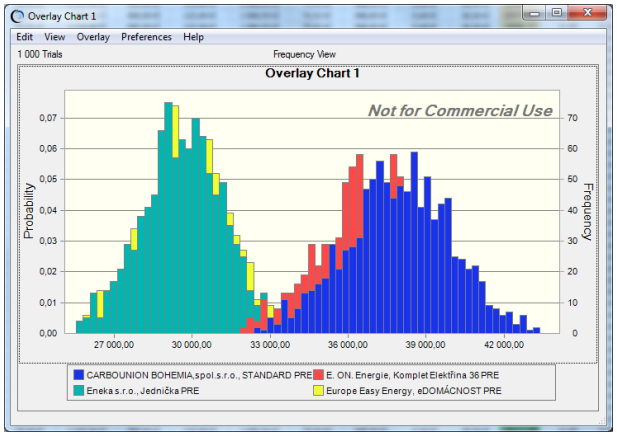


Figure 10: Histograms for the annual costs for the best and worst products in 2018

The annual electricity consumption average costs for the given household and its consumption oscillates around 30000 CZK (about 1130 EUR). The probability of the lowest annual cost (CARBOUNION BOHEMIA spol. s.r.o. STANDARD product) in 2017 is about 80% (Figure 11), but the same product in 2018 is the most expensive one and the probability that the average costs are lower than 30000 CZK goes to zero (Figure 12). The best product in 2018 (Eneka s.r.o. Jednička) has the probability of the cost lower than 30000 CZK about 58% (Figure 13).

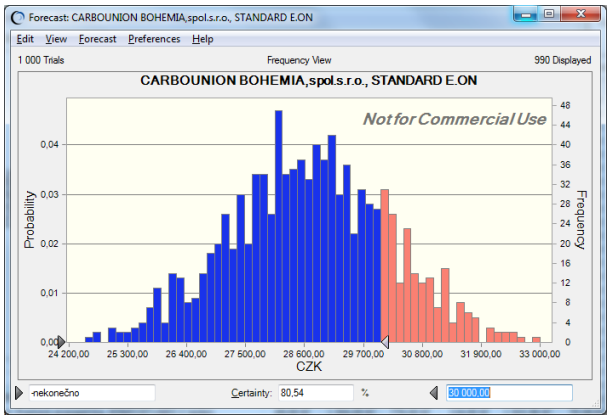


Figure 11: Histograms for the best product in 2017 and the probability of the costs lower than 30000 CZK

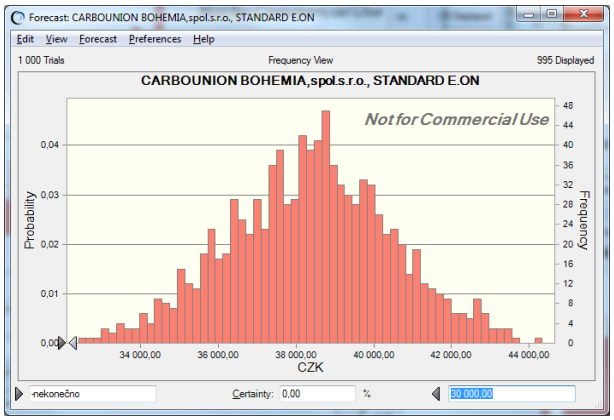


Figure 12: Histograms for the worst product in 2018 and the probability of the costs lower than 30000 CZK

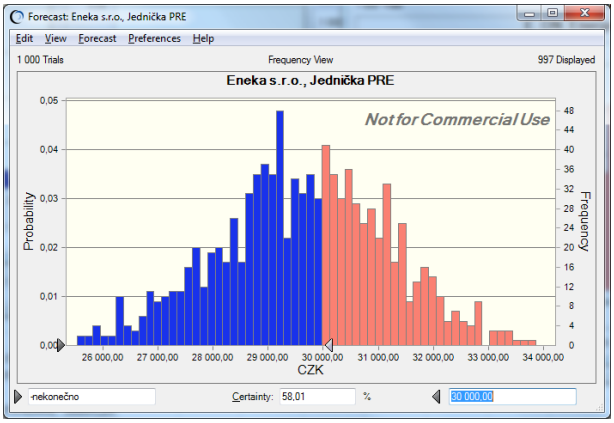


Figure 13: Histograms for the best product in 2018 and the probability of the costs lower than 30000 CZK

According to the results in 2017 and 2018, we see a big difference in the best (cheapest) products and also in the annual costs as well. When we use the data and results of our previous analysis (Kuncova 2015; Kuncova 2016; Kuncova, Seknickova 2015), the increase in annual costs is also evident (Figure 14) especially for the most expensive products which are on the same level as in the year 2013. On the other hand, the annual costs of the cheapest products are closed to the results from the year 2015. The differences between the cheapest and the most expensive products increased in 2018 compared with the previous years.

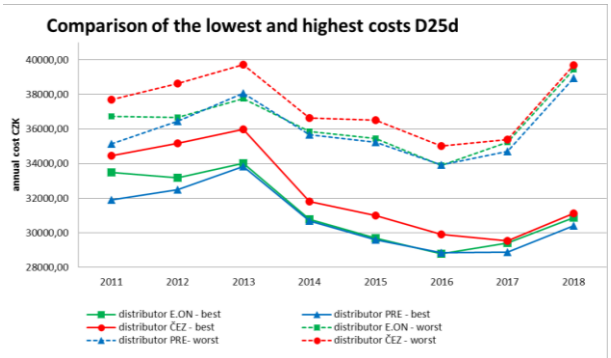


Figure 14: The time development of the highest and lowest annual cost of electricity savings

Table 5: The best products in the given years for E.ON

year	place	distributor E.ON
2014	1.	CARBOUNION KOMODITY
	2.	ELIMON eProdukt
	3.	Amper Market
2015	1.	CARBOUNION KOMODITY
	2.	ST Energy standard
	3.	Nano Energies Trade
2016	1.	ST Energy, Standard AKU 8
	2.	Amper Market, HOME_AKU
	3.	FOSFA, FEE e-TARIF
2017	1.	CARBOUNION BOHEMIA,spol.s.r.o., STANDARD
	2.	CARBOUNION BOHEMIA,spol.s.r.o., STANDARD 12
	3.	Nano Energies Trade s.r.o., Dobrý skutek
2018	1.	Eneka s.r.o., Jednička
	2.	Europe Easy Energy, eDOMÁCNOST
	3.	Eneka s.r.o., STANDARD

The differences are not only in the final annual costs but also in the best (cheapest) products. Table 5 shows the top 3 products in the five years in the E.ON distribution area. It is evident that nearly each year (except 2014 and 2015) the winner is different. Indeed, the differences between the annual costs may not be too large, but based on this analysis, it is advisable to analyze suppliers' costs and to use Monte Carlo simulation, which can show how much the annual cost of the selected products varies.

Many households still think that it is not necessary to change electricity suppliers and that due to the variable annual electricity consumption, the cheapest product cannot be estimated. However, Monte Carlo simulation is able to do this, and the results show that switching suppliers/products can save a considerable sum of money.

5. CONCLUSION

The situation on the electricity retail market in the Czech Republic is changing every year. All households can choose its supplier and the product based on their annual electricity consumption. The possibility of changing the supplier is also influenced by the setting up of a contract with the relevant supplier, where for a fixed contract length it is usually possible to switch to another supplier only at the price of the contractual penalty, which is quite high. As the annual electricity consumption is usually not fixed for each year, the Monte Carlo simulation model could be a good tool to use to be able to compare the products better. According to the results of this kind of simulation for the years 2017 and 2018 it is evident that the electricity prices were higher in 2018 and the increase in the annual electricity consumption costs in 2018 compared with 2017 could be higher than 25%. As the winning product of 2017 fell to the last place in 2018, it is evident that it is recommended to review annually the appropriateness of choosing a supplier and consider a possible change. Staying with the same supplier can increase the electricity consumption cost significantly.

The presented analysis was based on the real case study and related to the D25d tariff rate data. The conclusions are connected with this study and cannot be generalized for other tariff rates for which a similar analysis would be necessary. Equally, it can be expected that the savings in electricity supplier switching at low consumption levels will be different and probably lower. Again, it would be necessary to recalculate it. Anyway, with the knowledge of the products offered, it is always possible to use, for example, the Monte Carlo simulation to test for potential savings in different energy consumption situations.

Although this study is connected with the Czech conditions, the liberalization of the European electricity market is still work in progress (Boltz, 2013) and so similar analysis could be made for other countries with respect to their market conditions and rules.

ACKNOWLEDGMENTS

The research project was supported by the grant No. F4/66/2019 of the Faculty of Informatics and Statistics, University of Economics, Prague.

REFERENCES

Boltz, W. 2013. The Challenges of Electricity Market Regulation in the European Union. In. Sioshansi, F.P., eds. Evolution of Global Electricity Markets. Elsevier Inc. 199-224.

Brandimarte, P., 2014. Handbook in Monte Carlo Simulation – Applications in Financial Engineering, Risk Management, and Economics. USA, John Willey & Sons

ERU - Energy Regulatory Office, 2019. Price calculator. Available from: http://kalkulator.eru.cz/VstupniUdaje.aspx [accessed 26 March 2019]

ERU – Energy Regulatory Office, 2016. Yearly Report on the Operation of the Czech Electricity Grid. Available from: https://www.eru.cz/documents/10540/462820/Annual_report_electricity_2016.pdf/bfbc88b7-a814-498f-99bf-98b50148c805 [accessed 26 March 2019].

Eurostat, 2019a. Consumption-electricity-sector. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Consumption-electricity-sector-GWh-EU-28-1990-2016.PNG [accessed 26 March 2019].

Eurostat, 2019b. Electricity production, consumption and market overview. Available from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Electricity_production,_consumption_and_market_overview#Household_electricity_consumption [accessed 26 March 2019].

Hegazy, Y.G., Salama, M.M.A. and Chickhani, A.Y.: Adequacy assessment of distributed generation systems using Monte Carlo Simulation. Power Systems - IEEE Transactions 18/1 (2003), 48-52.

Karanfil, F. and Li, Y. 2015. Electricity consumption and economic growth: Exploring panel-specific differences. Energy Policy, 82, 264-277.

Kuncova, M, 2016. Analysis of the Electricity Market from the Perspective of Products Offered to Households in the Tariff Rate D25d in Period 2011-2016. ACC Journal, XXII/2, 45–56. Available from: http://acc-ern.tul.cz/images/journal/sbornik/ACC_JOURNAL_2016_B.pdf [accessed 10 April 2019].

Kuncova, M., 2015. Methods for the Electricity Supplier Selection - Case Study of the Czech Republic. International Journal of Mathematical Models and Methods in Applied Sciences, 9, 714-720.

Kuncova, M., Lizalova, L. 2012. Monte Carlo simulation – the bank account selection in the Czech Republic according to the bank charges. Proceedings of the International Conference on

Modeling and Applied Simulation (MAS). Wien, 2012, 153-161.
Kuncova, M., Seknickova, J. 2014. Analysis of the efficiency of the electricity supplier selection depending upon the price changes. Proceedings of the Conference Mathematical Methods in Economics pp. 542-547. September 10-12, Olomouc (Czech Republic).
Kuncova, M., Seknickova, J. 2015. Monte Carlo simulation used for the electricity supplier selection. Proceedings of the 14th International Conference on Modeling and Applied Simulation (MAS), pp. 105-109. September 21-23, Bergeggi (Italy).
Newbery, D. 2013. Evolution of the British Electricity Market and the Role of Policy for the Low-Carbon Future. In. Sioshansi, F.P., eds. Evolution of Global Electricity Markets. Elsevier Inc. 3-30.
Oracle Crystal Ball 2019. Crystal Ball. Available from: https://www.oracle.com/cz/applications/crystalball/ [accessed 7 July 2019]
OTE – Market operator, 2018. Statistics – changes of suppliers. Available from: https://www.ote-cr.cz/cs/statistika/mesicni-zprava-elektrina/zmeny-dodavatele?date=2018-01-01 [accessed 26 March 2019]
Razgaitis, R. 2003. Dealmaking Using Real Options and Monte Carlo Analysis. John Willey & Sons, New Jersey, 2003.
Turban , E., Meredith, J., R., 1994. *Fundamentals of Management science.* 6th ed. USA, Richard D.Irwin Inc.
Ventosa, M., Baíllo, Á., Ramos, A. and Rivier, M. 2005. Electricity Markets modeling trends, Energy Policy 33, 897–913.
Vlcek, T., Cernoch, F. 2013. The Energy Sector and Energy Policy of the Czech Republic. Massaryk University 2013. Available from: https://is.muni.cz/repo/1130969/The_Energy_Sector_and_Energy_Policy_of_the_Czech_Republic__Vlcek__T._-_Cernoch__F.__2013_.pdf [accessed 7 July 2019]

AUTHORS BIOGRAPHY

Martina Kuncova: She has got her degree at the University of Economics Prague, Czech Republic, at the branch of study Econometrics and Operational Research (1999). In 2009 she has finished her doctoral study at the University of West Bohemia in Pilsen, Czech Republic (Economics and Management). Since the year 2000 she has been working at the Department of Econometrics, University of Economics Prague, since 2007 also at the Department of Economic Studies of the College of Polytechnics Jihlava (Czech Republic). She is a member of the Czech Society of Operational Research, she participates in the solving of several grants. She is the coauthor of four books and the author of many scientific papers and contributions at conferences. She is interested in the usage of the operational research, simulation methods and methods of multi-criteria decision in real life problems.

STUDY ABOUT THE RIGOR OF TIGHTNESS TESTS AND THE EXISTANCE OF AN EMPIRICAL CONSTANT TO VALIDATE THEORETICAL SIMULATIONS TO APPROXIMATE THE REAL MODEL

João Arthur da Cruz Nunes[(a)], Angelo Roncalli Oliveira Guerra[(b)], Kleiber Lima de Bessa[(c)], Carlos Magno de Lima[(d)]

[(a),(b),(c),(d)]Universidade Federal do Rio Grande do Norte

[(a)]j.ac.n@hotmail.com, [(b)]profroncalli@gmail.com, [(c)]klbessa@ct.ufrn.br
[(d)]cmagno@ufrnet.br

ABSTRACT

Due to the great risk of contamination by leaking in underground fuel storage tanks (UST) of gas stations all over the world, the establishment of effective monitoring methods in this environment is extremely necessary. Among UST monitoring methods the tightness test is one of the most effective ones in identifying leaks, it can be done in two different ways, either wet part test or dry part test. But while both of the tests are permitted, they show a great difference in rigorousness, when it comes to approving or not a tank. This study envisions to deeply explore the causes of the difference of rigorousness between both tests, and discover ways in which simulations can approach the real situation. The research allowed us to identify not only the cause of such difference in rigor, but also to establish a constant that approximates the theory to the real situation.

Keywords: tank tightness, simulation, evaporation, vacuum.

1. INTRODUCTION

The notion of engineering and development being linked to sustainability and environment is an ever-growing trend. Researches with the goal of developing technologies to monitor and reduce pollution have been growing and gaining more and more space and visibility in the engineering world, as the connection between progress and environment has already become a necessity, rather than a luxury.

The vast number of leaking cases in gas stations, coming from USTs and piping systems, have caused damage to the environment, besides also harming the security, health and life quality of the population around these sites (SANDRES et al., 2002). This contamination can affect not only the soil, but also groundwater and cause explosion and fire risk.

In Brazil, there are more than 40 thousand gas stations and soil contamination is currently one of the main concerns, once around 30% of these stations present problems that can cause contamination. Multiple companies deal daily with the unforeseen coming from leaking and incorrect storage (TERRA BRASIL, 2017).

According to a research made by CETESB, gas stations are the main responsible for soil contamination in cities like São Paulo. The contamination occurs by leaking of fuel and gases due to bad installation of the USTs, which are fabricated in steel and do not possess protection against corrosion. Currently, multiple companies have invested in safer tanks, as seen that the investment in better installations it's much smaller than the costs of repairs for the problems caused by the contamination of the water and the soil (TERRA BRASIL, 2017).

Soil contamination by fuel its a big concern, seen that the fuel contains Benzene, Toluene, Ethylbenzene and Xylenes(BTEX) in its composition, that are all harmful to human health and can cause dangerous diseases. Soil contamination occurs when there isn't proper investment in quality equipment, mainly when it comes to metallic pipes and purely metallic tanks, because they are underground and it is necessary to pay attention to signs of corrosion and possible failures (TERRA BRASIL, 2017).

In a gasoline spill, one of the main concerns is the contamination of aquifers that are used as source of water for human consumption (TEIXEIRA, 2008). Due to the fact that it is very little soluble in water, spilled gasoline, containing more than 400 components, initially will be underground as a liquid of non-watery phase. In contact with underground water, gasoline will partially dissolve. The mono-aromatic hydrocarbons: benzene, toluene and xylenes, called BTEX are the components present in the gasoline that have the highest water solubility, thus, they are the first contaminants to reach the groundwater. These compounds are considered dangerous substances because they are depressor of the central nervous system. The benzene is proven carcinogenic, being able to cause leukemia (TEIXEIRA, 2008).

The gasoline commercialized in Brazil is mixed with alcohol in proportions that can go from 20% to 30%, according to the current legislation. That makes it different from the gasoline sold in other countries, where it isn't mixed with oxygenated compounds. The interaction between ethanol and BTEX can cause a rise

in mobility and solubility, while also hindering the natural biodegradation of these compounds (TEIXEIRA, 2008).

The monitoring ends up being one of the main weapons to avoid contamination. The monitoring equipment for leak detection helps avoiding possible contaminations to the environment and accidents at work. Amongst USTs monitoring methods, the tightness test is one of the most effectives in the identification of possible leaks, being able to be executed in two different ways: the wet part test(i.e. part filled with fuel) and the dry part test.

This study compared the results obtained from the application of both types of test. The motivation for this paper came from the necessity of supplying scientific evidence conjugated with experimental data and theoretical concepts, some laboratory tests and real tank testing practice on site (gas stations) to better compare the rigor of the methods of leak detection regulated in Brazil by the standard ABNT 13784 in force and understand the causes of the difference found in the rigor of both methods while finding a way to approximate theoretical simulations to the practical results. This way, on site comparisons were made where, initially a wet part test was performed, and following that, the same tank was emptied and a dry part test was performed on the same tank. The research allowed us to identify a better performance in terms of rigor from the dry part test and confirmed the initial suspicion of the great influence of the evaporation rate of the fuel on the testing of partially full tanks (i.e. hybrid situation where there is a coexistence of a dry and a wet part in the tank). Based on the results obtained in both the theoretical simulations and practical tests we were able to find a constant that helps correcting any false positives and approximates both situations. We have proved that a well-made simulation(with the constant) implies on a more realistic model, and that was validated.

2. PETROBRAS' COMMON GASOLINE

On this article, common gasoline was used as the study object, due to it having the highest evaporation rate among the liquid fuels commercialized by ANP, thus, the parcel of pressure increase allowed by a possible hole in the tank is lower, so the test executed with this fuel is essentially the most critic case.

Gasoline belongs in the group of the LNAPL (Light Non Aqueous Phase Liquids) e PMOS (Partially Miscible Organics Solubility) Mindrisz, et al., (2006). Derivative from petroleum, gasoline is composed by innumerable chemical compounds (olefins, aromatichydrocarbons etc.) among which stand out as the most water-soluble contaminants, the BTEX compounds, present in 18% of the gasoline weight. A particularity differentiates Brazilian gasoline from the ones in other nations, the considerable presence of ethanol, which currently corresponds to 27% of the volume (Portal G1, March 2016). This factor makes it possible for similar studies involving other countries' gasolines to be able to present different results when compared to the ones made in Brazil.

Table 1: Main properties of this fuel at 1 atm

Properties	
Starting Boiling Point	>35°C
Flash Point	<-43°C
Partition Coefficient – noctanol/water	Insoluble in water. Soluble in organic solvents
Auto-Ignition Temperature	Log kow 2-7
Density	0.73-0.77
Viscosity	0,6x10-6 m2/s
Volumetric thermal dilation coefficient	1,2 x 10-3 °C-1

3. TIGHTNESS TEST: METHODOLOGY, THORETICAL FOUNDATION AND EQUATIONS.

3.1. Equations for Wet Part Tightness Test

According to the international standards and the standard ABNT 13784 in force in Brazil, the wet part test requires that, considering the tank out of operations (i.e. sales over) and the liquid at rest. The behavior of the volume of liquid inside the tank shall be observed for one uninterrupted hour and verified to see if the volumetric variation has not exceeded the permitted limit in that time. Thus, for academic purposes and to make comparisons between laboratorial and on site tests, a computational program was developed, capable of calculating the diameter of an hypothetical hole necessary for the limit situation to occur, which is 378 ml(0.1 gallon) "leaked" in one hour. Any loss of liquid volume that surpasses this limit is enough to deem the tank unfit. Thus the flow rate limit of 378 ml/h is used as input data on the implemented algorithm. Methodology illustration (figure1) and equations subsequently described.

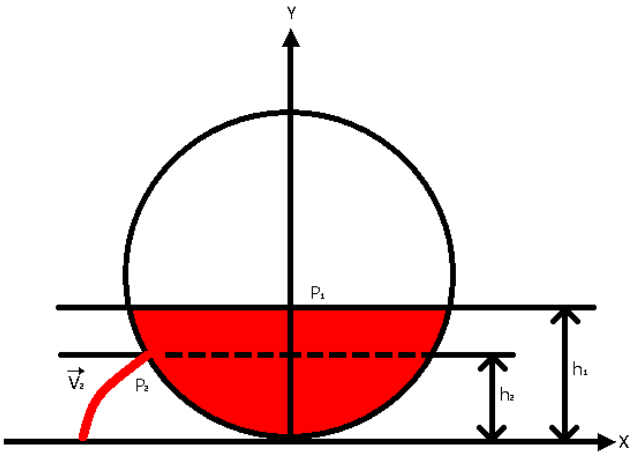


Figure 1 - Schematic representation of the methodology used to detect leaks in the wet part of a tank

Adopting a conservative approach, simplifying and considering, hypothetically, the fluid to be incompressible, non-viscous and a stationary flow(i.e. permanent flow can be considered due to the negligible variations of the water levels to obtain the admissible flow rate)., Using Bernoulli the following equations can be reached:

$$\frac{V_1^2}{2} + g * h_1 + \frac{P_1}{\rho} = \frac{V_2^2}{2} + g * h_2 + \frac{P_2}{\rho} \quad (1)$$

Where V_1 e V_2 are the velocities, P_1 **and** P_2 are the pressures, ρ is the specific mass of the fluid, h_1 is the height of the fuel at the start of the test and h_2 is the height of the theoretical hole.

It's observed on fig.1 that the pressures P_1 e P_2 are equal and manometric for the calculations, because they're in contact with the air and that the velocity V_1 is negligible, because the volume of liquid inside the tank is much bigger than what is leaking. Executing the algebraic manipulations necessary to isolate V_2:

$$V_2 = \sqrt{2 * g * (h_1 - h_2)} \quad (2)$$

To find out the diameter of the hole, the volumetric flow rate formula was used ($\dot{\forall}$):

$$\dot{\forall} = V * A \quad (3)$$

Where V is the flow velocity and A is the area of the hole described by:

$$A = \pi * \frac{d^2}{4} \quad (4)$$

Inserting eq.3 and eq.4 into eq.2:

$$\dot{\forall} = \sqrt{2 * g * (h_1 - h_2)} * \pi * \frac{d^2}{4} \quad (5)$$

Isolating d, the equation of the necessary hole diameter is reached. Given the heights of the hole and the fuel for a volumetric flow rate of 378 ml/h, the limit value for the permitted hole diameter is found:

$$d = \sqrt{\frac{4 * \dot{\forall}}{\sqrt{2 * g * (h_1 - h_2)} * \pi}} \quad (6)$$

3.2. Equations For Dry Part Tightness Test

For the execution of this type of test, a pump is coupled to the tank's breather, sealing all possible air inlets and cracks. With sales already over, as required per the ABNT 13784 standard, considering the tank completely sealed, the pump starts sucking air out of the tank untill a preassure drop between 90 to 100 mmHg is reached and then the pump is turned off. This procedure is repeated until after the pump's turning off the pressure drop value maintains itself stable inside the aforementioned range. After that te tank is put under observation for 30 minutes. During this time, the pressure raise cannot surpass 10mmHg (current standard) or 15mmHg (in study standard). In any case that the used limit value for the test is exceeded, the tank is deemed unfit. Thus, equally to the previous case of the wet part, for academic purposes, a computational program was developed, capable of calculating the diameter of the hypothetical hole necessary for the limit of pressure raise to be achieved. The procedures of the implemented algorithm, illustration of the methodology (fig.2) and equations are described next.

3.2.1. Dry Part Test on A Completely Empty Tank (i.e. only air)

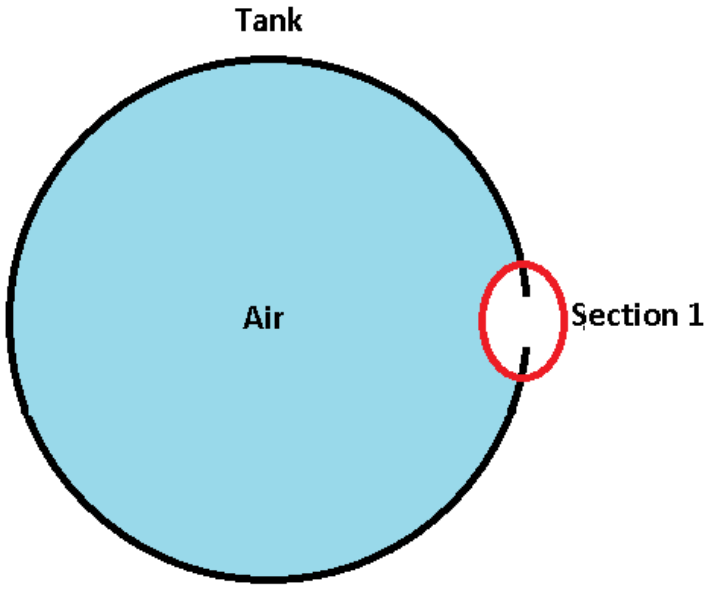


Figure 2: Schematic representation of the methodology to detect leaks on the dry part of a completely empty tank.

Initial considerations adopted:

(1) The air behaves as an ideal gas at pressures below 30 atm.
(2) The properties of air in the tank are uniform, but time dependent.
(3) Incompressible flow.

The continuity equation was used to approach the problem:

$$\frac{\partial}{\partial t}\int_{CV} \rho d\forall + \int_{CS} \rho \vec{V} d\vec{A} = 0 \quad (7)$$

Where: The first term represents the mass variation rate inside the control volume and the second term

represents the liquid rate of mass flow to the outside trough the control surface.
Once the properties in the tank are considered uniform, the specific mass (ρ) can be taken out of the integral:

$$\frac{\partial}{\partial t}\left[\rho_{CV}\int_{CV} d\forall\right] + \int_{CS}\rho\vec{V}d\vec{A} = 0, where, \int_{CV} d\forall = \forall \quad (8)$$

So,

$$\frac{\partial}{\partial t}(\rho\forall)_{VC} + \int_{SC}\rho\vec{V}d\vec{A} = 0 \quad (9)$$

The only place where mass crosses the control volume's boundary is at section 1, this way:

$$\int_{CS}\rho\vec{V}d\vec{A} = \int_{A1}\rho\vec{V}d\vec{A} \quad e \quad \frac{\partial}{\partial t}(\rho\forall)_{CV} + \int_{A1}\rho\vec{V}d\vec{A} = 0 \quad (10)$$

On the surface of section 1, the sign of $\rho\vec{V}d\vec{A}$ is negative,

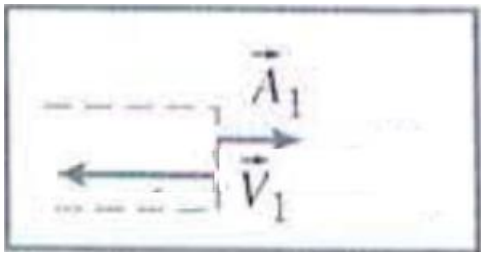


Figure 3: Surface 1

Thus,

$$\frac{\partial}{\partial t}(\rho\forall)_{CV} - \int_{A1}\rho V dA = 0 \quad (11)$$

As the flow is considered uniform on surface 1,

$$\frac{\partial}{\partial t}(\rho\forall)_{CV} - \rho_1 V_1 A_1 = 0 \quad or \quad \frac{\partial}{\partial t}(\rho\forall)_{CV} = \rho_1 V_1 A_1 \quad (12)$$

Once ∀ (tank's volume) isn't a function of time,

$$\forall\frac{\partial\rho}{\partial t} = \rho_1 V_1 A_1 \quad (13)$$

Isolating $\partial\rho/\partial t$, :

$$\frac{\partial\rho}{\partial t} = \frac{\rho_1 V_1 A_1}{\forall} \quad , where, A_1 = \pi * \frac{{d_1}^2}{4} \quad (14)$$

Where,

$\frac{\partial\rho}{\partial t} = Specific\ mass\ variation\ rate\ in\ time$
$\rho_1 = Specific\ mass$
$\forall = Tank's\ volume$
$V_1 = Air\ inlet\ velocity\ into\ the\ tank$
$A_1 = Section\ 1\ area$
$d_1 = Hole's\ diameter$

At this point the concept of stagnation pressure was used to find the velocity V_1 with which the air flows into the tank. Thus for an arbitrary point inside the tank, with the subscript "0" representing the stagnation conditions:

$$P_0 = P + \frac{\rho V^2}{2} \quad (15)$$

Isolating the velocity:

$$V = \sqrt{\frac{2*(P-P_0)}{\rho}} \quad , onde, \quad \rho = \frac{P[Pa]}{R_{air}[J*Kg^{-1}*K^{-1}]*T[K]} \quad (16)$$

Starting here, the subscript t_0 will be used to indicate the start of the test and $t_0 + \Delta t$ to indicate the end of it.
The velocities were obtained through the developed equations. Thus, for a vacuum (induced pressure drop) of 100 mmHg and a temperature of 27 C (300 K):

$$\rho_{t_0} = \frac{87992.76}{287*300} = 1.022 \quad (17)$$

$$V_{t_0} = \sqrt{\frac{2*(87992.6-101325)}{1.022}} = 161.257\frac{m}{s} \quad (18)$$

And for an increase of 10 mmHg of pressure according to the in vigor standard:

$$\rho_{t_0+\Delta t} = \frac{89325.99}{287*300} = 1.037 \quad (19)$$

$$V_{t_0+\Delta t} = \sqrt{\frac{2*(89325.99-101325)}{1.037}} = 152.09\frac{m}{s} \quad (20)$$

For a velocity of the sound on air at 27 C, $c = \sqrt{kRT}$,
Where: K is the Volumetric Elasticity Module of the air:

$$c = \sqrt{1.4 * 287 * 300} = 347.189\frac{m}{s} \quad (21)$$

And the Mach numbers,

$$M_{t_0} = \frac{161.257}{347.189} = 0.46524 \quad (22)$$

$$M_{t_0+\Delta t} = \frac{152.09}{347.189} = 0.438 \quad (23)$$

The Mach numbers obtained (M>0.3) suggest that the flow is actually compressible, different from what was initially assumed.
Correcting the approach to a compressible flow the stagnation conditions for a compressible flow are:

$$P_0 = P * \left(1 + \frac{k-1}{2} + M^2\right)^{\frac{k}{k-1}} \quad (24)$$

And

$$\frac{T_0}{T} = 1 + \frac{k-1}{2} * M^2 \quad (25)$$

Isolating the Mach number in the stagnation pressure formula:

$$M = \frac{\sqrt{2} * \sqrt{P_0 * \left(\frac{P_0}{P}\right)^{\frac{1}{k}} - P}}{\sqrt{k * P - P}} \quad (26)$$

The new Mach numbers will be:

$$M_{t_0} = \frac{\sqrt{2} * \sqrt{101325 * \left(\frac{101325}{87992.76}\right)^{\frac{1}{k}} - 87992.76}}{\sqrt{1.4 * 87992.76 - 87992.76}} = 0.4535 \quad (27)$$

$$M_{t_0+\Delta t} = \frac{\sqrt{2} * \sqrt{101325 * \left(\frac{101325}{89325.99}\right)^{\frac{1}{k}} - 89325.99}}{\sqrt{1.4 * 89325.99 - 89325.99}} = 0.4282 \quad (28)$$

To find the flow velocities it is necessary to obtain the sound velocity and the temperature inside the tank.
Thus, using the stagnation condition of the temperature:

$$T = \frac{T_0}{1 + \frac{k-1}{2} * M^2} \quad (29)$$

For the start and the end of the test:

$$T_{t_0} = \frac{300}{1 + \frac{1.4-1}{2} * 0.4535^2} = 288.15\ K \quad (30)$$

$$c_{t_0} = \sqrt{1.4 * 287 * 288.15} = 340.26\ \frac{m}{s} \quad (31)$$

$$T_{t_0+\Delta t} = \frac{300}{1 + \frac{1.4-1}{2} * 0.4282^2} = 289.39\ K \quad (32)$$

$$c_{t_0+\Delta t} = \sqrt{1.4 * 287 * 289.39} = 340.99\ \frac{m}{s} \quad (33)$$

In possession of the Mach numbers e the respective sound velocities it is possible to calculate the inlet velocities of air into the tank:

$$V_{t_0} = c_{t_0} * M_{t_0} = 154.31\ \frac{m}{s} \quad (34)$$

$$V_{t_0+\Delta t} = c_{t_0+\Delta t} * M_{t_0+\Delta t} = 146.01\ \frac{m}{s} \quad (35)$$

Back to the continuity equation:

$$\frac{\partial \rho}{\partial t} = \frac{\rho_1 V_1 \pi {D_1}^2}{4 \forall} \quad (36)$$

The time of the test can be represented by:

$$t = \frac{\Delta \rho}{\frac{\partial \rho}{\partial t}} \quad (37)$$

Where,

$$\Delta \rho = \rho_{t_0+\Delta t} - \rho_{t_0} \quad (38)$$

With,

$$\rho_{t_0} = \frac{P_{t_0}}{R * T_{t_0}} = \frac{87992.76}{287 * 288.15} = 1.064 \quad (39)$$

$$\rho_{t_0+\Delta t} = \frac{P_{t_0+\Delta t}}{R * T_{t_0+\Delta t}} = \frac{89325.99}{287 * 289.39} = 1.076 \quad (40)$$

Having $\Delta\rho$, it is possible to find the diameter of the hole in an iterative way, varying it until the time of the test reaches the desired 1800 seconds (30 minutes).

3.2.2. Dry Part Test on A Semi Empty Tank

This is a very common situation, which corresponds to a hybrid case with simultaneous coexistence of both wet part and dry part. The standardizing and procedures adopted to test tightness of the dry part of partially filled tanks is similar to the one described in the previous section (i.e. completely empty tank).However, a fundamental detail has to be considered in this case. As there is fuel present in the tank, part of the pressure increase will happen simply due to the evaporation rate of the fuel during the test, with this, not necessarily indicating a leaking through a hole or something similar. By neglecting this fact, the testing companies risk mistakenly deeming a tank unfit without knowing the real cause of the vacuum drop (pressure increase).
This way, the challenge question that comes with applying this method to semi empty tanks is the following: A result that deems a tank non-tight and unfit has that result because of a real hole or simply due to fuel evaporation at low pressures generating a false positive of untightness? To help solve this problem, it is necessary to determine the parcel of the pressure raise in the tank caused only by atmospheric air inlet trough possible holes in the tank and, this way, distinguish with reliability the two possible sources of pressure increase. Throughout the research process, multiple laboratory tests were made simulating diverse situations, including on site tests contemplating real empty tanks from gas stations, which were certified tight. As a result of these experiments an empiric constant was obtained. It represents the median pressure increase in a tank on critical temperature conditions due to natural reasons (i.e. a perfect vacuum doesn't exist) and leftover gas after the emptying process of the tank. The value found was between 9 and 10 mmHg. This empirical constant was used on this paper as the way to correct possible distortions in the application of the method and due to theoretical approximations, besides other ideal laboratory considerations applied to the equations of the study. Ratifying the fact that a perfect vacuum does not exist. Thus, the empirical constant above served to

attenuate the main differences between theoretical conditions and the real situations on the gas stations.
To validate the computational simulation and the laboratory tests trough practical dry part tests of semi empty tanks, the following considerations were taken:

(1) At the start of the test, the control volume is completely occupied by gasoline vapor, as it is denser than air, thus, it tends to expel it.
(2) The control volume is the dry part of the tank not filled with fuel, and the variations on its dimensions are negligible.
(3) A maximum natural increase of 10mmHg(empiric constant) is expected on the simulation and happens, even if there are no detectable holes to allow air inlet or increase due to evaporation rate of the fuel throughout the duration of the test.

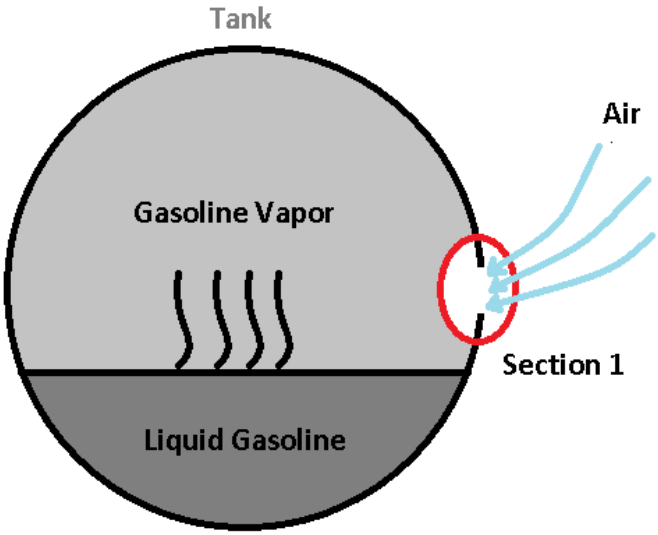


Figure 4 - Schematic representation of the semi empty tank at the start of the test

For this analysis, the Dalton law of partial pressures was used.

$$P_{Total} = P_{Air} + P_{GV} \quad (41)$$

$$P_{Air} * = \eta_{Air} * \frac{RT}{\forall}, air's\ partial\ pressure \quad (42)$$

$$P_{GV} = \eta_{GV} * \frac{RT}{\forall}, fuel\ vapour's\ partial\ pressure \quad (43)$$

As the objective is to know the parcel of the pressure increase for which the air inlet is solely responsible for, the equation to be used is:

$$P_{Air} = \eta_{Air} * \frac{RT}{V} \quad (44)$$

Which can be expanded as:

$$P_{Air} = \frac{m_{Air}}{M_{Air}} * \frac{RT}{V} \quad (45)$$

As the control volume does not contain any air at the start of the test, it is needed to find the final mass of air in the tank, which means, how much air entered the tank during the test.
To exemplify and validate the methodology and make the due approximations, it will be considered a tank with volumetric dimensions equal to the volume of the parcel of the tank not occupied by liquid gasoline (Ex.: for a tank with 30 thousand liters with 76 cm of column of liquid fuel, this theoretical volume will be of 22947 cubic meters) full of air. As the volume doesn't change, the variation of the mass of air will be given by:

$$\Delta m = \rho_{t_0+\Delta t} * \forall - \rho_{t_0} * \forall = Final\ Mass \quad (46)$$

The volume $\forall$ is the parcel of tank's volume that is not occupied by liquid gasoline.
In possession of the value corresponding to the mass of air that entered the tank the Eq. (45) can be used to calculate the increase of pressure by air inlet into the tank. This was always the procedure adopted on this research for multiple simulations e practical validation.
The temperature is the one at the end of the 30 minutes of the test. In this specific practical example the temperature was measured with the aid of well calibrated equipment of Veeder Root brand installed on the 30 de setembro gas station in Natal,RN, Brazil. The utilized tank in the simulation was of 30,000 liters with 76 cm of liquid column (common gasoline). The control volume (the volume not occupied by liquid fuel) corresponds to 22,497 liters.

Figure 5: Veeder Root Equipment

Having the value of pressure increase for which is responsible solely the inlet of air in the control volume, which is exactly what is necessary to certify tightness, a test situation of dry part test on an empty tank(only air) with a volume of 22,947 liters and a limit pressure raise corresponding to the pressure increase only by air inlet is simulated.

The approach used analyzing only the air is justified by the difficulties encountered in analyzing the pressure raise caused only by fuel evaporation (P_{VG}), due to the scarcity of data and lack of depth in studies about the properties of the gasoline in pressure conditions different from the atmospherical. Given these difficulties, the analysis with the Dalton law allows an approximation where only the air properties are necessary, these being easily accessible.

4. GENERAL RESULTS OF THE SIMULATIONS

The following tables present the data regarding the results obtained for the smallest hole diameter than can be identified by a specific test situation and the variables used in the simulation.

All the data from the simulations were obtained using the computational programs developed by the team simulating each of the previously detailed models (Wet part, empty dry part and semi-empty dry part).

Table 2: Results for the simulation of a wet part test with a completely full tank

Wet Part Test, Completely full tank (Theoretical)	
Volume of the Tank(m³)	30.607
Liquid Column Height(cm)	254
Volume of liquid(m³)	30.607
Flow Rate(l/h)	0.378
Gravity(m/ s2)	9.81
Diameter of the hole(mm)	0.1375

Table 3: Results for the theoretical ideal simulation of a dry part test with a completely empty tank

Dry part test, completely empty tank(Theoretical) (Neglecting empirical constant)	
Variation of specific mass(kg/m³)	0.017217
Volume of the tank(m³)	30.607
Initial Air Inlet Velocity(m/s)	154.3069
Initial Specific Mass(kg/ m³)	1.064019
Specific mass variation rate Calculated from the equation in the integral form (kg/m³/s)	9.53E-06
Time(s)	1800
Diameter of the hole(mm)	1.302

Table 4: Results for the theoretical simulation of a dry part test with a completely empty tank considering natural reasons

Dry part test on a completely empty tank(Theoretical) Considering the empirical constant	
Allowed pressure increase(mmHg)	15
Empirical constant(mmHg)	10
Variation of specific mass(kg/m³)	0.017217
Volume of the tank(m³)	30.607
Initial Air Inlet Velocity(m/s)	154.3069
Initial Specific Mass(kg/ m³)	1.064019
Specific mass variation rate Calculated from the equation in the integral form (kg/m³/s)	9.53E-06
Time(s)	1800
Diameter of the hole(mm)	0.753

Table 5: Results for the theoretical simulation of a wet part test with a semi empty tank

Wet Part Test, semi empty tank (Theoretical)	
Volume of the Tank(m³)	30.607
Liquid Column Height(cm)	76
Volume of liquid(m³)	7.660
Flow Rate(l/h)	0.378
Gravity(m/ s2)	9.81
Diameter of the hole(mm)	0.186

Table 6: Results for the real analysis of a dry part test with a semi empty tank

Dry part test, semi empty tank (Real)	
Mass Variation(Kg)	0.131979
Final temperature(K)	292.7
M(Molar mass) of the air(kg/Mol)	0.029
n(number of mos) of Air	0.102979
Allowed pressure increase(mmHg)	15
Empirical constant(mmHg)	10
Pressure increase by air inlet(Pa)	10.92
Diameter of the hole(mm)	0.097

With the obtained results, it is possible to verify the clear effect of the evaporation rate on the total pressure increase during the test. When comparing the smallest identifiable hole of the dry part test with a semi empty tank versus a completely empty one, the 0.753 mm from the latter is approximately 8 times bigger, thus less rigorous, than the 0.097 mm found on the former. It is also possible to state the superiority of the dry part test when compared to the wet part one, as seen that in the same situation (semi empty, 76 cm of liquid column of gasoline) the dry part test showed itself two times more rigorous than its rival, thus, much stricter when judging the integrity of the tank.
The difference in results between the tests using and neglecting the empirical constant is alarming, and shows how important its use is to approximate the simulation to the real situation.

5. CONCLUSIONS

This study verifies the suspicions of the great difference in rigorousness between the two types of tightness test for USTs in gas stations allowed by the in force standard in Brazil.

The results of the study confirmed the superior rigorousness of the dry past test. It also showed the necessity to adopt an empirical constant (obtained through extensive practical tests) to approximate the theoretical simulations and the real tests. This constant being the natural raise of pressure in any tank due to natural reasons, something that also allows us to reinforce the inexistence of a “perfect vacuum”. The results were also capable of confirming the initial suspicion of the great influence of the evaporation rate of the fuel on the total pressure increase throughout the duration of the test**.**
A major hindering factor on the dry part analysis of a semi-empty tank was the lack of information about the properties of the gasoline on non-standard conditions(atmospheric pressure and average temperature). Deeper studies into the properties of this fuel would allow greater precision in determining the empiric constant and are an interesting prospect.

REFERENCES

André T.S., Fontes F.A.O., Guerra A.R.O. and Barbosa C.R.F., 2016. Taxa De Evaporação Da Gasolina Comum Em Tanques De Postos De Combustíveis: Uma Avaliação Em Função Da Temperatura. <Proceedings of the Conference>, pp.1-7. August 21-25, Fortaleza (Ceará, Brasil)..

Fox R.W., Pritchard P.J., Mcdonald A.T., 2010. Introduction to fluid mechanics. [n.l.]: Ltc.

Teixeira R.M., 2008. Postos de combustíveis – contaminação de aquíferos e solos por vazamento em tanques subterrâneos. CENED. Available from: <http://www.cenedcursos.com.br/meio-ambiente/postos-de-combustiveis-contaminacao-de-aquiferos/> [Acessed 05 October 2018].

Terra Brasil, 2017. Contaminação do solo de posto de combustível. Terra Brasil. Available from: <http://terrabrasilambiental.com.br/blog/contaminacao-do-solo-de-posto-de-combustivel/> [Acessed 05 October 2018].

POLLUTION ROUTING PROBLEM WITH TIME WINDOW AND SPLIT DELIVERY

Sameh M. Saad [(a)], Ramin Bahadori [(b)]

[(a),(b)] Department of Engineering and Mathematics, Sheffield Hallam University
Howard Street, Sheffield, S1 1WB, UK

[(a)] S.Saad@shu.ac.uk , [(b)] R.Bahadori@shu.ac.uk

ABSTRACT

In most classic vehicle routing problems, the main goal is to minimise the total travel time or distance while, the green vehicle routing problem, in addition to the stated objectives, also focuses on minimising fuel costs and greenhouse gas emissions, including carbon dioxide emissions. In this research, a new approach in Pollution Routing Problem (PRP) is proposed to minimise the CO_2 emission by investigating vehicle weight fill level in length of each route. The PRP with a homogeneous fleet of vehicles, time windows, considering the possibility of split delivery and constraint of minimum shipment weight that must be on the vehicle in each route is investigated simultaneously. The mathematical model is developed and implemented using a simulated annealing algorithm which is programmed in MATLAB software. The generated results from all experiments demonstrated that the application of the proposed mathematical model led to the reduction in CO_2 emission.

Keywords: vehicle routing, green Vehicle Routing, split delivery, simulated annealing.

1. INTRODUCTION

The Vehicle Routing Problem (VRP) is part of a series of problems that are associated with determining a set of routes in which each vehicle starts moving from a particular warehouse, serving a set of specified customers, and returning to the same warehouse. This problem was first introduced by Dantzig and Ramser (1959) and solved by mathematical methods. In the sequel, Laport et al. (1992) developed a variety of approaches of the branch and bound method to solve the vehicle routing problem. Clarke and Wright (1994) proposed a savings algorithm for solving VRP, which was the basis for many further kinds of researches. With recent advances in solving these problems, and taking into account more complex assumptions and constraints, metaheuristic methods such as genetic algorithm (Potvin and Bengio 1994), Tabu Search algorithm (Ho and Haugland 2004), ant colony optimization algorithms (Reimann et al. 2002), simulated annealing (Saad and Bahadori 2010) and particle swarm optimization (Norouzi et al. 2017) are developed.

In 1990, Split Delivery Vehicle Routing Problem (SDVRP) and its mathematical model were introduced and presented by (Dror and Trudeau 1990) in which the economic aspect of the problem is considered when a customer is served with more than one vehicle. Dror et al. (1994) provided an integer program for the above problem and used the branch and bound algorithm to solve it. The real application of this problem is studied by (Mullaseril et al. 1997) for a food distribution network at a dairy farm in Arizona, USA. In their work, the delivery of goods to the customer is associated with a time limitation and proposed heuristics algorithm by (Dror and Trudeau 1990) is applied. Belfior (2009) applied SDVRP in a case study in Brazil, for a distribution network consisting of a central warehouse and 519 customers in 11 sectors using Neighbourhood Search algorithm. Tavakkoli-Moghaddam and Safaei (2007) developed the Simulated Annealing algorithm for SDVRP with the heterogeneous fleet.

Transportation has irreparable effects on the environment; Consumption of resources, toxic effects on ecosystems and humans, noise and emissions of greenhouse gases (GHG) and pollutants are examples of these risks. Apart from the aforementioned negative effects, emissions of greenhouse gases and carbon dioxide (CO_2) are directly linked to the health of the community and indirect to the destruction of the ozone layer (Bektaş and Laporte 2011). Most research has taken into account economic goals by minimising the distance, the time required or the number of vehicles needed and has neglected attention to environmental goals. Hence, the Green Vehicle Routing Problem (GVRP) has received the attention of scholars since 2006. In the following, two categories including Green-VRP (G-VRP) and Pollution Routing Problem (PRP) are dominated which have focused on reducing the energy consumption and CO_2 emissions respectively (Lin et al. 2014).

In terms of G-VRP, the following studies can be noted. Kara et al. (2007) modelled the Energy Minimising Vehicle Routing Problem (EMVRP) like the capacitated VRP (CVRP) with a new cost objective function based on the total load and Arc length. However, the details of the formulation of fuel consumption are not provided.

Peng and Wang (2009) modelled the VRP based on fuel consumption by considering just the load of the

vehicle. In their objective function, minimising both vehicle travel distance and fuel consumption are targeted. They suggested that in order to have lower fuel consumption, serving the customers with high demand must be prioritised rather than customers with lower demand. A formulation of fuel consumption is done by (Xiao et al. 2012). They added a Fuel Consumption Rate (FCR) as a load-dependent function into a CVRP model and developed a CVRP model with the objective of minimising fuel consumption. In their work, they investigated both the distance travelled and the truckload to determine the fuel costs. Kuo (2010) in addition to the travel distances and load weights, also added the transportation speed to the fuel consumption calculation model in time-dependent VRP. Norouzi et al. (2017) developed a new mathematical model based on time-dependent vehicle routing problem to reduce fuel consumption by using Particle Swarm Optimization (PSO) algorithm.

Among the studies that paid attention to PRP, Maden et al. (2010) considered VRP problem with time windows constraint and proposed and implemented the heuristic algorithm in a case study within the UK which received about a 7% saving in CO_2 emission. Palmer (2007) presented an integrated model for routing and carbon dioxide emissions. He considered the role of speed in reducing carbon dioxide emissions in various congestion scenarios with window time and a reduction of 5% in CO_2 emissions is achieved. However, the effect of the weight of the load was not considered in his problem. Bektaş and Laporte (2011) developed a comprehensive objective function of carbon emissions, driver's cost and fuel consumption within the PRP model with or without time windows. In this work, a minimum speed of 40 km/h is considered as an assumption, which is in contrast to the congestion situations. Continuing this research, Demir et al. (2012) investigated the optimal driving speed and showed that the reduction in CO_2 emissions could occur by changing the speed within the network. In this study, SD-VRP formulation is modified to consider the CO_2 emission and guarantee minimum vehicle weight fill level on board in order to formulate the new PRP optimisation model.

2. MATHEMATICAL MODEL

In this research, a Pollution Routing Problem (PRP) with a homogeneous fleet of vehicles, time window, considering the possibility of split delivery and constraint of minimum shipment weight that must be on the vehicle in each route is investigated simultaneously. In the following, an integer linear programming model of the problem is described:

2.1. Input parameters

V: Total number of customers; $V= \{0, 1, \ldots, n\}$; Where node 0 corresponds to the depot and the other nodes in this set of vertexes represent the customers.
A: sets of edges; $A= \{(i,j) \mid i, j\} \in V$ and $i \neq j\}$.
K: Number of available vehicles; $K= \{1, ..., k\}$.
Q_k= Capacity of k^{th} vehicle ($k \in K$).
D_i= Customers Demand ($i \in V$).
d_{ij}= Length of edge between the nodes i and j; ; $(i,j) \in$A and $d_{ij}= d_{ji}$
M_{sk} = Minimum Shipment weight that must be on the k^{th} vehicle in length of each route
At_{ik} = Arrival time of k^{th} vehicle to i^{th} customer
$t1_i$ = Order time of i^{th} customer
$t2_i$ = Order due time of i^{th} customer
C_{ijk}= CO_2 emission of moving k^{th} vehicle ($k \in K$) between the nodes i and j
Where

$$C_{ijk} = \left(\left(TW_k + W_{ijk} \right) \times E_k \right) \times d_{ij}$$

Where
TW_k =Tare Weight of k^{th} vehicle
W_{ijk}= Weight of shipments on board of k^{th} vehicle between the nodes i and j
E_k= CO_2 emission rate of k^{th} vehicle

2.2. Decision variables

$$x_{ijk} = \begin{cases} 1 & \text{if } j^{th} \text{ customer is served by } k^{th} vehicle \\ & \text{after } i^{th} \text{ customer} \\ 0 & otherwise \end{cases}$$

y_{ik}= the quantity of the demand of i^{th} customer which is delivered by the k^{th} vehicle.

2.3. Formulation

Therefore, the vehicle routing problem formulation by Dror et al. (1994) can be modified in order to consider the CO_2 emission and guarantee minimum vehicle weight fill level on board in order to formulate the proposed Green Vehicle optimisation model in this study.
The objective function represents minimisation of the total CO_2 emissions generated by the transportation fleet can be written as follows:

$$Min \sum_{i=0}^{n} \sum_{j}^{n} \sum_{k=1}^{K} C_{ijk} x_{ijk} \ , i \neq j \qquad (1)$$

The ten model constraints considered in the proposed model are:

a) Constraint represented in equation (2) ensures that each customer is visited at least once which guarantees the possibility of a split delivery.

$$\sum_{i=0}^{n} \sum_{k=1}^{K} x_{ijk} \geq 1, \ \ j = 1, \ldots, n, \qquad (2)$$

b) Equation (3) is about entrance and exit flows (p), where if a node i is visited by vehicle k, then the amount of product from vehicle k that enters and leaves that node must equal the

demand at that node. Conversely, if node *i* is not visited by vehicle *k*, then the amount of product from vehicle *k* that enters and leaves that node must be 0. In fact, this constraint guarantee that any vehicle enters each node will definitely leave it.

$$\sum_{i=0}^{n} x_{ipK} - \sum_{j=0}^{n} x_{pjK} = 0\,, p = 0, \ldots, n;\ k = 1, \ldots, K, \tag{3}$$

c) Equation (4) guarantees that vehicle cannot continue to serve more customers in length of each route if the weight of its shipment on board, coming down is from a specified minimum shipment weight.

$$\sum_{i=0}^{n} W_{ijk} \geq M_{sk}\,,\ j = 1, \ldots, n; \quad k = 1, \ldots, K \tag{4}$$

d) Equation (5) ensures that the i^{th} customer's demand is completed if at least one vehicle passes through it.

$$y_{ik} \leq \sum_{j=0}^{n} x_{ijk}\,,\ i = 1, \ldots, n;\ k = 1, \ldots, K \tag{5}$$

e) Constraint displayed by equation (6) indicates that all customers demand is entirely fulfilled.

$$\sum_{k=1}^{K} y_{ik} = D_i, \quad i = 1, \ldots, n \tag{6}$$

f) Equation (7) imposes that the loading process on any route should not exceed the capacity of the vehicle.

$$\sum_{i=1}^{n} y_{ik} \leq Q, \qquad k = 1, \ldots, K \tag{7}$$

g) Constraint demonstrated by equation (8) presents the sub tour elimination constraints where (*S*) refers to any collection of customers having at least *2* and at most *n-1* members.

$$\sum\nolimits_{i,j \in S} x_{ijk} \leq |S| - 1\ \ (S \subset \{1, \ldots, n\});\ |S| \geq 2 \tag{8}$$

h) Equation (9) enforces that all customers are serviced no later than their order due time.

$$t1_i \leq At_{ik} \leq t2_i, i = 1, \ldots, n; k = 1, \ldots, K \tag{9}$$

i) Equation (10) guarantees the decision variables x_{ijk} to be binary.

$$x_{ijk} \in \{0,1\}, i = 0, \ldots, n\ ;\ j = 0, \ldots, n; k = 1, \ldots, K \tag{10}$$

j) Equation (11) guarantees that the decision variable yik is positive.

$$y_{ik} \geq 0, i = 1, \ldots, n;\ k = 1, \ldots, K \tag{11}$$

3. SIMULATED ANNEALING

Simulated annealing algorithm is an effective meta-heuristic optimization algorithm for solving optimization problems which is presented by (Kirkpatrick et al., 1983) and adapted from the Metropolis–Hastings algorithm (Metropolis et al., 1953). They proposed a gradual freezing technique to solve the hard optimisation problems. The main advantage of the simulated annealing algorithm is its ability to do not remain at the optimal local point and move to the global optimum point.

In generic term, the algorithm consists of two loops: the first loop reduces the temperature from the initial temperature to the final temperature and the second loop identifies the number of repetitions at each temperature. The factors affecting the timing of temperature reduction include the initial temperature, the final temperature, how to reduce the temperature and the number of repetitions in each temperature.

Simulated annealing algorithm starts from an initial solution and then in a repeat loop it moves to neighbouring solutions. If the neighbour's solution is better than the current one, the algorithm puts it as the current solution. Otherwise, the algorithm accepts that solution with the probability of *exp* (*-ΔE / T*) as the current solution. In this regard, *ΔE* is the difference between the objective function of the current solution and the neighbour's solution and *T* is a parameter called temperature. At each temperature, several repetitions are performed, and then the temperature is slowly reduced. In the initial steps, the temperature is set very high, so it is more likely to accept worse solutions. With the gradual decrease of temperature, in the final steps, there will be fewer probabilities for accepting worse solutions, and so the algorithm converges to a good solution. Figure 1 illustrates the general structure of the simulated annealing algorithm.

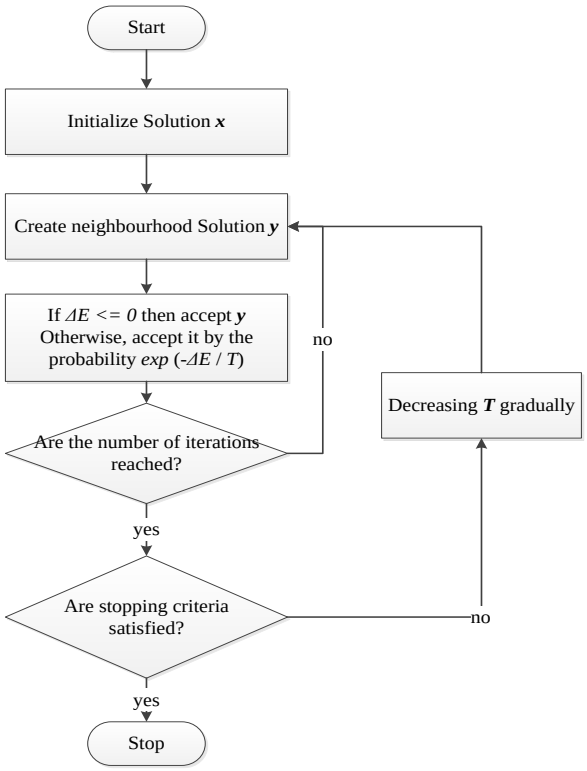


Figure 1: Simulated annealing algorithm flow chart

4. COMPUTATIONAL RESULTS

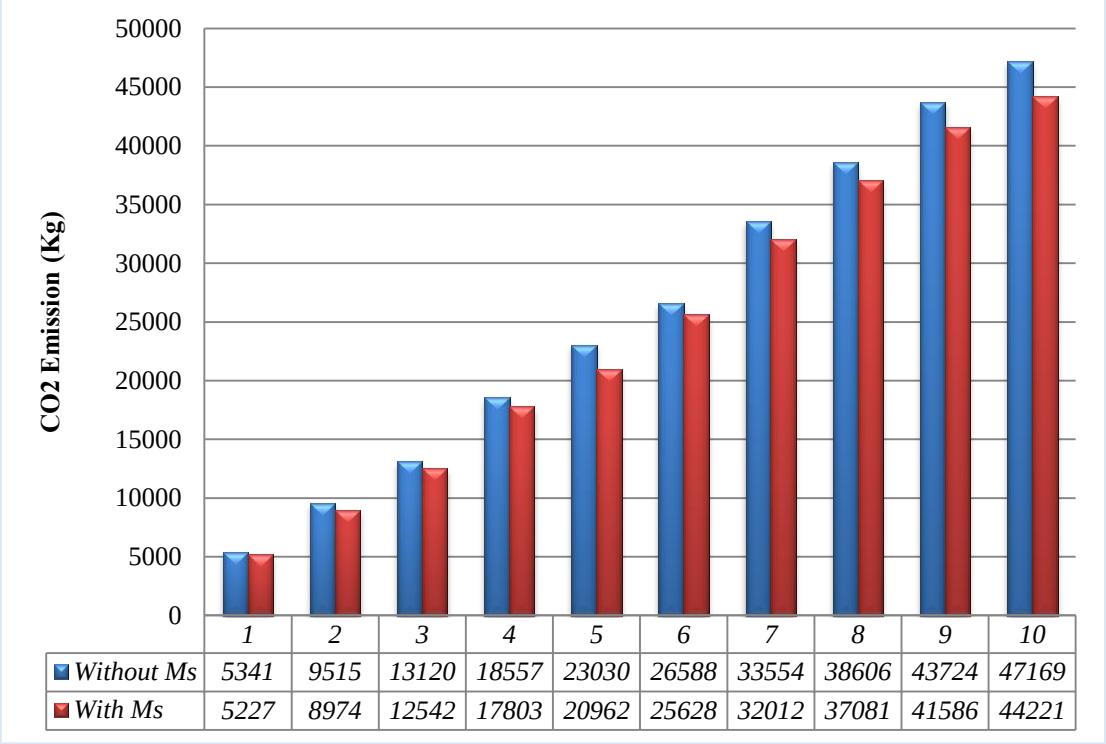
To evaluate the efficiency of the proposed mathematical method, numerical experiments with a single central depot and customer numbers of 20, 40, 60, 80, 100, 120, 140, 160, 180 and 200 are generated randomly; this will lead to ten different experiments. Customer's locations are randomly generated in the space of up to 200 in the *x* and *y* directions and depot's location is concentrated in the centre of this space randomly. Customers' demands weights (kg) are randomly selected from minimum 500 to maximum 3500 Kg to allow split delivery as the capacity of the vehicle is determined to be 3000 kilograms with a CO_2 emission rate of 0.0005442 kg per km (Hill et al., 2012). The number of the vehicle is considered equal to the number of customers through each experiment. The customer's order time (the date that order is placed) is determined randomly to be between January 1th, 2019 and January 7th, 2019.

Order due time of each customer is specified to be the maximum three days. It means customer's order must be filled within three days after that its order is placed.

In order to evaluate the efficiency and verify the proposed model, it was performed using simulation annealing through two approaches; one by considering the minimum weight of shipments on board (*Ms*) (equation 4) and one without its consideration. Comparison of the results shows in all experiments, by applying the *Ms* (Minimum Shipment weight) constraint, the obtained outputs from the proposed model are improved in terms of both mileage and CO_2 emission (see Table 1). In Figures 2 and 3, the quality of the generated results through the ten experiments is shown in the improvement of the two criteria of CO_2 emission and mileage.

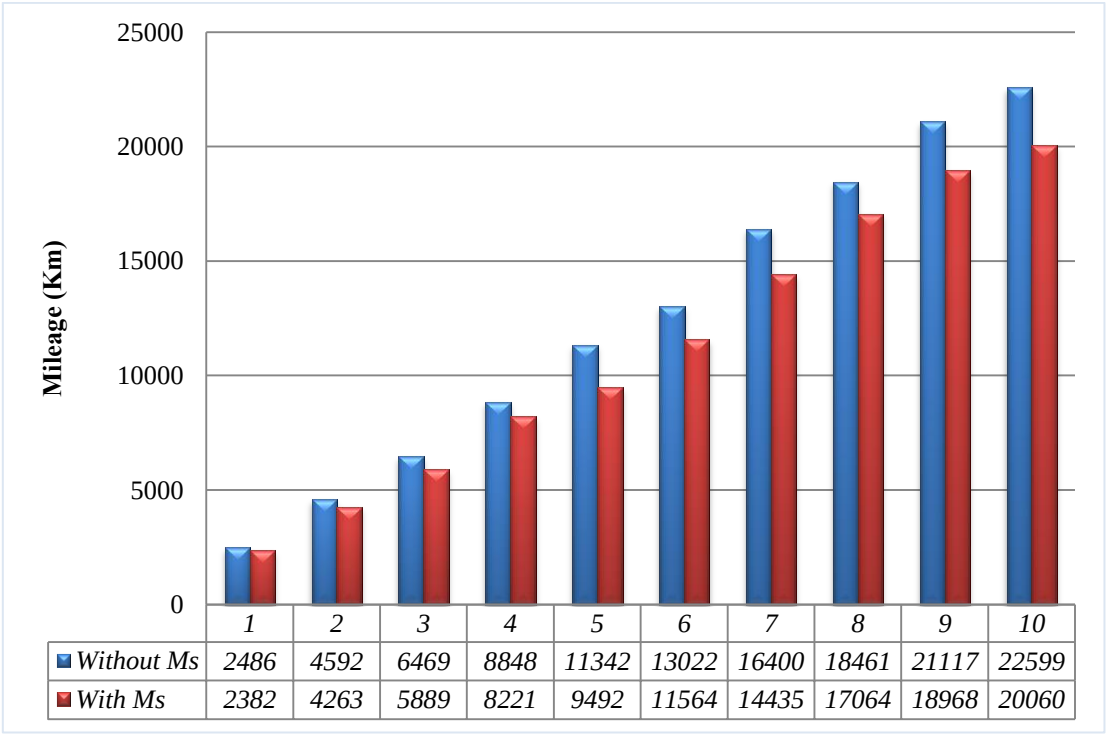
Table 1: The obtained reduction rate of mileage and CO_2 emission by applying the *Ms* (constraint represented by equation 4)

Experiment Number	Mileage reduction (%)	CO_2 emission reduction (%)
1	4.2	2.1
2	7.2	5.7
3	9.0	4.4
4	7.1	4.1
5	16.3	9.0
6	11.2	3.6
7	12.0	4.6
8	7.6	4.0
9	10.2	4.9
10	11.2	6.2



	1	2	3	4	5	6	7	8	9	10
Without Ms	5341	9515	13120	18557	23030	26588	33554	38606	43724	47169
With Ms	5227	8974	12542	17803	20962	25628	32012	37081	41586	44221

Figure 2: Comparison of the generated results in terms of the CO_2 emission criterion



	1	2	3	4	5	6	7	8	9	10
Without Ms	2486	4592	6469	8848	11342	13022	16400	18461	21117	22599
With Ms	2382	4263	5889	8221	9492	11564	14435	17064	18968	20060

Figure 3: Comparison of the generated results in terms of the Mileage criterion

5. CONCLUSION

In this paper, a new approach for Pollution Routing Problem (PRP) with time window and the possibility of split delivery was proposed to minimise the CO_2 emission by investigating minimum shipment weight that must be on the vehicle in length of each route.

The proposed mathematical model was developed and implemented using a simulated annealing algorithm in MATLAB software.

The results demonstrated that the application of the proposed model has led to a reduction in CO_2 emission.

In relation to future work, the proposed PRP will be developed with heterogeneous fleet and limit the number of vehicle for further evaluation of its effectiveness.

REFERENCE

Bektaş, T. and Laporte, G., 2011. The pollution-routing problem. Transportation Research Part B: Methodological, 45(8), 1232-1250.

Clarke, G. and Wright, J.W., 1964. Scheduling of vehicles from a central depot to a number of delivery points. Operations research, 12(4), 568-581.

Dantzig, G.B. and Ramser, J.H., 1959. The truck dispatching problem. Management Science, 6(1), 80-91.

Demir, E., Bektaş, T. and Laporte, G., 2011. A comparative analysis of several vehicle emission models for road freight transportation. Transportation Research Part D: Transport and Environment, 16(5), 347-357.

Dror, M. and Trudeau, P., 1990. Savings by split delivery routing. Transportation Science, 23(2), 141-145.

Dror, M., Laporte, G. and Trudeau, P., 1994. Vehicle routing with split deliveries. Discrete Applied Mathematics, 50(3), 239-254.

Hill, N., Walker, H., Choudrie, S. and James, K., 2012. Guidelines to Defra/DECC's GHG Conversion Factors for Company Reporting: Methodology Paper for Emission Factors, Department for Environment. Food and Rural Affairs (Defra).

Kara, I., Kara, B. and Yetis, M.K., 2007. Energy minimizing vehicle routing problem. Combinatorial optimization and applications, 62-71. Springer, Berlin, Heidelberg.

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P., 1983. Optimization by simulated annealing. science, 220(4598), 671-680.

Kuo, Y., 2010. Using simulated annealing to minimize fuel consumption for the time-dependent vehicle routing problem. Computers & Industrial Engineering, 59(1), 157-165.

Laporte, G., 1992. The vehicle routing problem: An overview of exact and approximate algorithms. European journal of operational research, 59(3), 345-358.

Maden, W., Eglese, R. and Black, D., 2010. Vehicle routing and scheduling with time-varying data: A case study. Journal of the Operational Research Society, 61(3), 515-522.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E., 1953. Equation of state calculations by fast computing machines. The journal of chemical physics, 21(6), 1087-1092.

Mullaseril, P.A., Dror, M. and Leung, J., 1997. Split-delivery routeing heuristics in livestock feed distribution. Journal of the Operational Research Society, 48(2), 107-116.

Norouzi, N., Sadegh-Amalnick, M. and Tavakkoli-Moghaddam, R., 2017. Modified particle swarm optimization in a time-dependent vehicle routing problem: minimizing fuel consumption. Optimization Letters, 11(1), 121-134.

Palmer, A. 2007.The development of an integrated routing and carbon dioxide emissions model for goods vehicles. PhD thesis, School of Management, Cranfield University, Bedford, United Kingdom.

Peng, Y. and Wang, X., 2009, April. Research on a vehicle routing schedule to reduce fuel consumption. In Measuring Technology and Mechatronics Automation, 2009. ICMTMA'09. International Conference on (Vol. 3, pp. 825-827). IEEE.

Potvin, J.Y. and Bengio, S., 1996. The vehicle routing problem with time windows part II: genetic search. INFORMS Journal on Computing, 8(2), 165-172.

Reimann, M., Stummer, M. and Doerner, K., 2002, July. A savings based ant system for the vehicle routing problem. In Proceedings of the 4th Annual Conference on Genetic and Evolutionary Computation (pp. 1317-1326). Morgan Kaufmann Publishers Inc.

Saad, S. and Bahadori, R., 2019. Development of a dynamic information fractal framework to monitor and optimise sustainability in the food distribution network. International Journal of Food Engineering, in press.

Tavakkoli-Moghaddam, R., Safaei, N., Kah, M.M.O. and Rabbani, M., 2007. A new capacitated vehicle routing problem with split service for minimizing fleet cost by simulated annealing. Journal of the Franklin Institute, 344(5), 406-425.

Xiao, Y., Zhao, Q., Kaku, I. and Xu, Y., 2012. Development of a fuel consumption optimization model for the capacitated vehicle routing problem. Computers & Operations Research, 39(7), 1419-1431.

Yoshizaki, H.T.Y., 2009. Scatter search for a real-life heterogeneous fleet vehicle routing problem with time windows and split deliveries in Brazil. European Journal of Operational Research, 199(3), 750-758.

AUTHORS BIOGRAPHY

Professor Sameh M. Saad, BSc (Honours), MSc, PhD, PGCHE, CEng, MIET, MILT, FHEA, FCILT. He is the Professor of Enterprise Modelling and Management and also a Postgraduate Research Coordinator and MSc/MBA Course Leader in the Department of Engineering, Faculty of Science, Technology and Arts (STA), Sheffield Hallam University, UK. His research interests and experience include fractal supply chain, modelling and simulation, design and analysis of manufacturing systems, production planning and control, reconfigurable manufacturing systems and next generation of manufacturing systems including fractal and biological manufacturing systems. He has published over 150 articles in various national and international

academic journals and conferences, including keynote, address a book and four patents.

Dr Ramin Bahadori BSc, MSc, PhD is a lecturer in the Department of Engineering, Faculty of Science, Technology and Arts (STA), Sheffield Hallam University, UK. His research interest and experience include fractal supply chain, food supply chain sustainability, CO_2 emission optimisation, modelling and simulation, technology road mapping, logistics capabilities, inventory optimisation, logistics cost optimisation, information system, responsiveness optimisation, supply network integration, communication and collaboration within supply network, and multi-criteria decision-making fuzzy AHP.

OPTIMIZING ENERGY EFFICIENCY IN DISTRIBUTED MHEALTH NETWORKS

Philipp Skowron[(a)], Michael Aleithe[(b)], Bogdan Franczyk[(c)]

[(a),(b)] Leipzig University, Grimmaische Straße 12, 04109 Leipzig, Germany
[(c)] Wroclaw University of Economics, Wroclaw, Poland

[(a)]skowron@wifa.uni-leipzig.de, [(b)]aleithe@wifa.uni-leipzig.de,
[(c)]bogdan.franczyk@ue.wroc.pl

ABSTRACT
Energy efficiency in mobile health applications is a relevant problem for long-term monitoring and user acceptance. Various parameters influence the runtime of the system to some degree. One of the parameters is the sampling rate of the individual distributed sensors. Increasing the sampling rate can lead to an increase in energy consumption within the system. By contrast, a reduction can lead to a loss of the data quality, which reduces the informative value of the results of algorithms that use this data. Using optimization methods from reinforcement learning and deep learning to adaptive adjust the sampling rates during runtime, energy efficiency could be improved in only 40 training runs without losing data quality during sampling.

Keywords: mHealth, energy efficiency, therapy systems, reinforcement learning

1. INTRODUCTION

Monitoring and evaluation of daily activities, measurement and control of nutritional habits, monitoring of glucose levels, long-term ECGs for the detection of ischemia signs up to the mobile early detection of emotional break-ins of patients suffering from chronic depression. This is just a small overview of mHealth solutions in the end-customer market, which according to Reuters is expected to grow by more than 35% over the next three years and is already worth 23 billion dollars (Orbis Research 2017). This can be seen especially in the trend that end customers want to live more consciously and healthily and want to have full control over their physical health. In the same time, this trend makes it possible to expand medical care, even where there is no doctor or hospital nearby.

Many solutions use the already existing infrastructures at the end customer's side to perform their services. These include personal smartphone and wearables such as fitness trackers or smart watches. These are either extended with external sensors or the existing sensors in the devices are used. In order to be able to collect data and draw conclusions from this, the sensors in the devices must be used in different ways, which varies depending on the particular use case, i. e. so-called sampling rates are defined for each sensor, which collects data in different time intervals ranging from a few milliseconds to hours. The more fine-grained the sampling rate time intervals are, the higher the power consumption of the individual sensors. This also reduces the battery life of the entire system, which can have a negative impact on the acceptance and long-term use of the system for the end users (both patients and medical staff, depending on the application). Thus, in this way, user acceptance is also a significant aspect of the dissemination of mHealth solutions.

In particular, algorithms such as deep neural networks, which depend on large amounts of data in order to deliver good results, require high sampling rates in order to obtain the necessary data in order to be trained effectively. In order to achieve a long runtime of the overall system, the energy efficiency of the application can be increased by adjusting the sampling rates. This can be described as an optimization problem, whereby the objective function is the runtime of the overall system. The dependent variables of the optimization problem are the sampling rates on the one hand and a data quality measure of the used algorithms that process the measured data and depend on them on the other. Such a data quality measure should avoid that the meaningfulness of the algorithms is impaired during optimization.

The work is organized as follows. Section 2 discusses the existing approaches. Section 3 deals with the optimization problem between total runtime and sampling rates. In section 4 the concept is introduced to be able to do the optimization. Section 5 describes the use case to which the concept was applied. The results of the simulation are presented and discussed in section 6 and section 7 presents the future considerations in detail.

2. RELATED WORKS

In previous work, several approaches to saving energy resources have already been developed. (Nguyen et al. 2008) describes a simulation system with which strategic decisions for the selection of monitoring components, like sensors etc., can be determined to reduce the power consumption of remote monitoring systems. Here, the focus of the work is on the technical description of the simulation system as well as the architectural design of

the simulation system. The actual optimization, however, is carried out manually and not algorithmically.

(Zois, Levorato and Mitra 2013) presents an advanced approach that enables and disables sensors of a distributed monitoring system based on their power consumption and profile. This means that algorithmically, based on quality characteristics, it is determined when which sensor must be used. Like this paper, at (Zois, Levorato and Mitra 2013) the sensors are switched off completely, but no configuration elements of the sensors are included.

Other approaches such as that of (Chai et al. 2014) do not consider the sensors and their power consumption per se, but rather the data transmission and their frequency. It is algorithmically determined when and how many data are to be transmitted from which sensor. Thus, the focus of the paper is on optimizing the transmission intervals between the base station and the sensors without specifically addressing the underlying algorithms and their optimization.

As the approach of (Zois, Levorato and Mitra 2013), (Chattopadhyay and Mitra 2017) looks at the stochastic selection of transmitting sensors in a distributed sensor network based on a modified Gibbs algorithm to determine the incoming data under uncertainty. However, as with (Chai et al. 2014), the focus here is again on optimizing the transmission times between base station and sensors, without going into the optimal selection of sensors at a certain point in time, based on the measured data.

Like (Chai et al. 2014), (Zhang, Zhang and Zhang 2017) uses the transmission rate as a variable for optimizing energy consumption. They maximize the lifetime of the system through a heuristic approach that allows transmit power through multi-hop transmission of data from sensor to sensor. Although the system is working, it can only be used when a multi-hop application is possible.

(Aleithe et al. 2018) developed a simulation framework for troubleshooting mHealth systems. They investigated the runtime on the bases of the energy consumption of the sensors used. Although the approach identifies precise problem sources on the basis of the FMEA, the adjustment is still manual and cannot be changed dynamically.

(Wu et al. 2019) consider the energy consumption when transferring data to a personal device. The sensors are all equipped with an energy recovery module, which allows them to generate their own power. During transmission, no more energy should be consumed than is produced, without causing excessive delays in transmission. Even if the energy consumption is taken into consideration, due to the energy production of the sensors, this is not the main focus for the optimization.

3. OPTIMIZATION PROBLEM

An interesting parameter for improving distributed body area networks is the overall system runtime. Longer system runtime, e. g. due to reduced power consumption, reduces charging times, battery change, etc. This can be achieved by reducing the power consumption of distributed sensors, for example, by influencing certain parameters of the sensors. For example, sensors can be switched on and off or their scanning frequency can be changed. The disadvantage of switching the sensors on and off is that not all sensors can be completely disconnected from the power supply and reconnected later. In addition, there may be other functionalities of the systems that require certain sensors independent of the system under consideration. The scanning frequency of the sensors is thus an effective instrument to reduce power consumption in the overall system without compromising other functionalities.

The sampling interval for sensors determines how often data are collected from the sensors and transferred to the system, where each sensor has its own sampling interval. The sampling interval is defined with the interval $0 < T < +\infty$ and valid under the condition that $T \in \mathbb{R}$.Depending on the nature of the sensor, this sampling interval can take any positive value. The higher the value of the sampling interval, the fewer data from the sensors are collected in some time frame. The overall description of T_n is shown in Equation (1), where t_n is a moment in a time series which is generated by the sensor.

$$T_n = t_{n+1} - t_n | \mathrm{n} \in \mathbb{N}\{0 \leq \mathrm{n} < +\infty\}, \quad \mathrm{t} \in \mathbb{R}\{-\infty < \mathrm{t} < +\infty\} \tag{1}$$

If the sampling interval of a sensor is set to four, for example, the initial value is read in time zero and every four time steps new values are read in each subsequent time (Yang 2014). This fact can be illustrated in Figure 1 more clearly.

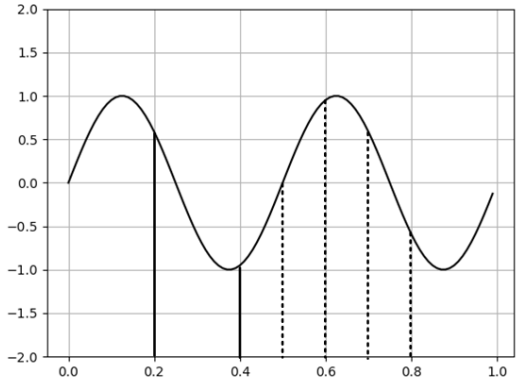


Figure 1: With a sampling interval of 0.2 (black line) only rough fragments of the actual signal are read and processed. If the sampling interval is decreased to 0.1 (dotted line), considerably more characteristics of the signal can be read and processed at the same time.

Now, T is defined as sampling interval, which can increase or decrease. This is an important fact for the optimization problem and the depending variables of T. Therefore, it must be defined how T behaves when it increases or vice versa.

$$T_{dec} := t_n(T) > t_n(T_{dec}) \tag{2}$$

Equation (2) defines, on the one hand, the decreasing of T (T_{dec}). This means that more data are collected in a shorter time than before.

$$T_{inc} := t_n(T) < t_n(T_{inc}) \quad (3)$$

Equation (3), on the other hand, shows that if T increases (T_{inc}), it collects less data at the same time as before, because it takes more time.

In a continuous space, as with physical data such as heart rates etc., the sampling rate determines how accurately the real world can be approximated. Different sampling rates with different sensors, therefore, map the signals and thus the values from the real world with different precision.

The selection of the optimal sampling rates for a maximum total runtime of the system can be described as an optimization problem. The runtime *dur* of the system represents the variable to be optimized and the sampling rates T of any number of sensors represent the parameters. In empirical experiments, it was proved, that the duration of a system is dependent on the sampling rates of all sensors in the system. Therefore Equation (4) defines the duration of the system.

$$dur(T) := d \in \mathbb{N} | 0 < \mathrm{d} < +\infty \quad (4)$$

The problem with this definition of the optimization problem is that no clearly defined optimum exists. In this way, the optimization would converge against the highest possible sampling rates, which would cause almost a stagnation of the data collection. One possibility to define an optimum is the total runtime of the system without sensor reading. However, functionalities based on the collected data would be affected. A better definition of the optimum is, therefore a quality measure, which reflects the quality of functionality or an algorithm output and at the same time is used as an optimization parameter. For predictive algorithms, for example, such a quality measure could be the accuracy of the algorithm, cf. Equation (5), with the measured data based on the sampling rates of the sensors.

$$acc(T) := a \in \mathbb{R} | 0 \leq \mathrm{a} \leq 1 \quad (5)$$

However, these quality measures must be defined individually for each algorithm and each function in order to measure the best possible representation of the quality of the algorithm's output. These quality measures limit the optimization of the sampling rates to the degree that the sampling rates can only take those values which are within a certain tolerance range of the quality measure.

$$maximize(acc, dur) \quad (6)$$

Both, the accuracy of the quality measure and the system duration, should be maximized after the optimization taking into account the dependencies between both variables.

To describe the dependency between *acc* and *dur* Equations (7) and (8) are provided.

$$acc(T) = \begin{cases} acc'(T) \geq 0, if\ T_{dec} \\ acc'(T) \leq 0, if\ T_{inc} \end{cases} \quad (7)$$

$$dur(T) = \begin{cases} dur'(T) \geq 0, if\ T_{inc} \\ dur'(T) \leq 0, if\ T_{dec} \end{cases} \quad (8)$$

Both equations show, that if T increases or vice versa, *acc* and *dur* are reverse related. This is because *acc* rises if more data are available to distinguish between multiple data samples. The connection between *acc* and *dur* will be further described in section 4 and 5.

Thus, the optimization in this case consists of the optimization algorithm, mapped as a reinforcement learning (RL) problem, the total runtime of the system, the sampling rates of the sensors as optimization parameters and the quality measures of the functions and algorithms, which should be included as further optimization parameters and ensure the quality of the system.

4. MODEL CONCEPT

As described in the previous chapter, the model concept presented here consists of several components. These components include the simulated environment with the associated sensors, the restrictive algorithms with their quality measures, which are formulated as boundary conditions for optimization, and the optimization algorithm itself. Figure 2 illustrates the exchange of information between the individual components.

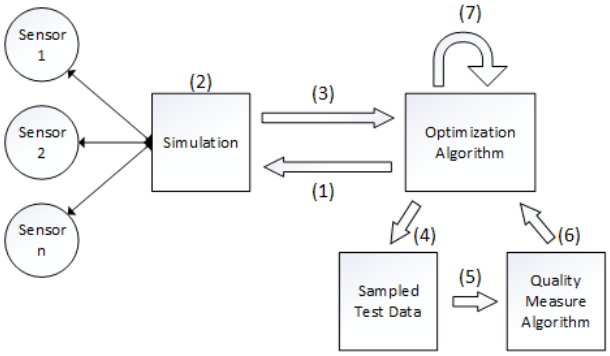


Figure 2: The optimization algorithm gives the sampling rates in the simulation (1). The simulation propagates the sampling rates to the sensors (2) and simulates the total runtime, which is passed back to the optimization algorithm (3). Based on the sampling rates, new test data are then generated (4) which are used in the quality measurement algorithm (5) to return the accuracy based on the current sampling rates to the optimization algorithm (6). The optimization algorithm then updates its network (7) and generates new sampling rates.

As it is shown in Figure 2, the environment simulation is responsible for generating the duration of the total system runtime under the given sampling rates. The restrictive algorithms are also re-tested according to the sampling rates to calculate the change in the quality measure. The

optimization algorithm then calculates new sampling rates and the loop starts again. The individual components of the model architecture are explained in more detail below.

The model of the environment maps the simulated data and the associated sensor network. The data generated for the sensors are based on different distributions to ensure variance between the individual sensors. This data is based on a continuous space, like most of the values measured in the real world. In a concrete example, these values could be the pulse rate or position determined via GPS. Due to the continuous nature of the measured data, the choice of sampling rate also determines the accuracy of the approximation of the environmental values. This, in turn, has an influence on pattern recognition of different algorithms whose data basis is subjected to this approximation.

The quality measurement algorithms in the model consist of both the algorithm itself and the quality measure. Depending on the algorithm, the measurement of quality is defined differently. In this process, the quality measures flow indirectly into the calculation of the optimum to varying degrees, depending on the weighting in the reward function of the optimization algorithm. There is no restriction on the number of quality measures to be used, but there is an increase in the complexity of the environment and the difficulty of finding a global or near-global optimum when using multiple quality measures at once.

As mentioned above, a reinforcement learning (RL) algorithm is defined as an optimization algorithm. Reinforcement learning is a semi-supervised learning method in the field of machine learning. This method consists of the continuous exchange between an action of the agent as a reaction to a state of the environment and the subsequent new state of the environment and a reward signal. Through the continuous exchange of status, reward and action, the agent develops a policy that attempts to optimize a variable in interaction with the environment.

These RL algorithms have the property of reducing the state and action space by learning a function where other approaches such as brute force algorithms fail due to the complexity of the environment and the resulting runtime. In addition, RL algorithms using neural networks enable the learning of non-linear contexts. Since the state space, i. e. the measured data from the sensors, is subject to continuous distribution, policy-gradient functions that can deal with continuous environments are suitable for the selection of appropriate RL algorithms (Sutton and Barto 2018). Recurrent Neural Networks can be used to map the temporal dependency between runtime and selected sampling rates (Hochreiter and Schmidhuber 1997). The output layer is a linear output consisting of three output nodes (total number of sensors in this use case). In contrast to classical classification methods, no classification is carried out here, but all values are returned directly. This has the advantage that the algorithm can adjust several sampling rates simultaneously during a one-time step.

Now that the general structure of the model concept has been described, the next chapter will explain the use case and simulation structure.

5. USE CASE SIMULATION

In the last chapter, the components of the presented model were introduced and explained in general. This chapter applies this model to a concrete use case from a current research project. The use case uses wearables and the participants' smartphones to identify mood patterns. Measured values can be obtained via the sensors from the smartphone as well as via the sensors of the wearables, which can be used to determine firmly classified mood levels. For this purpose, different sampling rates need to be set for different sensors. The problem here is the tradeoff between the long overall runtime of the system and the accuracy of the mood detection. If the sampling rate is maximized, the total runtime of the system drops to a few hours. If the sampling rates are increased simultaneously, the system can operate for up to one day. However, this reduces the accuracy of mood detection as predictive algorithms depend on the measured data.

Real data are not yet available for use in the simulation since they will be collected later in the research project, assumptions about the data must be made in order to prove the convergence and thus the functionality of the approach. The sensor data of the wearables were generated randomized and normally distributed. To be able to predict three different moods with the quality measurement algorithms, three slightly different distributions were added to the simulated data.

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} * e^{\frac{-}{2\sigma^2}} \tag{9}$$

Equation (9) shows the used normal (Gaussian) distribution for generating the sample data. Depending on the mood, different mean values for each mood. Mood one got a mean of 2.0, mood 2 a mean of 0 and mood 3 a mean of -2.0. All have the same variance of 1.

The used quality measure algorithm is a fully-connected two hidden layer neural network with classification output. The quality algorithm was pre-trained with data collected from the sensors. All data were labeled with the corresponding class labels of the three moods in a one-hot-encoding. The simulation of the sensors and the associated continuous environment was implemented in Matlab Simulink® and generates the total runtime of the system. For the optimization algorithm, an actor-critic reinforcement learning model has been chosen, as this is a policy-gradient model that converges well with complex, continuous state-action spaces.

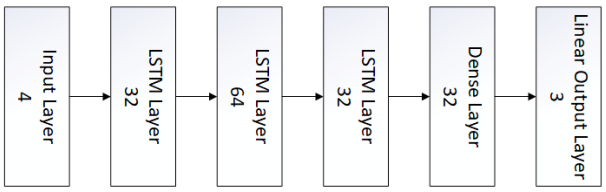


Figure 3: Illustrate the architecture of the deep neural network for the reinforcement learning algorithm

The neuronal network for the reinforcement learning algorithm, displayed in Figure 3, was built up empirically and consists of several layers of LSTM cells to map the temporal dependencies. The Reward hypothesis is formulated as follows:

$$E(r) = \delta * \Big((a_t - a_{t-1}) + \big(a_t - a_{t_0}\big)\Big) + (d_t - d_{t-1})) * y^{-1} \quad (10)$$

a describes the accuracy of the quality measurement and *d* the total runtime of the simulated system. δ and γ are hyperparameters for balancing the importance of the two parameters runtime and accuracy. In this use case, the total runtime was measured in seconds, resulting in a magnitude of 10^4. Variations in the accuracy of the quality measure were in the range from 10^{-3} to 10^{-1}, which required an adjustment of δ and γ to an appropriated level.
The loss function is a standard policy gradient update function similar to the publication in this paper (Mnih 2016). However, a synchronous variant like the one in (Wang et al. 2016) was used. This decision was made due to the limited resources on the one hand and also based on the considerations and arguments and test results of (Wu et al. 2017) on the other hand.

6. RESULTS & DISCUSSION

After a few episodes, the result of several training runs with different configurations was a convergence of the total runtime. As can be seen in Figure 4 the algorithm successfully converged after about ten episodes against a longer runtime than initially indicated.
In 40 episodes the sampling rates were adjusted 15 times in each episode. When averaging the episodes, Figure 4, a short slump within an episode of the total runtime can be considered. Afterward, the expected convergence to an optimal overall runtime happens. A view at the sampling rates of the sensors in Figure 4 over the episodes shows that they also oscillate to an optimum in the applied use case. However, a dynamic environment has not been included, in which the sampling rates do not seek a general optimum state but are adjusted again depending on the changes in the environment, e. g. through user interaction.
The presented results are a proof of concept which shows that the approach leads to an optimum in principle. The fast convergence and the high duration in the example are probably due to the simplicity of the environment, which consisted of only three sensors and a limiting algorithm. A more complex environment would require more effort from the RL algorithm.

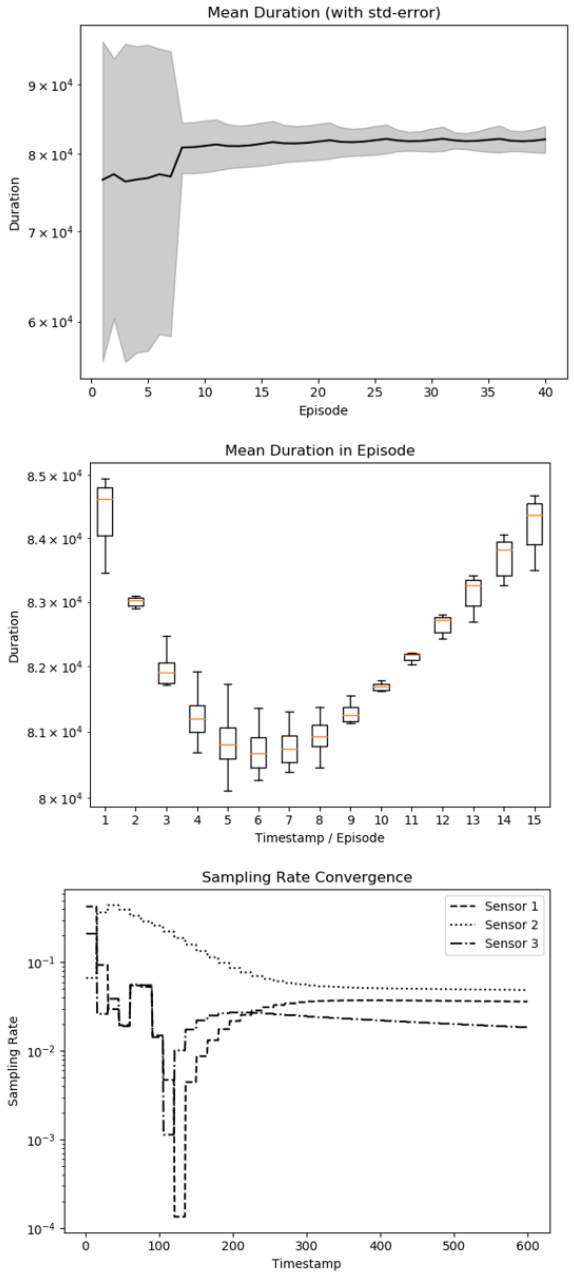


Figure 4: The first diagram shows the average total system runtime over the episodes with standard deviation as an error rate. The second diagram shows the development of the total runtime within an episode during the adjustment steps. The last illustration shows the sampling rates of the sensors and their development across the episodes.

Nevertheless, despite simulative data with fictitious distributions, these outcomes can be interpreted as representative results, since the concept will behave in the same way even with sufficient underlying real data. The same applies to determine the accuracy of the quality measurement algorithms. These fundamentally influence the sensitivity of the algorithm to changes in the dataset. A change of the dataset to real data would only shift the

optimum between total runtime and accuracy, but not the convergence of the RL algorithm, by adjusting the hyperparameter in the reward hypothesis. Thus, the use of the concept for optimizing the overall runtime in mHealth applications can be described as basically successful. The following section describes the next steps for further evaluation of the concept to test its stability under real conditions.

7. CONCLUSION & FUTURE WORK

This paper presented a concept to increase the overall runtime of a distributed sensor system by adjusting the sampling rates of the sensors. A reinforcement learning algorithm was used to optimize the ratio between sensor sampling rates and system runtime. In order to ensure that the quality of the collected data is respected, quality measurement has also been introduced. The result was a convergence of the overall system runtime to an optimal level without reducing the quality of the data and the resulting knowledge out of the data for further algorithms in the system. All calculations were performed on simulated data since the evaluation of real data can only be carried out after the proof of concept.

The aim is to replace the simulated data with real test data to validate the functionality of the system under real conditions. In addition, a dynamic environment will be tested in future observations. For this purpose, everyday user interactions should be considered when calculating the optimum solution. Also, it is planned to test the algorithm on a field test with patients of a depression study.

REFERENCES

Aleithe M., Skowron P., Carell A., Boettger D., Goblirsch T., Franczyk B., 2018. Simulation Framework for Mobile Patient Monitoring Systems. Proceedings of the International Workshop on Innovative Simulation for Health Care (IWISH), 17.09.2018, Budapest, Hungary.

Chai R., Wang P., Huang Z., Su C., 2014. Network lifetime maximization based joint resource optimization for Wireless Body Area Networks. 2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), 1088–1092. 02.09.2014, Washington DC, USA.

Chattopadhyay A., Mitra U., 2017. Optimal Dynamic Sensor Subset Selection for Tracking a Time-Varying Stochastic Process. arXiv preprint arXiv:1711.10610.

Hochreiter S., Schmidhuber J., 1997. Long Short-Term Memory. Neural Computation 8:1735–1780.

Mnih V., Badia A.P., Mirza M., Graves A., Lillicrap T.P., Harley T., Silver D., Kavukcuoglu K., 2016. Asynchronous Methods for Deep Reinforcement Learning. Proceedings of The 33rd International Conference on Machine Learning. 19.06.16, New York, NY, USA.

Nguyen K.D., Cutcutache I., Sinnadurai S., Liu S., Basol C., Sim E., Phan L.T.X., Tok T.B., Francis L.X., Tay E.H., Mitra T., Wong W.-F., 2008. Fast and accurate simulation of biomonitoring applications on a wireless body area network. 2008 5th International Summer School and Symposium on Medical Devices and Biosensors, 145–148. 01.06.2008, Hong Kong, China.

Orbis Research, 2017. mHealth Market Worth $23 Billion in 2017 and Estimated to Grow at a CAGR of more than 35% over the next three years. Reuters. Available from: https://www.reuters.com/brandfeatures/venture-capital/article?id=4640 [01.18].

Sutton R.S., Barto A.G., 2018. Reinforcement Learning. Cambridge:The MIT Press.

Wang J.X., Kurth-Nelson Z., Tirumala D., Soyer H., Leibo J.Z., Munos R., Blundell C., Kumaran D., Botvinick M., 2016. Learning to reinforcement learn. Proceedings of the 38th Annual Conference of the Cognitive Science Society. 10.08.16, Philadelphia, Pennsylvania, USA.

Wu Y., Mansimov E., Liao S., Grosse R., Ba J., 2017. Scalable trust-region method for deep reinforcement learning using Kronecker-factored approximation. Advances in neural information processing systems. 5279–5288. 04.12.2017, Long Beach, CA, USA.

Wu G., Chen Z., Zhang D., Jiaqi L., 2019. Resource allocation algorithm with worst case delay guarantees in energy harvesting body area networks in Peer-to-Peer Networking and Applications 12:74-87.

Yang S.-H., 2014. Wireless sensor networks. London:Springer.

Zhang Y.; Zhang B.; Zhang S., 2017. A Lifetime Maximization Relay Selection Scheme in Wireless Body Area Networks in Sensors, 17:1267-1287

Zois D.-S., Levorato M., Mitra U., 2013. Energy-Efficient, Heterogeneous Sensor Selection for Physical Activity Detection in Wireless Body Area Networks. IEEE Transactions on Signal Processing 7:1581–1594.

AUTHORS BIOGRAPHY

Philipp Skowron studied Business Information Systems at Harz University of Applied Sciences and afterwards he continued his studies at University of Leipzig in the same area of study. Currently, he is writing on his PHD about the simulation of cyber physical systems using artificial intelligence in the mHealth area, which is also his specialization and research interest. Supervisor is Prof. Dr.-Ing. Bogdan Franczyk.

Michael Aleithe studied Systemdesign at Ernst-Abbe University for Applied Sciences and specialized at Modeling and Simulating of Complex Technical Systems. In the past he was employee at Xceptance Software Technologies GmbH and works on agentbased imulations. His research-interest include interconnected data, simulation and mHealth. He is currently working on

a PhD thesis in the area of designing mHealth systems at Leipzig University, Germany supervised by Prof. Dr.-Ing. Bogdan Franczyk.

Bogdan Franczyk is professor at Leipzig University since 2002. In the past he was the research director of Intershop AG as well as professor in Swinburne, Australia and Wroclaw, Poland. His research-area includes especially information management systems.

Author's index

ISBN 978-88-85741-37-9; Bruzzone, Janosy, Nicoletti and Zacharewicz Eds.