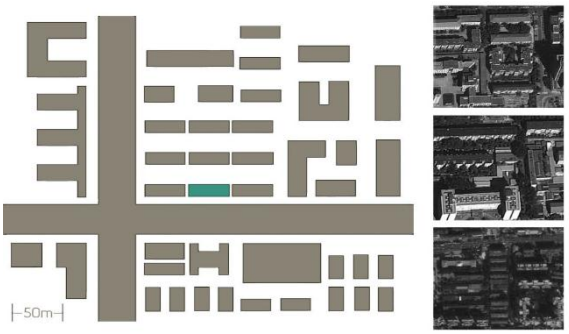
Journal of Energy Challenges and Mechanics

ISSN 2056-9386

http://www.nscj.co.uk/JECM/

Volume 1, Issue 1
September 2014



Featured article:

The effects of future climate change on energy consumption in residential buildings in China and retrofitting measures to counteract

D. H. C. Chow[1], M. Kelly[2], Huiyang Wang[1], J. Darkwa[2]*

[1]School of Architecture, University of Liverpool, Liverpool L69 7ZN, UK

[2]Centre for Sustainable Energy Technologies (CSET), University of Nottingham Ningbo China, 315100, China

Journal of Energy Challenges and Mechanics, volume 1, pages 1-7

North Sea

North Sea Conference & Journal LTD
2 Charlestown Walk, Cove Bay, AB12 3EZ, Aberdeen, Scotland, United Kingdom
http://www.nscj.co.uk/JECM/ | jecm@nscj.co.uk | +44(0)1224 875635

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1

TABLE OF CONTENTS pages

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 1

The effects of future climate change on energy consumption in residential buildings in China and retrofitting measures to counteract

未来气候变化对中国住宅的能耗影响及相对改造措施

D. H. C. Chow[1*], M. Kelly[2], Huiyang Wang[1], J. Darkwa[2]

[1]*School of Architecture, University of Liverpool, Liverpool L69 7ZN, UK*
[2]*Centre for Sustainable Energy Technologies (CSET), University of Nottingham Ningbo China, 315100, China*

David.Chow@liverpool.ac.uk

Accepted for publication on 22th July 2014

Abstract - **At present, China is going through a rapid rate of mass urbanisation, and this poses a number of challenges for the building sector. On one hand, under new directives from the government, new buildings will have stricter energy requirements and existing buildings will also need to lower their rate of energy consumption, on another hand, the lifetime of buildings are now intended to last longer, meaning that building designers will also need to account for effects of future climate change when assessing the performance of building schemes. This paper investigates the effect of future climate change on energy consumption in typical residential buildings in different climate regions of China. These include the "Cold" region in the north, which includes Beijing; the "Hot Summer Mild Winter" region in the south, which includes Guangzhou, and two regions from the "Hot Summer Cold Winter", one along the coast in the east, which includes cities such as Shanghai and Ningbo; the inland region, which includes cities such as Wuhan and Chengdu. Using data from the climate model, HadCM3, Test Reference Years are generated for the 2020s, 2050s, and 2080s, for various IPCC future scenarios for these cities. These are then used to assess the energy performance of typical existing residential buildings, and also the effects of retrofitting them to the standard of the current building codes. It was found that although there are reductions in energy consumption for heating and cooling with retrofitting existing residential buildings to the current standard, the actual effects are small compared with the extra energy consumption that comes as a result of future climate change. This is especially true for Guangzhou, which currently has very little heating load, so there is little benefit of the reduction in heating demand from climate change. The overall effects of retro-fitting in other selected cities depend largely on the specification of current existing buildings. In general, more improvements in building standards in all four regions are required to significantly reduce the effects of future climate change.**

Keywords - **Climate Change, Energy Consumption, Residential Buildings, Retrofitting, Urbanisation.**

I. INTRODUCTION

Since the opening of its market in 1978, the Chinese economy has grown at a staggering pace, leading to a drastic increase in energy consumption. The building sector is responsible for around 27.5% of the national total energy consumption [1-3]. And this could even rise to 40% over the next 20 years as more buildings are constructed [4]. Occupants too, are demanding a higher level of indoor comfort which will lead to a steeper increase in heating and cooling loads. In 2010, China has over 43 billion m^2 of constructed area, but just 4-5% meet the national building energy standards, the rest are classified as "intensive energy consumers" [5-6].

In China, the lifespan of most buildings is relatively short, compared to European and North American standards. This high turnover rate, along with the recent construction boom, has resulted in an existing building stock that is fairly young. Zhu and Lin [7] project that by 2015, half of China's existing building stock will have been built after 2000.

Effects are being made by the government to reduce the amount of energy consumed by buildings in China with more and more stringent building codes for residential, commercial and public buildings. However, despite the comparatively high rates of construction and demolition, most buildings in China do not comply with the latest national building codes. This paper aims to investigate the benefits of retrofitting existing residential buildings to the standard of the current building code in four main cities (Beijing, Shanghai, Wuhan and

Guangzhou) in China's main climate regions, and investigate if retrofitting to the current standard can nullify effects of future climate change in these regions in the next 100 years. Data from the climate prediction model, HadCM3, and impacts of climate change in the next 100 years will also be investigated.

1.1. Climate Zones in China

There are five main climatic zones in China: "Severe Cold", "Cold", "Hot Summer and Cold Winter", "Moderate", and "Hot Summer and Warm Winter", as shown in figure 1 [8]. They have distinctive characteristics, and thus different national building codes. This paper will investigate the situation in the "Cold" zone by using weather data for Beijing, the "Hot Summer and Warm Winter" zone with the city of Guangzhou. For the "Hot Summer and Cold Winter" zone, two cities are investigated, Shanghai which is a coastal city, with less varying temperatures between summer and winter, and Wuhan, which is located inland, and prone to even more extreme weather.

Fig.1, Climate zones in China and cities investigated

II. Climates in the Four Chosen Cities

The cities of Beijing, Shanghai, Wuhan and Guangzhou were selected to represent situations in three different climate zones in China. The buildings in each zone have very different heating and cooling needs. For example, Guangzhou in the "Hot Summer Warm Winter" zone has very little heating demand, as its winters are relatively mild. Wuhan, on the other hand, has significant heating and cooling loads. Shanghai is in the same climate region as Wuhan, but as it is located on the coast, thus it can be considered relatively milder over the whole year. Beijing also has hot summers, especially around August and early September, and winters are extremely harsh too. The changes in the next 100 years due to climate change for the three climate zones are also significantly different.

2.1. Current Climates

Long running series of real observed data from weather stations in the four cities were not available for this study, so in-depth study of the characteristics of different weather parameters could not be conducted. However, data can be extracted from the existing Test Reference Years, from the Energy Plus program. Fig 2 shows the average monthly range and values for daily maximum and minimum temperatures for the four selected cities.

2.2. Future Climate Change

The Hadley Centre model (HadCM3) [9] is used to provide future climate data for this study. HadCM3 is a global climate model developed at the Hadley Centre of the Met Office in the UK. It is a Coupled Atmosphere-Ocean General Circulation Model (AOGCM), in which the globe is divided into grid boxes, each measuring 2.50° x 3.75°. The gridboxes used in this paper are gridbox numbers 1952 (which encloses the area with latitude from 38.75°N to 41.25°N, and longitude from 114.375°E to 118.125°E, and includes the city of Beijing), 2337 (which encloses the area with latitude from 28.75°N to 31.25°N, and longitude from 118.125°E to 121.875°E, and includes the city of Shanghai), 2335 (which encloses the area with latitude from 28.75°N to 31.25°N, and longitude from 110.625°E to 114.375°E, and includes the city of Wuhan), and 2623 (which encloses the area with latitude from 21.25°N to 23.75°N, and longitude from 110.625°E to 114.375°E. and includes the city of Guangzhou).

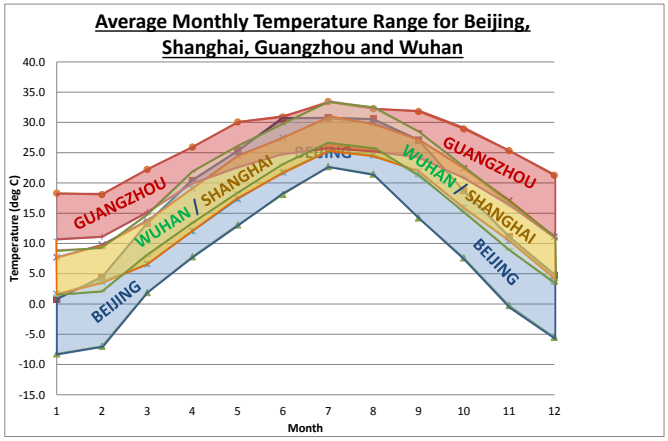


Fig.2, Current monthly average temperature range for Beijing, Shanghai, Wuhan and Guangzhou

Fig. 3 shows the extent of the gridbox for Shanghai. Unlike weather data from typical weather years, HadCM3 only provide daily values for parameters such as maximum, minimum and average temperatures, humidity, wind speed and downward short-wave flux (solar radiation), based on 4 main future scenarios on carbon emissions, A1F, A2, B2 and B1 [10]. For example, the A2 scenario describes a very heterogeneous world where slow and fragmented economic growth is assumed, together with a continuation of population growth and continued increase in CO_2 emission into the twenty-first century [11].

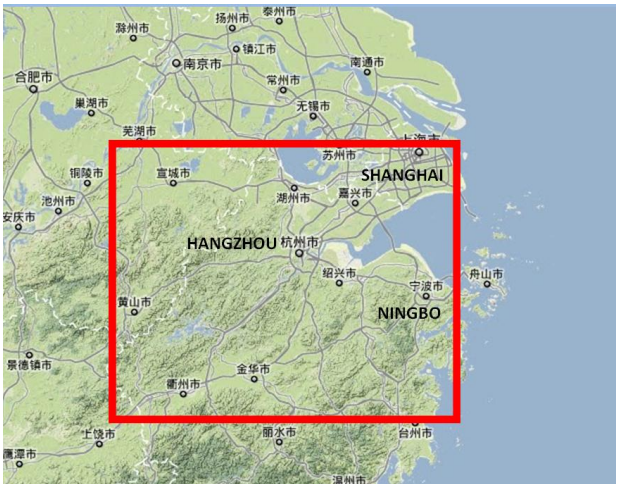


Fig.3, Area covered by gridbox 2337 of HadCM3

2.3. Compilation of Future Test Reference Years

Future Test Reference Years for Beijing, Shanghai, Wuhan and Guangzhou were constructed using the "morphing method" [12], which uses differences between monthly averages from "historical periods" and "future periods", and impose these onto existing Test Reference Year. For this study, data from HadCM3 are separated into four periods: 2000s (which includes all data from 1990-2009); 2020s (which includes data from 2010-2039); 2050s (which includes data from 2040-2069) and 2080s (which includes data from 2070-2099), and the average monthly temperatures under the A2 scenario from resulting TRYs for Shanghai is shown in Fig. 4

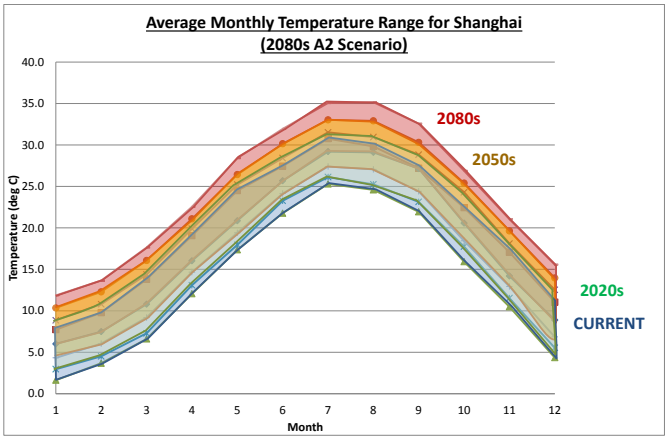


Fig.4, Changes in Temperature Range for Shanghai in current and future TRYs (A2 Scenario)

III. METHODOLOGY

The study used a typical residential building in the four cities and the whole year heating and cooling demands were calculated for buildings under the existing building specifications as well as standards complying with the latest building codes. This was conducted for current Test Reference Years and also Test Reference Years for 2020s, 2050s and 2080s, under different IPCC future scenarios.

3.1. Typical Building Used for Study

This study targets a mid-size multi-family residential apartment complex built in 2003, the first year that residential building energy codes were adopted in all four study cities. The selection of this particular building form is intended to be both realistic and strategic. A cheap and quick-to-construct answer to increasing housing needs, these types of buildings are ubiquitous in Chinese cities. These apartment buildings are in a 'grey area' of cultural value—although lacking in the prestige of historic structures or the glamour of new designs, they form the foundation for daily urban life for many people.

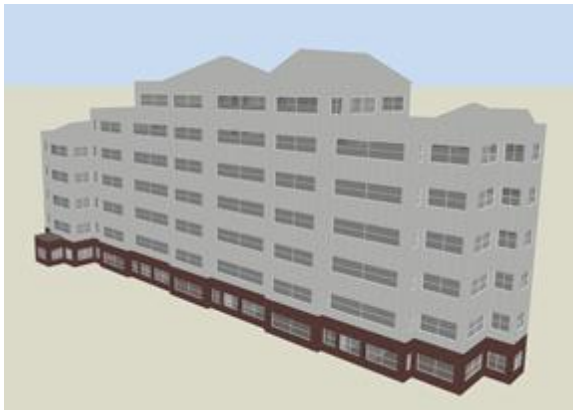


Fig.5, Isometric view of the study building

The standard of early building codes within which these buildings were constructed are quite low compared to current codes a decade later, and they are likely to be inadequately prepared for future climate change. Additionally, buildings constructed in the midst of the construction boom, which focused on speed over quality, are now showing severe maintenance problems which require refurbishment and repair. Figure 5 shows an isometric view of the building type being studied.

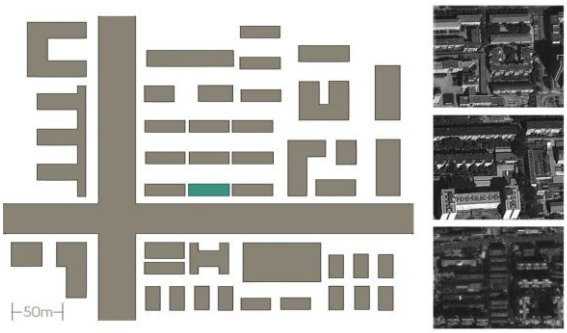


Fig.6, Theoretical site for the study building

TABLE 1, UPDATED ENVELOPE ELEMENT REQUIREMENTS IN THE MOST RECENT BUILDING CODES

U-Value ($W / m^2 K$)	Beijing		Shanghai / Wuhan		Guangzhou	
	Existing Building	Current Reg (JI2070-2012)	Existing Building	Current Reg (JGJ-134-2010)	Existing Building	Current Reg (DBJ15-51-2007)
Windows	2.7	1.8	3.2	2.8	3.8	3.5
Roof	0.54	0.35	0.8	0.8	1.0	0.9
Ext. Walls	0.76	0.40	1.0	1.0	1.5	1.5
Ground	0.53	0.45	1.5	1.5	1.5	1.5

The building and project site are both designed to be as generic as possible to facilitate high levels of comparison between the study cities. The buildings surrounding the project site are a composite of similar sites in each of the four cities, and represent a primarily-residential neighbourhood. The apartment complex itself, located on the corner of the main streets, consists of nine identical buildings with a north-south orientation. Each building is six stories tall and contains five apartments per floor. Based on the average values from surveys of similar apartment buildings conducted by Chen et al. [13, 14], Gu et al. [15], and Hu et al. [16], each apartment has 85 m2 of floor space, including two bedrooms, a bathroom, and a living room.. One building in the apartment complex, highlighted in blue in Figure 6 was used for the evaluation of retrofitting options.

3.2. Specification for "Current" Buildings

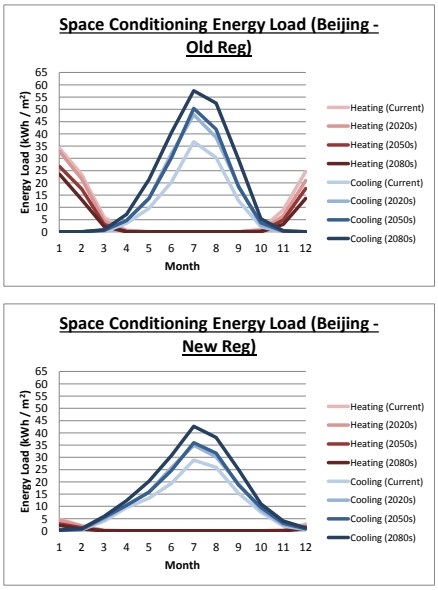
Although Rousseau and Chen [17] describe older apartment buildings as having solid brick construction with concrete floors, bricks were banned for construction in 1999 due to the increasing environmental impacts associated with their demand. Fernández [18] and Huang, et al. [19] report that most multi-family residential buildings constructed in the past two decades are made from reinforced concrete. The demand for concrete in China's building sector in general [20] also lends confidence that reinforced concrete is the most representative material for this building type.

3.3. Specification for Renovated Buildings

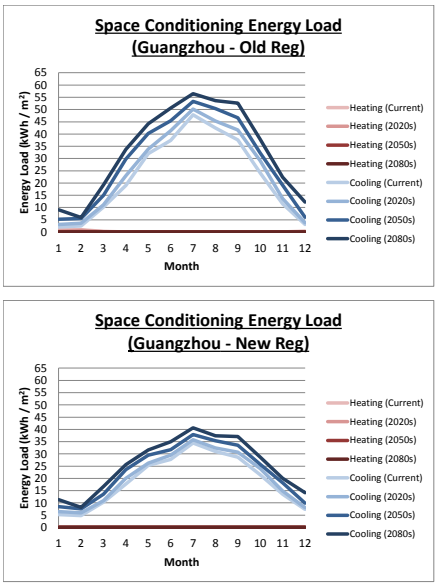
While the most recent residential building codes in China are stricter than those in 2003, they have nonetheless been developed based on historic climate knowledge. While lower current energy consumption may indicate greater climate resiliency, it cannot be taken for granted that buildings built to current codes will maintain higher performance levels under future conditions. To compare the climate vulnerability of the study building against a newer structure, an equivalent but code-compliant building was also modelled. This building represents both a building of the same size constructed in 2014, which could replace the existing structure entirely, and the study building if it were retrofitted to meet cur-rent codes without any other measures. Table 1 lists the envelope requirements which have improved since 2003. The comfort range where no heating and cooling would be supplied to the building was set as 16°C to 24°C for all four cities, and the design air-tightness in renovated buildings are improved from 3.0 ach to 1.0 ach. Heating was assumed to use gas boiler, with a total efficiency of 0.83, and fan-coil units used for cooling, with a COP of 1.67.

IV. RESULTS AND DISCUSSION

Fig 7a shows the monthly space conditioning (heating and cooling loads) requirements for the tested existing and retrofitted residential building in Beijing, under the A2 future scenario. It can be seen that with the existing building specification, the ratio between cooling and heating starts to drift in the direction of more cooling with climate change, and with the same building, but retrofitted to the latest building regulation for Beijing, there is a major decrease in heating demand, but the effect on cooling demand is relatively small in comparison. The cooling period also appears to have extended more to include the months of March and November. Similar results were achieved in Shanghai and Wuhan, though with a much smaller peak cooling load in Shanghai than Wuhan and Beijing. The results for Guangzhou is shown in fig 7b, where it can be seen that the benefit of the new building code does not benefit the reduction of heating much, as Guangzhou has minimal heating demand anyway. The peak cooling demand in summer for a retrofitted building reaches approximately the same level as an existing building under the current climate, so the retrofitting can just about nullify the effects of future climate change here.



(a)



(b)

Fig.7, Changes in Heating and Cooling Loads in the Residential Building in Beijing (a) and Guangzhou (b) under Future Climate Change (A2 Scenario)

Fig 8 shows the range of changes in cooling consumption in future periods under various scenarios for typical existing residential buildings (in grey), and those that have been retrofitted to the standard of the current building code for the respective region (in blue). Under the current climate, the energy consumption for cooling is actually higher for a retrofitted building in Beijing and Shanghai. The overlap occurs in the 2020s and after that the retrofitted building performs better. No overlap occurs for either Wuhan or Guangzhou. It is also worth noting that the range for cooling energy consumption is smaller in a retrofitted building, suggesting that these are less susceptible to different future scenarios.

The effect of retrofitting has a major effect on the reduction of heating demand, and there are large gaps between the ranges of heating energy consumption between existing and retrofitted buildings in all four cities, usually by a factor of 6 to 8. Apart from the reduction in U-Values of some of the external elements, the main contribution to a much lower heating demand is attributed to the better air-tightness specification of the retrofitted buildings. The effect of which on reducing cooling load is not as effective (solar radiation can still enter, and it is also more difficult to remove other internal heat gains with better insulated external elements).

V. Conclusion

This paper studied the effects of retrofitting residential buildings in cities in three major climate zones (with four cities) of China, from building standards set in 2003 to the current building codes, and seeing how this would perform under climate change in the next 100 years. The effects of retrofitting was most profound for reducing heating energy consumption, but this may be due to the effects of having introduced much tighter building envelopes which may not be as well incorporated in practice. Guangzhou, with practically no heating demand in its warm winters, does not benefit from the reduced heating demand with climate change, and with retrofitting to the current code, there is only a small amount of reduction in cooling loads, which is not enough to counter the effects of future climate change, regardless of which scenario the world follows in the future. The same could be said for cooling in all four cities as even with the least severe future scenario, the cooling energy demand in a retrofitted building is higher than that for a non-retrofitted building under the current climate. This happens before the 2020s for Beijing and Shanghai, slightly later for Wuhan (just before the 2050s) and at around 2050s for Guangzhou. With more severe future scenarios, this will happen even quicker.

With all four cases, it is clear from the results that the current building standards will not be able to nullify the effects of future climate change, as far as cooling is concerned, and more studies are required to investigate optimum building standards for each different climate region in China, which must also be economically feasible. This study has not taken the effects of urban heat islands, which will make the quest for producing low-energy buildings even more challenging, and further work is required to investigate the effects of more reasonable assumptions of air-tightnesss in residential buildings in China.

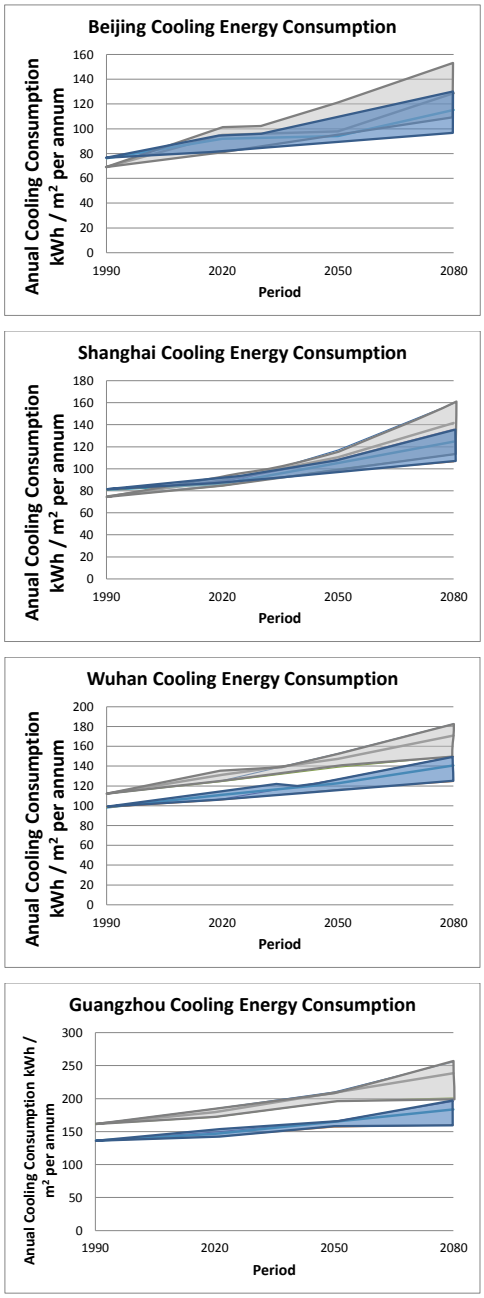


Fig.8, Changes in Heating and Cooling Loads in the Residential Building in Guangzhou under Future Climate Change (A2 Scenario)

REFERENCES

[1] Y. Wu, "China Building Energy Efficiency: Current Status, Issues, and Policy Recommendations" China Ministry of Construction, 2003.

[2] N. Zhou, "Energy Use in Commercial Buildings in China: Current Situation and Future Scenarios" 8th ECEEE Summer Study, Lawrence Berkeley National Laboratory, 2007. pp. 1065-1071.

[3] D. G. Fridley, N. Zheng, N. Zhou, "Estimating Total Energy Consumption and Emissions of China's Commercial and Office Buildings" IBNL-248E, Lawrence Berkeley National Laboratory, 2008.

[4] J. Zhong, J. "Development Approach of Renewable Energy and Building Integration Technology" Solar Energy Vol. 5, 2007. pp. 40-43.

[5] Y. Zheng, "Energy-Efficiency in Buildings:- An Imminent Task" Environmental Economy Vol. 11, 2007. pp. 23-24.

[6] T. Hong, "A Close Look at the China Design Standard for Energy Efficiency of Public Buildings" Energy and Buildings, Vol. 41, Issue 4, 2009. pp. 426-435.

[7] Zhu, X. and Lin, B., "Sustainable housing and urban construction in China" Energy and Buildings, Vol. 36, 2004, pp. 1287-1297.

[8] G. Heggelund, A. Meier, S. Ohshita, S. Wiel, "Cooperative Climate", International Institute for Sustainable Development, 2008, Chapter 5.

[9] C. Gordon et al. "A Simulation of SST, Sea Ice Extents Ocean Heat Transports in a Version of the Hadley Centre Coupled Model without Flux Adjustments", Climate Dynamics, Vol. 16, 2000. pp. 147-168.

[10] N. Nakicenovic "Emissions Scenarios" Intergovernmental Panel on Climate Change, Cambridge University Press. ISBN 0521 80493 0. 2000.

[11] WMO, UNEP "IPCC Climate Change 2001, Summary for Policy Makers and Technical Summary of the Working Group I Report", 2001. pp. 63-65.

[12] S. Belcher, J. Hacker, D. Powell, "Constructing design weather data for future climates" Building Services.Engineering Research and Technology Vol. 26.1, (2005) pp.49–61.

[13] Chen, S., Yoshino, H., Levine, M. D. & Li, Z., "Contrastive analyses on annual energy consumption characteristics and the influence mechanism between new and old residential buildings in Shanghai, China, by the statistical methods." Energy and Buildings, Volume 41, 2009. pp. 1347-1359.

[14] Chen, S., Yoshino, H. & Li, N., "Statistical analyses on summer energy consumption characteristics of residential buildings in some cities of China." Energy and Buildings, Volume 42, 2010. pp. 136-146.

[15] Gu, Z., Sun, Q. & Wennersten, R., "Impact of urban residences on energy consumption and carbon emissions: An investigation in Nanjing, China." Sustainable Cities and Society, Volume 7, 2013. pp. 52-61.

[16] Hu, T., Yoshino, H. & Jiang, Z., "Analysis on urban residential energy consumption of Hot Summer & Cold Winter Zone in China." Sustainable Cities and Society, Volume 6, 2013. pp. 85-91.

[17] Rousseau, D. & Chen, Y., "Sustainability options for China's residential building sector." Building Reearch & Information, Vol. 29, Issue 4, 2001. pp. 293-301.

[18] Fernández, J. E., "Resource Consumption of New Urban Construction in China." Journal of Industrial Ecology, Vol. 11, Issue 2, 2007. pp. 99-115.

[19] Huang, T. et al., "Materials demand and environmentla impact of buildings construction and demolition in China based on dynamic material flow

analysis." Resources, Conservation and Recycling, Volume 72, 2013. pp. 91-101.

[20] Chang, Y., Ries, R. J. & Wang, Y., "The embodied energy and environmental emissions of construction projects in China: An economic input-output LCA model." Energy Policy, Vol. 38, 2010. pp. 6597-6603.

[21] Wang, Y., Huang, Z. & Heng, L., "Cost-effective assessment of insulated exterior walls of residentail buildings in cold climate." International Journal of Project Management, Vol. 25, 2007.pp. 143-149.

[22] Aden, N., Qin, Y. & Fridley, D., "Lifecycle Assessment of Beijing-Area Building Energy Use and Emissions: Summary Findings and Policy Applications (LBNL-3939E)", Berkeley, CA: China Energy Group, Lawrence Berkeley National Laboratory. 2010.

[23] Chen, S. et al., "Statistical analyses on winter energy consumption characteristics of residential buildings in some cities of China." Energy and Buildings, Vol. 43, 2011. pp. 1063-1070.

[24] Anderson, B., "Thermal properties of buliding structures." In: K. Butcher, ed. CIBSE Guide A: Environmental Design. 7th ed. Norwich: The Chartered Institution of Building Services Engineers, 2006. pp. 3-1 - 3-46.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 2

A multiscale-multiphysics strategy for numerical modeling of thin piezoelectric sheets

薄压电片多尺度多物理场的数值模拟

Claudio Maruccio[1*], Laura De Lorenzis[2], Luana Persano[3] and Dario Pisignano[3,4]

[1]*Department of Innovation Engineering, University of Salento, via Monteroni, Lecce, Italy, claudio.maruccio@unisalento.it*

[2]*Institut für Angewandte Mechanik, Technische Universität Braunschweig, Germany, l.delorenzis@tu-braunschweig.de*

[3]*National Nanotechnology Laboratory, CNR-Istituto Nanoscienze, Lecce, via Arnesano, Italy, luana.persano@nano.cnr.it*

[4] *Department of Mathematics and Physics "E. De Giorgi", University of Salento, via Monteroni, Lecce, Italy, dario.pisignano@unisalento.it*

Accepted for publication on 26th July 2014

Abstract - **Flexible piezoelectric devices made of polymeric materials are widely used for micro- and nano-electro-mechanical systems. In particular, numerous recent applications concern energy harvesting. Due to the importance of computational modeling to understand the influence that microscale geometry and constitutive variables exert on the macroscopic behavior, a numerical approach is developed here for multiscale and multiphysics modeling of piezoelectric materials made of aligned arrays of polymeric nanofibers. At the microscale, the representative volume element consists in piezoelectric polymeric nanofibers, assumed to feature a linear piezoelastic constitutive behavior and subjected to electromechanical contact constraints using the penalty method. To avoid the drawbacks associated with the non-smooth discretization of the master surface, a contact smoothing approach based on Bézier patches is extended to the multiphysics framework providing an improved continuity of the parameterization. The contact element contributions to the virtual work equations are included through suitable electric, mechanical and coupling potentials. From the solution of the micro-scale boundary value problem, a suitable scale transition procedure leads to the formulation of a macroscopic thin piezoelectric shell element.**

Keywords – **Electromechanical Coupling, Multiphysics Modeling, Multiscale Modeling, Piezoelectricity, Shell elements.**

I. INTRODUCTION

The discovery of appreciable piezoelectricity on ceramic and polymeric materials such as ZnO and Polyvinylidene fluoride (PVDF) and its copolymers has led to the development of a series of Nano Wire-based piezoelectric nanogenerators [1]. Most recent models have showed the capability of powering small electronic devices. In particular, PVDF is a polymeric piezoelectric material with good piezoelectric and mechanical properties. Its piezoelectric coefficient is around -30 pC/N, which is higher than that of ZnO. Its flexible polymeric nature allows very high strains to be applied to a PVDF beam/sheet, thus high piezoelectric potentials can be expected. Far-field electrospinning was recently used [2] to produce piezoelectric sheets of PVDF nanofibers directly patterned onto the substrate (Fig.1). The stretching force due to the strong electric field applied between the nozzle and the substrate surface can pole the PVDF nanofibers into the β phase, which has a resulting component of the global polarization along the longitudinal direction of the resulting fibers. In previous experiments, the PVDF sheet was directly applied between two electrodes on a flexible plastic substrate for piezoelectric output measurement [2] .

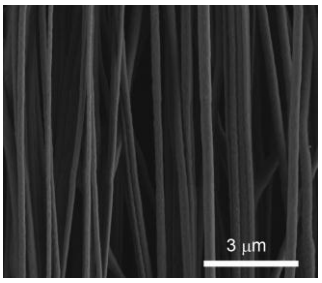


Fig.1. Scanning electron micrograph of a piezoelectric PVDF sheet.

The objective of this paper is to model the behavior of the aforementioned PVDF sheets made of polymeric nanofibers by a multiscale and multiphysics approach. This requires the definition of a representative volume element (RVE) at the microscale, the formulation and solution of a microscale boundary value problem (BVP), and the development of a

suitable micro-macro scale transition. Due to the small thickness, $h_M << l_M$, a discretization with shell elements at the macroscale is numerically efficient (Fig.2).

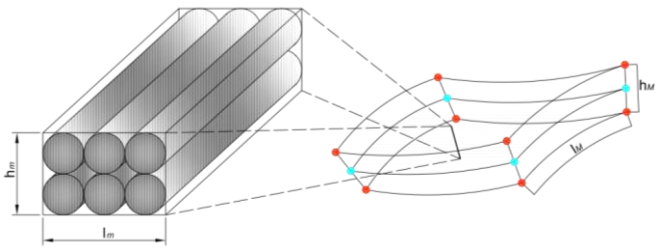


Fig.2. Determination of an RVE. h_m and l_m are the RVE thickness and length at the microscale, while h_M and l_M are the thickness and length at the macroscale, $l_m << l_M$.

The first part of the paper describes the kinematic behaviour of a piezoelectric shell following [3-5]. In the second part a microscale RVE element is defined and the theory of linear piezoelasticity [6] is briefly introduced along with its finite element formulation [7,8] at the microscale. Moreover some details about the formulation and implementation of suitable electromechanical contact elements using smoothing techniques [9,10] are provided. The third part describes numerical homogenization procedures for piezoelectric nanocomposites [11-15]. A multiscale approach is introduced to perform the scale transition between the micro- and macroscales. Details regarding the linearization of the finite element equations and implementation of this multiscale formulation are also provided. Finally, based on the presented multiscale and multiphysics framework the RVE geometry is analyzed to determine the effective material properties suitable for a macroscopic shell element definition. Advanced symbolic computational tools available in the AceGen-AceFem finite element environment within Mathematica [16,17] are used throughout this work, with the advantage that the tasks related to the finite element implementation are largely automated.

II. Macroscale Formulation

Based on the analysis of the shell kinematics, the Green-Lagrangean strains and the electric field components in convective coordinates can be arranged in a generalized strain column vector:

$$\boldsymbol{\varepsilon} = [\varepsilon_{11}, \varepsilon_{22}, 2\varepsilon_{12}, \kappa_{11}, \kappa_{22}, 2\kappa_{12}, \gamma_1, \gamma_2, E_1, E_2, \varepsilon^0_{33}, \varepsilon^1_{33}, E^0_3, E^1_3]^T \quad (1)$$

where the membrane strain components $\varepsilon_{\alpha\beta}$ and the change of curvature $\kappa_{\alpha\beta}$ read:

$$\varepsilon_{\alpha\beta} = \frac{1}{2}(\boldsymbol{\psi}_{,\alpha} \cdot \boldsymbol{\psi}_{,\beta} - \boldsymbol{\psi}_{\mathbf{0},\alpha} \cdot \boldsymbol{\psi}_{\mathbf{0},\beta}); \quad (2)$$

$$\kappa\alpha\beta = \frac{1}{2}(\boldsymbol{\psi}_{,\alpha} \cdot \boldsymbol{d}_{,\beta} + \boldsymbol{\psi}_{,\beta} \cdot \boldsymbol{d}_{,\alpha} - \tfrac{h_0}{2}\boldsymbol{\psi}_{\mathbf{0},\alpha} \cdot \boldsymbol{g}_{,\beta} - \tfrac{h_0}{2}\boldsymbol{\psi}_{\mathbf{0},\beta} \cdot \boldsymbol{g}_{,\alpha}); \quad (3)$$

In the previous equations the comma indicates partial derivation, Greek indices take the values 1, 2; ξ_1, ξ_2 are the natural coordinates of the shell middle surface, $\boldsymbol{\psi}, \boldsymbol{\psi_0}$ are respectively the current and initial position vectors of the shell middle surface, **g** and **a** are the initial and current normal (shell directors) and h_0 is the initial shell thickness. Moreover we introduced the quantity:

$$\boldsymbol{d} = \frac{h_0}{2}\lambda(\xi_1, \xi_2)\boldsymbol{a}(\xi_1, \xi_2); \quad (4)$$

where $\lambda(\xi_1, \xi_2)$ is the thickness stretch. The shear strain components take the form:

$$\gamma_\alpha = (\boldsymbol{\psi}_{,\alpha} \cdot \boldsymbol{d} - \frac{h_0}{2}\boldsymbol{\psi}_{\mathbf{0},\alpha} \cdot \boldsymbol{g}) \quad (5)$$

The electric components read:

$$E_\alpha = \frac{\partial \phi}{\partial \xi_\alpha}, \quad (6)$$

where ϕ is the electric potential. Furthermore, $\varepsilon^0_{33}, \varepsilon^1_{33}$ are the constant and linear components of the thickness strain and E^0_3, E^1_3 represent the constant and linear parts of the electric field along the thickness direction.

III. Microscale Formulation

At the microscale, the RVE consists in piezoelectric polymer fibers that feature a linear piezoelastic constitutive behavior and are subjected to electromechanical contact constraints. Piezoelectric problems are those in which an electric potential gradient causes deformation, and viceversa. The governing equations are the Navier equations and the strain-displacement relations for the mechanical field and Gauss and Faraday laws for the electrostatic field [6,7]. Moreover, the constitutive equations are:

$$a)\ T_{ij} = C_{ijkl}S_{kl} - e_{kij}E_k;\ \ b)\ D_i = e_{ikl}S_{kl} + \tilde{\epsilon}_{ik}E_k \quad (7)$$

where C_{ijkl}, e_{ikl}, and ϵ_{ik} are respectively the elastic, piezoelectric, and permittivity constants, whereas S_{ij}, T_{ij} are the strain and stress components and D_i, E_i are the electric displacement and the electric field components, respectively. Clearly, the coupling between mechanical and electric fields is determined by the piezoelectric coefficients. The 3D electromechanical frictionless contact element is implemented with the following main characteristics:

- the contact formulation is based on the master-slave concept;
- Bézier patches are used for smoothing of the master surface;
- The impenetrability condition is extended to the electromechanical setting by imposing equality of the electric potential in case of closed contact. The electromechanical constraints are regularized with the penalty method.

For each slave node, the normal gap is computed as:

$$g_N = (\boldsymbol{x_s} - \boldsymbol{x_m}) \cdot \boldsymbol{n} \quad (8)$$

where $\boldsymbol{x_s}$ is the position vector of the slave node, $\boldsymbol{x_m}$ is the position vector of its normal (i.e. minimum distance) projection point onto the master surface, and $\boldsymbol{n}$ is the outer normal to the master surface at the projection point. The sign

of the measured gap is used to discriminate between active and inactive contact conditions, a negative value of the gap leading to active contact. The electric field requires the definition of the contact electric potential jump:

$$g_\phi = (\phi_s - \phi_m) \tag{9}$$

where ϕ_s and ϕ_m are the electric potential values in the slave node and in its projection point on the master surface. A tensor product representation of one-dimensional Bézier polynomials is used to interpolate the master surfaces in the contact interface for a three-dimensional problem according to the relation:

$$x(\xi_1, \xi_2) = \sum_{k=0}^{m} \sum_{l=0}^{m} \boldsymbol{d_{kl}} B_k^m(\xi_1) B_l^m(\xi_2) \tag{10}$$

where $B_k^m(\xi_1)$ and $B_l^m(\xi_2)$ are the Bernstein polynomials and $\boldsymbol{d}_{kl}$ the coordinates of the control points. With the same procedure, to interpolate the electric potential on the master surface the following relation is introduced:

$$\phi(\xi_1, \xi_2) = \sum_{k=0}^{m} \sum_{l=0}^{m} \phi_{kl} B_k^m(\xi_1) B_l^m(\xi_2) \tag{11}$$

where ϕ_{kl} is the potential evaluated at 16 control points as a function of the potential at the auxiliary points $\hat{\phi}_{kj}$ and at the master nodes ϕ_{ij}^2. According to standard finite element techniques, the global set of equations can be obtained by adding to the variation of the energy potential representing the continuum behavior the virtual work associated to the electromechanical contact contribution provided by the active contact elements. If $\hat{\boldsymbol{u}}$ is the set of degrees of freedom (DOF) used to discretize the displacement field $\mathbf{u} = \mathbf{u}(\hat{\boldsymbol{u}})$, and $\hat{\boldsymbol{\phi}}$ is the set of DOF used to discretize the electric potential field $\varphi=\varphi(\hat{\boldsymbol{\phi}})$, so that $\hat{\boldsymbol{u}} \cup \hat{\boldsymbol{\phi}}$ is the vector of all nodal DOF, Π_{global} is the global energy of the discretized system, the residual vector and the stiffness matrix terms resulting from the finite element discretization are determined according to:

$$\begin{aligned}
&\boldsymbol{R}_{u_i} = \frac{\delta \Pi_{global}}{\delta \hat{\mathbf{u}}_i}\,;\; R_{\phi_i} = \frac{\delta \Pi_{global}}{\delta \hat{\phi}_i}\,;\\
&\boldsymbol{K}_{uu_{i,j}} = \frac{\delta R_{u_i}}{\delta \hat{\mathbf{u}}_i}\,;\; K_{\phi\phi_{i,j}} = \frac{\delta R_{\phi_i}}{\delta \hat{\phi}_j}\,;\\
&\boldsymbol{K}_{u\phi_{i,j}} = \frac{\delta \boldsymbol{R}_{u_i}}{\delta \hat{\phi}_j}\,;\; \boldsymbol{K}_{\phi u_{i,j}} = \frac{\delta R_{\phi_i}}{\delta \hat{\mathbf{u}}_j}\,.
\end{aligned} \tag{12}$$

IV. Homogenization procedure

The main idea of homogenization is to find a globally homogeneous medium equivalent to the original composite, where the equivalence is intended in an energetic sense as per Hill's balance condition. Coupling between the macroscopic and microscopic scales is here based on averaging theorems, see [18,19]. Formulated for the electromechanical problem at hand, Hill's criterion in differential form reads:

$$\bar{T}_{ij}\delta\bar{S}_{ij} + \bar{D}_i\delta\bar{E}_i = \frac{1}{V}\iiint T_{ij}\delta S_{ij}dV + \frac{1}{V}\iiint D_i\delta E_i\,dV \tag{13}$$

and requires that the macroscopic volume average of the variation of work performed on the RVE is equal to the local variation of work on the macroscale. In the previous equation: $\bar{T}_{ij}$, $\bar{S}_{ij}$, $\bar{D}_i$ and $\bar{E}_i$ represent respectively the average values of stress, strain, electric displacement and electric field components. Hill's lemma leads to the following equations:

$$\begin{aligned}
\bar{T}_{ij} &= \frac{1}{V}\iiint T_{ij}dV; \quad \bar{S}_{ij} = \frac{1}{V}\iiint S_{ij}dV;\\
\bar{D}_i &= \frac{1}{V}\iiint D_i dV; \quad \bar{E}_i = \frac{1}{V}\iiint E_i dV
\end{aligned} \tag{14}$$

Classically three types of boundary conditions are used for an RVE: prescribed displacements, prescribed tractions and periodic boundary conditions. Among them, periodic boundary conditions lead to a more reasonable estimation of the effective properties and therefore are often preferred. In this work, periodic boundary conditions are expressed as linear constraints and are implemented in Acegen as multipoint constraints using the Lagrange multiplier method. For simplicity, meshing of the RVE is performed uniformly such that identical nodes are present on all faces of the RVE. The final constitutive equation in the homogenized setting reads:

$$\begin{pmatrix}\bar{\boldsymbol{T}}\\ \bar{\boldsymbol{D}}\end{pmatrix} = \begin{pmatrix}\bar{\boldsymbol{C}} & -\bar{\boldsymbol{e}}\\ \bar{\boldsymbol{e}} & \bar{\bar{\boldsymbol{\epsilon}}}\end{pmatrix}\begin{pmatrix}\bar{\boldsymbol{S}}\\ \bar{\boldsymbol{E}}\end{pmatrix} \tag{15}$$

or in compact form:

$$\mathbb{T} = \mathbb{D}_{\boldsymbol{solid}}^{\boldsymbol{macro}}\mathbb{S} \tag{16}$$

where $\bar{C}_{ijkl}$, $\bar{e}_{ikl}$, and $\bar{\epsilon}_{ik}$ are respectively the homogenized elastic, piezoelectric, and permittivity constants. With some algebra and after integration on the shell thickness the previous equation becomes:

$$\mathbb{L} = \mathbb{D}_{\boldsymbol{shell}}^{\boldsymbol{macro}}\,\boldsymbol{\varepsilon} \tag{17}$$

with:

$$\mathbb{L} = [n_{11}, n_{22}, n_{12}, m_{11}, m_{22}, m_{12}, q_1, q_2, -d_1, -d_2, n_{33}^0, n_{33}^1, -d_3^0, -d_3^1]^T \tag{18}$$

where $n_{\alpha\beta}$ are the membrane forces, $m_{\alpha\beta}$ are the bending moments, q_α are the shear forces and $n_{33}^0, n_{33}^1, -d_3^0, -d_3^1$ are the constant and linear components of membrane force and dielectric displacement in the thickness direction.

V. Results

If an RVE with PVDF fibers aligned in one direction is considered, the final material constitutive equations will be those of a transversely isotropic piezoelectric solid, and consequently the stiffness matrix, the piezoelectric matrix and the dielectric matrix will simplify so there remain in all 10 independent coefficients. In this work it is assumed that the elastic modulus E =150N/mm^2 and Poisson ratio ν= 0.29, while the piezoelectric strain coefficients d_{31}, d_{32}, d_{33} are equal to $20*10^{-12}$, $3*10^{-12}$, $-35*10^{-12}$ m/V, respectively, and the permittivity coefficients $\bar{\bar{\epsilon}}_{11}$, $\bar{\bar{\epsilon}}_{22}$, $\bar{\bar{\epsilon}}_{33}$ are all equal to 12 $\bar{\bar{\epsilon}}_0$, where $\bar{\bar{\epsilon}}_0$ = $8.854*10^{-12}$ F/m is the void permittivity. According to the experimental results, each fiber in the RVE is assumed to have a radius R = 1.0μm. Hence, the RVE has a cubic geometry with a side length (L=2R) of 2 micron. For the evaluation of the effective properties, suitable boundary conditions have to be applied to the unit cell in such a way that,

apart from one component of the strain/electric field vector, all other components are equal to zero [20, 21]. Then each effective coefficient can be easily determined by multiplying the corresponding row of the material matrix by the strain/electric field vector. In Fig.3 the mesh and the boundary conditions used during the analysis are illustrated. In particular different colors correspond to different degrees of freedom constrained during the analysis. Moreover, in figures 4 and 5 the contours of displacements and stresses are provided for boundary conditions allowing only stretching along the fiber longitudinal axis.

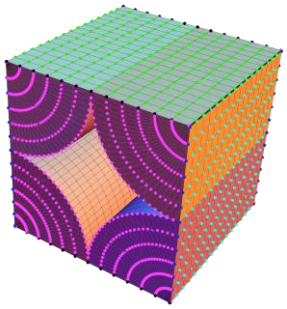
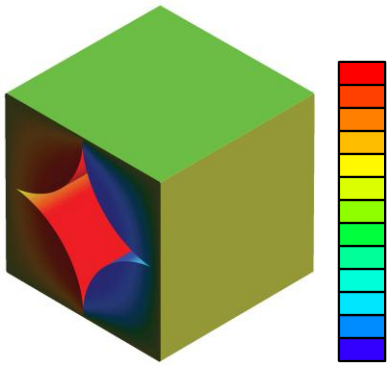
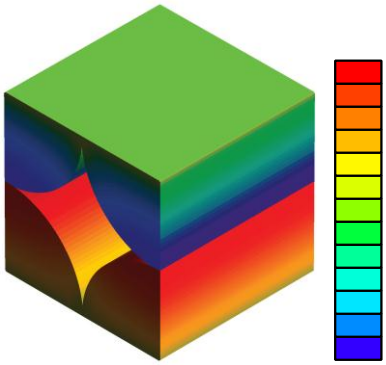


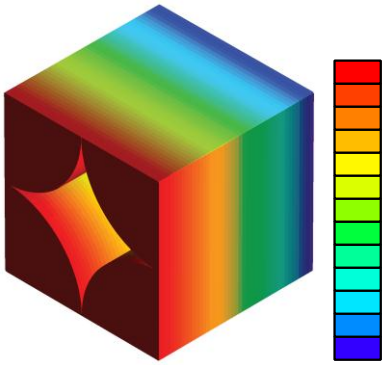
Fig.3. RVE mesh and boundary conditions.



a) u_1

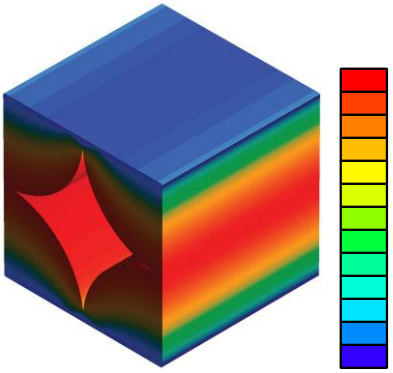


b) u_2

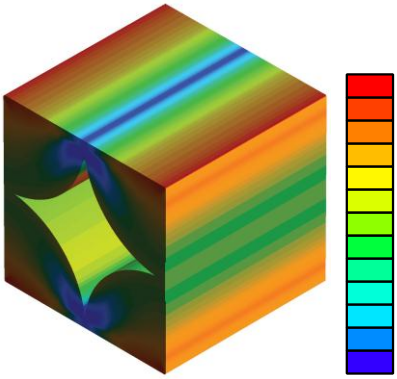


c) u_3

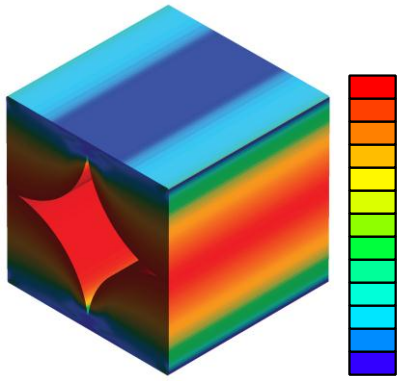
Fig.4. Microscale behavior, countour levels of displacement distribution in the RVE. Maximum and minimum corresponding to the vertical color scales are: a) +/- 0.104 [μm]; b) +/- 0.18 [μm]; c) -0.1/0 [μm].



a) T_{11}



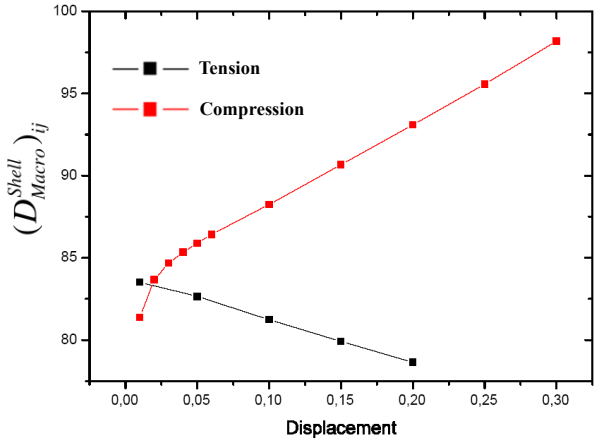
b) T_{22}



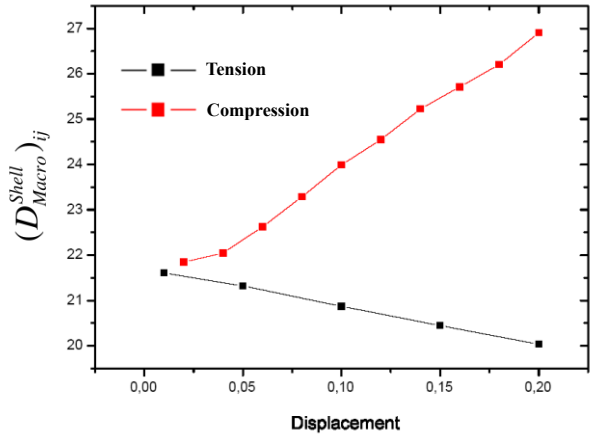
c) T_{33}

Fig.5. Microscale behavior, countour levels of stress distribution in the RVE. Maximum and minimum corresponding to the vertical color scales are: a) -2.39/+0.6 [N/mm^2]; b) -2.34/+0.48 [N/mm^2]; c) -6.43/-4.97 [N/mm^2].

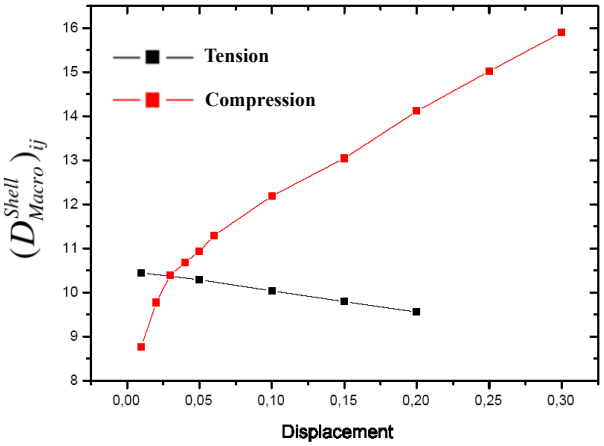
Following the aforementioned procedure, it is possible to determine the evolution of the material coefficients under prescribed increasing displacements. The plots of the coefficients $D_{1,11}$, $D_{1,1}$ and $D_{2,2}$, $D_{1,2}$, $D_{11,11}$, are provided in Fig.6. Nonlinear behavior is due to the fiber interaction through electromechanical contact. In the figures, displacements are in micron and coefficients in MPa.



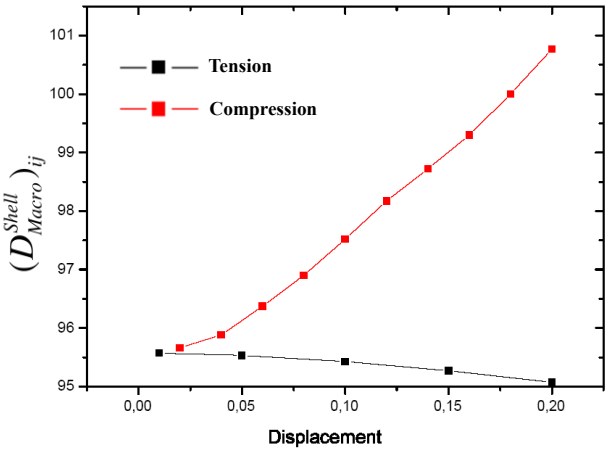
a) i=1=j=11



b) i=1 and j=2 or i=1 and j=1



c) i=1 and j=11



d) i=1 and j=1 or i=2 and j=2

Fig.6. Evolution of the effective coefficients of the shell: $D_{1,11}$, $D_{1,1}$ and $D_{2,2}$, $D_{1,2}$, $D_{11,11}$ as functions of the applied displacements. Coefficients are in N/mm^2 and displacements are in μm.

VI. Conclusion

In this paper a multiscale and multiphysics homogenization method to model fibrous piezoelectric materials is presented. For each macroscale material point a fine scale boundary value problem is solved to determine the local material behavior on the macroscale. On the macroscopic level thin array of fibers are characterized by a large area to thickness ratio, such that a discretization with shell elements is numerically efficient. The material behavior is strongly influenced by the heterogeneous microstructure. Within homogenization based approaches the macroscopic constitutive behavior of the inhomogeneous material is modelled by means of an appropriate microscale RVE. The RVE consists in piezoelectric polymer fibers, assumed to feature a linear piezoelastic constitutive behavior and subjected to electromechanical contact constraints. A two-step homogenization scheme is applied to transfer the microscopic response to the macrolevel. Theoretical aspects are discussed and preliminary results for periodic structures are presented. Hence the computational homogenization

procedure proposed is suitable for numerical modeling of flexible piezoelectric devices for energy harvesting technologies. Possible applications range from powering of remote sensors to extensive collection of energy. In the next future the development of these technologies will promote energy microgeneration and production of "Km 0 energy" in opposition to "macrogeneration" produced in massive plants that have a huge negative impact on the surrounding environment.

ACKNOWLEDGEMENTS

Claudio Maruccio acknowledges the support from the Italian MIUR through the project FIRB Futuro in Ricerca 2010 Structural mechanics models for renewable energy applications (RBFR107AKG). Laura De Lorenzis, Dario Pisignano and Luana Persano acknowledge the support from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013), ERC Starting Grants INTERFACES (L. De Lorenzis, grant agreement n. 279439) and NANO-JETS (D. Pisignano and L. Persano, grant agreement n. 306357).

REFERENCES

[1] Y. Gao and Z.L. Wang. Electrostatic potential in a bent piezoelectric nanowire. The fundamental theory of nanogenerator and nanopiezotronics. *Nano Lett.* 7, 2499-2505, 2007.

[2] L. Persano, C. Dagdeviren, Y. Su, Y. Zhang, S. Girardo, D. Pisignano, Y. Huang, J. A. Rogers. High performance piezoelectric devices based on aligned arrays of nanofibers of PVDF. *Nature Communications*, 1633, 4, 2013.

[3] K. Schulz, S. Klinkel, W. Wagner. A finite element formulation for piezoelectric shell structures considering geometrical and material non-linearities. *Int. J. Numerical Methods in Engineering* 2011; 87:491–520.

[4] S. Klinkel, W. Wagner. A piezoelectric solid shell element based on a mixed variational formulation for geometrically linear and nonlinear applications. *Computers and Structures* 2008; 86:38–46.

[5] S. Klinkel, F. Gruttmann, W. Wagner. A mixed shell formulation accounting for thickness strains and finite strain 3d material models. *Int. J. Numerical Methods in Engineering* 2008; 75:945–970.

[6] J. Yang. An introduction to the theory of piezoelectricity. *Springer Science + Business Media Inc*, 2005.

[7] J. Yang. The mechanics of piezoelectric structures. *Springer*, New York, 2006.

[8] A. Benjeddou. Advances in piezoelectric finite element modeling of adaptive structural elements: a survey. *Computers and Structures*, 76, 2000, 347-363.

[9] C. Weissenfels and P. Wriggers. Numerical modeling of electrical contacts. *Computational Mechanics* 46, 301-314, 2010.

[10] J. Lengiewicz, J. Korelc, S. Stupkiewicz. Automation of finite element formulations for large deformation contact problems. *International Journal for Numerical Methods in Engineering*, Vol. 85, 1252-1279, 2011.

[11] J. Schroeder and M. Keip. Two-scale homogenization of electromechanically coupled boundary value problems - Consistent linearization and applications. *Computational Mechanics*, 50(2): 229-244, 2012.

[12] V. Kouznetsova, M. G. D. Geers and W. A. M. Brekelmans. Multi-scale constitutive modelling of heterogeneous materials with a gradient-enhanced computational homogenisation scheme. *Int. J. Numer. Meth. Eng.* 54, 1235-1260, 2002.

[13] C. Miehe. Computational micro-to-macro transitions for discretized microstructures of heterogeneous materials at finite strains based on the minimization of averaged incremental energy. *Comput. Methods Appl. Mech. Eng.* 2003;192:559–91.

[14] S. Fillep, J. Mergheim, P. Steinmann. Computational modelling and homogenization of technical textiles. *Eng. Structures* 50, 2013; 68–73.

[15] E. W. Coenen, V. G. Kouznetsova and M. G. D. Geers. Computational homogenization for heterogeneous thin sheets. *International Journal for Numerical Methods in Engineering*, Vol. 83, 1180-1205, 2010.

[16] J. Korelc. Automation of primal and sensitivity analysis of transient coupled problems. 54(2-3), 161–167, 2001. *Computational Mechanics*, 44(5):631-649, 2009.

[17] J. Korelc. Multi-language and multi-environment generation of nonlinear finite element codes. *Engineering with Computers*, 18(4):312-327, 2002.

[18] V. Kouznetsova, W. A. M. Brekelmans and F. P. T. Baaijens. An approach to micro-macro modeling of heterogeneous materials. *Comput. Mech.* 27, 37-48, 2001.

[19] M. G. D. Geers, V. G. Kouznetsova and W. A. M. Brekelmans. Multi-scale first-order and second-order computational homogenisation of microstructures towards continua. *Int. J. Multiscale Comput. Eng.* 1, 371-386, 2003.

[20] H. Berger, U. Gabbert, H Koeppe, R. Rodriguez-Ramos, J. Bravo-Castillero, G. R. Diaz, J. A. Otero and G. A. Maugin. Finite element and asymptotic homogenization methods applied to smart composite materials. *Comput. Mech.* 33 61-7, 2003.

[21] H. Berger, S. Kari , U. Gabbert, R. Rodriguez-Ramos, R. Guinovart-Diaz, J. A. Otero and J. Bravo-Castillero. An analytical and numerical approach for calculating effective material coefficients of piezoelectric fiber composites. *Int. J. Solids Struct.* 42 5692-714, 2004.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 3

Third party damages of offshore pipeline
海底管道的第三方损伤

Nurul Sa'aadah Sulaiman*, Henry Tan （谭鸿来）

LRF Centre for Safety and Reliability Engineering, School of Engineering, University of Aberdeen, Aberdeen AB24 3UE, UK

r01nsas@abdn.ac.uk, h.tan@abdn.ac.uk

Accepted for publication on 26th July 2014

Abstract - **Risk assessment is established to assist authorities in determining the priority of maintenance using risk which integrates both safety and failure. An efficient pipeline risk assessment should be able to characterize and calculate the risk associated with the pipeline. Unfortunately, the calculation of risk requires knowledge about the probability of failure and the consequence of failure. Both of which are difficult to estimate and in practical, the system under analysis cannot be characterized exactly. Numerical or objective data are often inadequate, highly uncertain and sometimes not available to perform calculations. To deal with this kind of situation effectively and consistently, a rigorous method of quantifying uncertainty using provided data is needed as well as to update existing information when new knowledge and data become available.**

In this paper, a probability analysis model of offshore pipeline failure due to third party damages is presented. The interaction between ship anchors, dropped objects and fishing gears are discussed. Bayesian networks model is proposed to determine the probability of third party damages to subsea pipelines. To generate the probabilities of different kind of nodes in a Bayesian network, a systematic probability approach is proposed with an emphasis on eliciting the conditional probability tables with multi-parents. The UK PARLOC database and DNV reports were used for the work. The paper concluded that Bayesian Network is a superior technique for risk analysis of pipeline failure. It is envisaged that the proposed approach could serve as a basis for decision making of pipeline maintenance.

Keywords – **Bayesian Network, Third party damage, Pipeline failure, Risk assessment**

I. INTRODUCTION

Oil and gas transmission pipelines are demonstrably safe and reliable means of transporting hydrocarbons. This is due to the combination of good design, materials and operating practices. The material used to design a pipeline is basically for them to operate under severe stress condition. However, diverse unavoidable factor may lead to pipeline failure as the limiting stress condition is achieved [1].

Review of historical subsea pipeline incident data and the literature summarizing pipeline failures suggest that failure causes can be grouped into four categories [2, 3]. Fig. 1 shows the distribution of offshore pipeline failure causes. It is illustrated by the historical data that the external interference (38%) is a significant contributor to potential pipelines failures followed by corrosion (36%). Most common factors of outside force or third party damage are from impact and offshore anchoring activities. The consequences of any pipeline failure may cause a serious impact to the life safety, environment and economy. In order to mitigate the severity of pipeline consequences, regular inspection and maintenance are crucial. However, maintenance activities can represent a significant cost as pipeline may cover large distances, often located in inaccessible areas and aging pipeline system incurred more maintenance frequency. Thus, risk assessment is established to assist authorities in determining the priority of maintenance using risk which integrates both safety and failure. There are several researches were conducted in developing framework to analyze risks connected to pipelines [4-7]. Nevertheless, the most significant drawback of existing frameworks for pipeline risk assessment is that they have not been performed from a causal perspective as a proper risk assessment requires a holistic outlook that embraces a causal view of interconnected event.

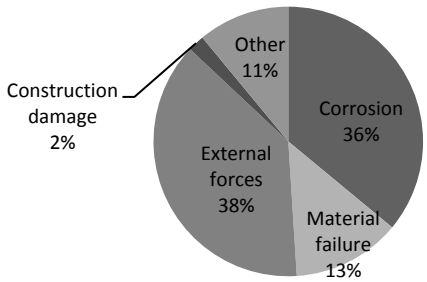


Fig.1, Offshore pipeline failure causes.

In spite that, engineering problems are subjected to significant uncertainty which is an inseparable part of real world systems. It is necessary to be able to represent, treat and manage the uncertainty in a consistent manner. Furthermore, in the decision making process, the uncertainty related to system assumptions is of tremendous importance. It is as

important to be well accounted for in the calculation of risks as the degree of uncertainties and their dependencies might influence the assessed risks [8]. Several approaches being developed to represent and express uncertainties in risk assessment for pipeline failure such as fuzzy logic [9] and Bayesian statistics [1]. Yet, in practical risk assessment, the most common approach to treat the uncertainty is by probabilistic approach, in their Bayesian formulation for the treatment of rare events and poorly known process typical of high-consequence technologies [10].

II. BAYESIAN NETWORK

Bayesian Networks (BN) is a probabilistic graphical model for a set of variables $A=\{A_1,...,A_n\}$, which consist of qualitative and quantitative components. Qualitatively, a BN is formed from the variables together with the directed edges or arcs. In the model, each node represents a variable (discrete or continuous) that can be in one of a finite state. Meanwhile, the arc linking two variables designate causal or influential relationships between them. The structure of the BN is explicitly represents the dependence and independence relationship among the set of variables.

As the structure is defined, the strength of relationships among the variables can be achieved from the joint probability distribution of all the variables numerically which represents a BN. The probability of independencies among the model can be described efficiently by this distribution. Each variable in the graph has an associated probability distribution (PD) conditional on its direct predecessors (parents) also known as conditional probability table (CPT).

Conditional independence implies due to absence of an arc between two nodes. However, by taking into account the distribution of its parents, the conditional probability of a node can be determined. By implementing the concept, it is thus possible to specify the joint probability of the entire network structure. The relationship can be calculated by applying the chain rule for Bayesian networks. A unique joint probability distribution of the entire network over all the variables is given by the product of conditional distributions attached to each node as in Eq. (1):

$$P(A_1, \dots, A_n) = \prod_{i=1}^{n} P(A_i | parents(A_i)) \quad (1)$$

The probability statement of model parameters is made up from some initial or prior belief (probability) about an event. Once new information on evidence/event is observed, the state of knowledge of the prior probability can be updated by calculating revised probabilities which also known as posterior probabilities. In general, this rule is a repeating process about an event A given information about event B every time new or additional evidence/information becomes available. The Eq. (2) is given as:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \quad (2)$$

The term P(A) is called prior or marginal probability of A. It is prior in the sense that it predicts any information about B and this is what causes all the arguments. P(A|B) is called the posterior probability of A given B as it is derived from the specified value of B. Whereas, P(B|A) is the conditional probability of B given A. P(B) is called prior or total probability of B and it is the one that provides evidence of interest for the probability update of A.

Bayesian networks can as well be transformed to a valuable decision model called influence diagrams which are highly intuitive in the decision making process [11]. It consists of belief networks with two additional node types, namely decision nodes and utility nodes. Decision nodes correspond to controllable decisions that have an effect on the system. Utility nodes on the other hands represent criteria for making choices and are used to assign values to particular outcomes.

Influence diagrams can be built and manipulated using program package such as *Hugin Expert*. The outcome of an influence diagram includes the marginal probability distribution of all variables in the domain and the expected utilities for the decision. If evidence is observed and propagated, expected utilities for the decision variable will be computed and updated. Thus, an influence diagram provides a dynamic decision tool presenting the optimal strategy, possibly conditional on a set of knowledge [12].

III. RISK ASSESSMENT USING BN MODEL

A risk based ranking of pipeline segment is valuable to assist authorities in determining the priority of inspection and maintenance of pipeline. The common method reported from the literature to quantify the risk is by applying the risk formula of;

Risk = probability of failure x *probability of consequences*

This method however, decomposing the risk into two components and so as the uncertainties. Moreover, the value can only be obtained with detailed analysis of the variables involved.

In this paper, the suggested risk assessment is by quantifying the risk in causal analysis. This approach treats a risk as an event that can be characterized by a causal chain involving (at least)[13] :

- The event
- At least one consequences event that characterises the impact
- One or more initiating event
- One or more control events which may stop the initiating event from causing the risk event
- One or more mitigating events which may help avoid the consequence event

With the causal perspectives approach suggested, a risk is not treated as a single event but by a set of event.

IV. THIRD PARTY DAMAGE

"Third-party damage" is known as damages due directly to acts of man and includes all activities not directly related to the pipeline of study. Three possible hazardous scenarios that may

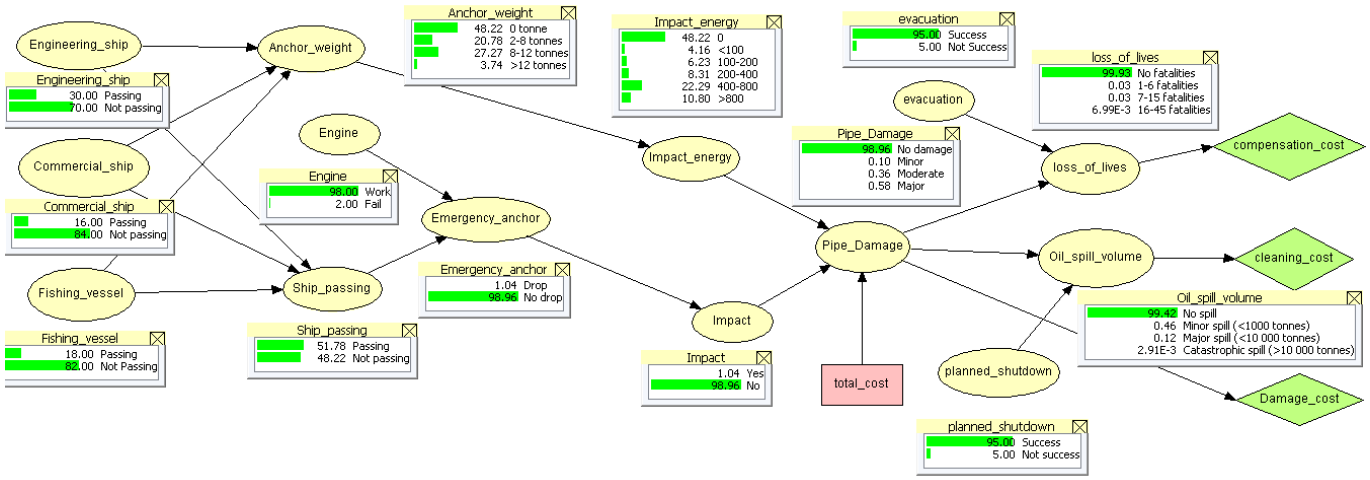


Fig 2, Proposed Bayesian network model for section 1.

threaten the pipeline are dropped object, anchoring and fishing activities.

Dropped object is the event where an object can be lost from vessels. The most frequent incident in the open sea is the lost of container from commercial ships crossing the pipeline route. However, only a small fraction of containers sunk due to heavy load or to a lack of water tightness while others will stay floated [14]. Thus the risk of this dropped object can be neglected.

Fishing activities on the other hand, threaten the pipeline which mainly due to bottom trawling. Load from the trawl gear is associated with prompt impact, subsequent overdraw or hook up. Yet, experience with numerous offshore pipelines in the North Sea show that fishery and offshore pipelines can co-exist safely as protected pipeline can withstand trawl gear interaction [15].

Despite that, historical data shows that 22/96 leaked incidents were caused by anchor impact [16]. Thus, only anchoring activities are considered in this study. The major cause leading to drop of anchor is unplanned or emergency anchoring as planned anchoring is made in authorized areas of pipeline absent.

The possible damage induced by an external impact is based on deformation level of the pipeline at the end of interaction process [14]. This deformation occurs depend on the impact energy towards the pipeline. Throughout this study, impact energy was used to define damage of pipeline as each pipeline has its own characteristics.

V. Anchor-Pipeline Interactions Scenario

There are three types of ships possibility will pass through the pipeline in open sea which are commercial ships, engineering ships and fishing vessel. The weight of anchor is depending on the type of ships and it then affects the probability of impact energy. Impact on pipeline happens when ship's engine crossing through a pipeline is not working. Subsequently, anchor is dropped under emergency condition. Damage to a subsea pipeline is integrated by impact energy and impact on the pipeline. The potential consequences due to pipeline damages are loss of lives, oil spillage, production loss and repair. The amount of oil spill also depends on the success of production shut-down operations; meanwhile, loss of lives depends on the success of operations to evacuate victims during the event.

The network is then expanded into an influence diagram by introducing decision node and utility functions which gather information for the potential benefits and enabling the 'oil estimation of expected total cost. Given the outcome state of spill', a value node of 'environmental cleaning cost' is created, given the outcome state 'loss of lives' a value node of 'compensation and loss of reputation' is introduced and given the outcome of 'pipe damage' a value node of 'damage cost' referring to production and cost to repair damage is created. All the value nodes are introduced to estimate the total expected cost as a function of the decision denoted as 'total cost'. The scenario is then transformed into BN model as in Fig. 2.

VI. Application Example: Nord Stream Pipeline

In this section, the utilization of developed BN model framework is discussed based on example of Nord Stream pipeline interaction with ships along the Baltic Sea involving five countries; Russia, Sweden, Germany, Finland and Denmark [17]. The properties of the pipeline sections with high density ship traffic were defined as in Table 1.

TABLE 1, Total expected cost for each section

Pipeline section	Section length (km)	No of ship passing (per year)		
		Fishing ship	Commercial ship	Engineering ship
1	10	400	805	775
2	15	1000	1960	860
3	10	790	2180	920

The proposed model shown in Fig. 2 is used and the methods for CPTs elicitations are discussed. For two states events, conversion of Boolean operations (disjunction, v and conjunction, ∧) into CPTs is applied. Disjunction is applied for node ship passing and type of ships (*ship passing = fishing vessel v engineering ship v transport ship*) as in Fig. 3 indicating 1 is ship passing and 0 is ship not passing. Meanwhile, conjunction is applied for node ship passing, engine and emergency anchor (emergency anchor = ship passing ∧ engine) as shown in Fig. 4 with 1 is anchor drop and 0 for no drop.

Edit Functions View

Ship_passing

Engineering...	Passing				Not passing			
Fishing_vessel	Passing		Not Passing		Passing		Not Passing	
Commercial...	Passing	Not passing	Passing	Not passing	Passing	Not passing	Passing	Not passing
Passing	1	1	1	1	1	1	1	0
Not passing	0	0	0	0	0	0	0	1

Fig.3, Disjunction CPTs of fishing vessel, engineering ship and transport ship.

Edit Functions View

Emergency_anchor

Ship_passing	Passing		Not passing	
Engine	Work	Fail	Work	Fail
Drop	0	1	0	0
No drop	1	0	1	1

Fig.4, Conjunction CPTs for emergency anchor.

The CPTs for impact energy and pipe damage are elicited according to historical data from recommended practice for pipeline protection, DNV-RP-F107 presented by [18]. Fig. 5 shows the probability of impact energy given anchor weight based on the historical data proposed for a pipeline with normal protection requirement. The damages to the pipeline are classified and defined according to [18] report as follows:

i. No damage.
ii. Minor damage: no repair required and no hydrocarbons release.
iii. Moderate damage: repair required but not leads to hydrocarbons release.
iv. Major damage: damage leading to release of hydrocarbons. Immediate stop of pipeline operation and to be repaired if the pipe in ruptured. Damage section need to removed and replaced.

Edit Functions View

Impact_energy

Anchor_weight	0 tonne	2-8 tonnes	8-12 tonnes	>12 tonnes
0 kJ	1	0	0	0
<100 kJ	0	0.2	0	0
100-200 kJ	0	0.3	0	0
200-400 kJ	0	0.4	0	0
400-800 kJ	0	0.1	0.7	0.3
>800 kJ	0	0	0.3	0.7

Fig.5, Probability of impact energy given anchor weight.

From the state of pipeline, consequences due to the failure are assessed. In pipeline mid-line zone, releases of hydrocarbon may endanger third party personnel. In this case, only major release scenarios (rupture) may threat the personnel. The personnel here is refers to the crew and passengers of vessels operating in the vicinity of a pipeline. Ignition will only occur if the gas above the sea surface is flammable concentration and possible ignition sources are present within this cloud. The outcome of the event is often difficult to predict accurately. It may be assumed in major release events, 1-10% of these events the gas release will ignite [18].

The number of fatalities is estimated based on quantification of societal risk resulted from the consequences analysis of the third party activities (dropped object, dropped anchor, dragged anchor and fishing) reported for subsea gas pipeline [15]. The societal risk result was plotted on F-N curves and it is used to develop the CPT for node 'loss of lives'. It is demonstrated from the report that the level of risk is broadly acceptable when compared with agreed risk tolerability criteria.

In determining the cost valuation of human life, several approaches have been proposed. These include Willingness to pay (WTP), Value of statistical life (VSL), the CSX-value and Human capital approach [19]. For this example, the valuation of loss of life is estimated from work done by [20]. They suggested that for developed countries, the average value of a statistical life according to Life Quality Index and macro-economic valuation is in between £ 0.8 M – £3.3 M. The quantitative inputs for compensation cost utility node is £11.2 Million, £27.9 Million and £83.7 Million for 1-6 fatalities, 7-15 fatalities and 16-45 fatalities respectively.

To simplify the model, damage to environment is measured only in terms of oil spill volumes. Both minor and major release scenarios may give an impact to the environment. The pollution from the spillage affects the eco-system in the water, shoreline environment, sea birds and fishes. The environmental consequences are usually expressed as clean – up efficiency or estimated time for achievement of full recovery of the affected area. The estimation of cleanup costs for oil spill is according to average cleanup cost per tonne spilled based on analysis of oil spill cost data from the OSIR International Oil Spill data [21]. From the report, the average cleanup cost is £ 5281.38 per tonne for Europe region. The cleanup-cost is £41.2 Million for minor spill, £47.5 Million for major spillage and £80 Million in case of catastrophic spill.

Lastly, the damage cost which is the summation of production loss and pipeline repair cost is evaluated. The assumption value is found from the evaluation done by [22]. Both costs have been assumed to be linearly related to time taken for repair. In this assumption, material cost for repairs have been neglected. The estimated costs are as follows:

- Loss of production £ 2 M per day
- Repair cost of £ 0.1 M per day

It is assumed to repair a moderate damage is up to 16 days (clamp repair) and for large damage (newspool piece installed using mechanical connectors) 30 days is required. Addition of 3 days vessel mobilization has been assumed for every case. The damage cost for moderate and major event is £39.9 Million and £69.3 Million, respectively.

The BN model is analyzed and the total cost incurred for each section is shown in Table 2. From the data used in this example, the major contributor to the expected cost of failure is product losses and repair cost with 90%, 86% and 85% for

pipe section 1, 2 and 3 respectively. The contribution of other cost to the total of expected failure cost is relatively low. Result from Table 2 is utilized in order to compare and rank the risk of different pipeline section. Section 3 has the highest cost incurred thus this section has the highest priority for maintenance operations followed by Section 2 and Section 1.

In addition, a number of entered evidence situation for this model can be investigated and the total expected cost will be recalculated. As an example, in worst case scenario where pipeline is 100% in major damage, the total expected cost incurred is £83.78 provided the planned shutdown is 95% success. From the analysis, the main contributors for the major damage event to occur are 'engine fail' and 'engineering ship passes'.

TABLE 2, TOTAL EXPECTED COST FOR EACH SECTION

Pipe Section	Total cost incurred (£ Million)
1	0.63
2	0.86
3	0.90

VII. CONCLUSION

This paper has demonstrated the application of Bayesian network modelling for risk assessment of subsea pipeline due to anchor interaction. Firstly, BN model is developed for pipeline-anchor interaction. The network then is expanded into an influence diagram by introducing decision node which is the 'total cost' and utility functions ('compensation cost', 'cleaning cost' and 'damage cost') which gather information for the potential benefits and enabling the estimation of expected total cost. This is then followed by the elicitation of CPTs. Generally, there are many ways to establish the CPTs. To evaluate the proposed model, a case study from Nord Stream is adopted and most of the CPTs are elicited from statistical and historical data. Probability prediction and evidence propagation were conducted to analyze damage probability and to highlight problematic area respectively. The proposed model illustrate that risk analyst are able to figure out the causes of pipeline damage and the consequences such as loss of live, oil spill volume and repair cost and production loss. In addition, the developed influence diagram offer a tool to support risk ranking and risk reduction measures based on expected failure cost incurred for a given segment of a pipeline.

REFERENCES

[1] T. Breton, J. C. Sanchez-Gheno, J. L. Alamilla and J. Alvarez-Ramirez, "Identification of failure type in corroded pipelines: A Bayesian probabilistic approach," *J. Hazard. Mater.*, vol. 179, pp. 628-634, 2010.
[2] P. K. Dey and S. S. Gupta, "Risk-based model aids selection of pipeline inspection, maintenance strategies," *Oil Gas J.*, vol. 99, pp. 54-60, 2001.
[3] D. Yuhua and Y. Datao, "Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis," *J Loss Prev Process Ind*, vol. 18, pp. 83-88, 2005.
[4] M. Dziubiński, M. Fratczak and A. S. Markowski, "Aspects of risk analysis associated with major failures of fuel pipelines," *J Loss Prev Process Ind*, vol. 19, pp. 399-408, 2006.
[5] Z. Y. Han and W. G. Weng, "An integrated quantitative risk analysis method for natural gas pipeline network," *J Loss Prev Process Ind*, vol. 23, pp. 428-436, 5, 2010.
[6] A. J. Brito, A. T. de Almeida and C. M. M. Mota, "A multicriteria model for risk sorting of natural gas pipelines based on ELECTRE TRI integrating Utility Theory," *Eur. J. Oper. Res.*, vol. 200, pp. 812-821, 2010.
[7] Y. Jo and B. J. Ahn, "A method of quantitative risk assessment for transmission pipeline carrying natural gas," *J. Hazard. Mater.*, vol. 123, pp. 1-12, 8/31, 2005.
[8] JCSS, "Risk assessment in engineering, principles, system representation & risk criteria." 2008.
[9] M. Singh and T. Markeset, "A methodology for risk-based inspection planning of oil and gas pipes based on fuzzy logic framework," *Eng. Failure Anal.*, vol. 16, pp. 2098-2113, OCT, 2009.
[10] T. Aven and E. Zio, "Some considerations on the treatment of uncertainties in risk assessment for practical decision making," *Reliability Engineering and System Safety*, vol. 96, pp. 64-74, 2011.
[11] A. G. Eleye-Datubo, A. Wall, A. Saajedi and J. Wang, "Enabling a powerful marine and offshore decision-support solution through bayesian network technique," *Risk Analysis*, vol. 26, pp. 695-721, 2006.
[12] A. Friis-Hansen, "Bayesian Network as a Decision Support Tool in Marine Applications," 2000.
[13] N. Fenton, M. Neil and D. A. Lagnado, "A General Structure for Legal Arguments About Evidence Using Bayesian Networks," *Cognitive Science*, vol. 37, pp. 61-102, 2013.
[14] R. Carpaneto and M. Tominez, "Risk analysis of subsea pipelines hazards induced by human offshore activities," in *Proc. First International Offshore and Polar Engineering Conf.*, Edinburgh, UK. 1991, pp. 372-377.
[15] Nord Stream. (2009, Feb.). Nord Stream environmental impact assessment documentation for consultation under the Espoo Convention (Nord stream Espoo report). Nord Stream AG, Zug, Switzerland. [Online]. Available: https://www.nord-stream.com
[16] Mott MacDonald Ltd. (2003, Jun). PARLOC 2001: The update of loss of containment data for offshore pipelines. HSE UK, Croydon, UK. [Online]. Available: http://www.hse.gov.uk/ parloc-2001-report.pdf
[17] L. Vitali, F. Candiracci, C. Crea, R. Bruschi and W. Rott, "Nord stream project - pipeline safety against ship traffic related threats: Quantitative risk assessment approach," in *Proc. 22 International Offshore and Polar Engineering Conf.*, Rhode, Greece.2012, pp. 75-82.
[18] Det Noske Veritas, "Risk assessment of pipeline protection", Norway. Recommended practice DNV-RP-F107, Oct. 2009.

[19] A. W. Dawotola, "Risk based maintenance of petroleum pipelines," Ph.D dissertation, Delft University of Technology, Netherland, 2012.
[20] R. B. Jongejan, S. N. Jonkman and J. K. Vrijling, "An overview of methods for the economic valuation of human life," in 2005, pp. 973-980.
[21] D. S. Etkin, "Estimating cleanup costs for oil spills," in *Int. Oil Spill Conf.*, 2005, pp. 2625-2634.
[22] Y. Bai and Q. Bai, *Subsea pipelines and risers*. Amsterdam: Elsevier, 2005, pp. 798

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 4

Electrical recovery and fatigue degradation phenomena in cracked silicon cells

断裂硅电池中的电力恢复和疲劳退化现象

Marco Paggi[1*], Irene Berardone[2], Andrea Infuso[2] and Mauro Corrado[2]

[1]*Research Unit on Multi-scale Analysis of Materials, IMT Institute for Advanced Studies Lucca, Lucca, Italy*

[2]*Department of Structural, Geotechnical and Building Engineering, Politecnico di Torino, Torino, Italy*

marco.paggi@imtlucca.it

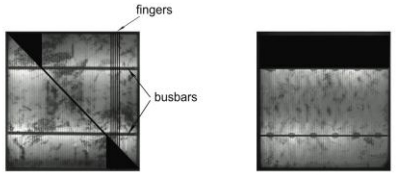
Accepted for publication on 26th July 2014

Abstract - **An experimental study based on the electroluminescence technique is herein proposed to demonstrate the existence of a coupling between mechanical deformations and the intensity of the electric field due to cracks in monocrystalline Silicon cells embedded in photovoltaic modules. In spite of the very brittle nature of Silicon, due to the action of the encapsulating polymer and residual compressive stresses resulting from the lamination stage, cracks experience crack closure and contact during mechanical unloading, partially recovering their original electric response. Crack propagation in case of cyclic loading, as, e.g., in the case of vibrations due to transportation and use, have also been reported for the very first time. The research results pinpoint the need of improving electric predictions based on the estimation of inactive cell areas since worst case scenarios not accounting for electro-mechanical coupling are too conservative.**

Keywords – **Photovoltaics, Cracks, Fatigue, Electromechanical coupling**

I. INTRODUCTION

Photovoltaic (PV) modules are supposed to have a lifetime longer than 20 years under the exposure to environmental conditions like temperature variations, wind vibration and snow pressure. Thermo-mechanical loads induce stresses into the components of the module, especially into the crystalline Silicon (Si) solar cells, which are affected by cracking [1-9]. Cracks on the millimetre or centimetre size are mostly invisible by naked eye but they can be localized according to the electroluminescence (EL) technique [10]. Such cracks can lead to electrically inactive cell areas, thus reducing the power output of the module. For instance, cracks inserted in solar cells by the application of a uniform pressure to simulate snow can lead up to 1.5% of power loss [3]. After the subsequent application of 200 humidity freeze cycles according to standard specifications [11], such cracks propagate, the electrically disconnected areas increase in size and up to 10% of power loss has been reported [3]. Potentially, if a crack crossing conductor (called finger) is sufficiently open, then the finger may fail and the electric flow to the busbar can be interrupted. Therefore, portions of Si cells can be potentially deactivated by cracks and their impact on power-loss reasonably depends on their inclination and position with respect to the busbars, see Fig. 1a,b. For instance, a crack parallel to the busbar on the upper side of the cell could lead up to 25% of electrically inactive area (Fig. 1b). According to this pure geometrical criterion which does not take into account neither physical mechanisms such as thermo-mechanical deformation, nor the fact that the cells are embedded in the composite PV module, worst-case scenarios have been predicted by considering all the experimentally detected [3] or the numerically simulated [8] cracks as perfectly insulated lines.



(a) 6% of potentially inactive cell (b) 25% of potentially inactive cell

Fig. 1, Amount of potentially inactive cell areas depending on the orientation of the crack with respect to the busbars [3,8] (a,b).

In reality, it is reasonable to expect intermediate configurations where cracks may partially conduct depending on the relative crack opening displacement at crack faces. In analogy with the cohesive zone model used in computational fracture mechanics to depict crack growth in crystalline materials [12,13], where cohesive tractions opposing to the relative displacement of the crack faces are decreasing functions of the opening and sliding displacements, an electric flux dependent on crack opening might be postulated. A very

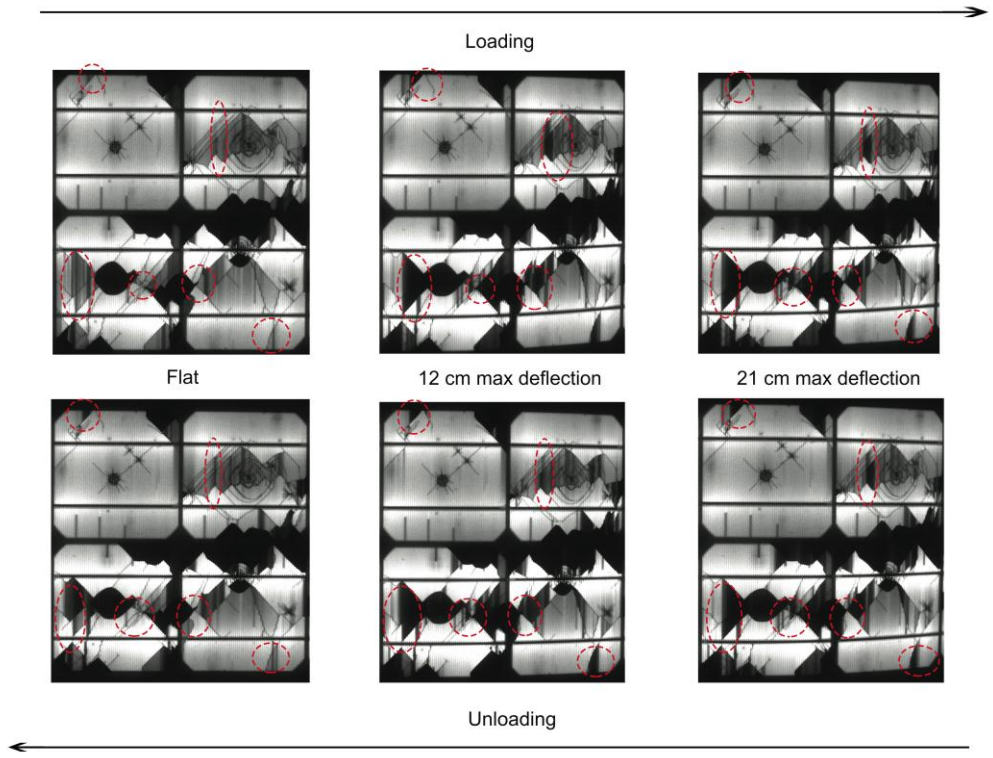


Fig. 2, Evolution of Electroluminescence (EL) signal during bending loading and unloading for different deflections.

preliminary experimental observation supporting coupling effects due to cracking regards the highly oscillating electrical response in time of a PV string containing a cracked cell, depending on the cell temperature [7].

II. Experimental tests

The evidence of coupling between thermoelastic and electric fields has important consequences from the modelling point of view. At present, electric models of solar cells do not consider this form of coupling induced by cracking.

In the most refined versions [14,15], a discretization of the solar cell is made in the plane and a two-diode model is applied to each node of the mesh to predict the electric response of the semiconductor. In case of hot spots, however, it was indeed necessary to modify the value of the series resistance in the nodes close to a crack. Hence, to achieve a predictive stage useful for power-loss predictions and durability assessment of PV systems exposed to environmental loads, series resistance values used as input of the circuit model should be related to the thermoelastic stress state in the solar cell.

To make an insight into this phenomenon, a three-point bending test with a mid-span displacement control on a rectangular PV module has been performed by monitoring cracking at different deformation levels by taking EL images. A semi-flexible module made of 2 columns of 5 monocrystalline Si cells (156x156 mm wide) is used. This type of module, which has a certain degree of flexibility, can be used in many applications where the substrate we would like to bond the module is curved. The partially symmetric arrangement of the layers through the thickness (0.265 mm of polyethylene terephthalate, 0.600 mm of epoxy-vinyl-acetate, 0.166 mm of Silicon, 0.400 mm of epoxy-vinyl-acetate and 0.345 mm of backsheet) leads to Silicon cells just above the neutral axis of the cross-section. Therefore, although large transversal displacements are imposed to the thin plate, tensile stresses inside Silicon are moderate and comparable with those experienced in classical PV modules under conventional bending tests. In those composites, having a thick glass cover (4 mm) and an unsymmetric structure, Si cells are much more distant from the neutral axis and therefore are subjected to a tensile stress state comparable with that simulated in the present test, in spite of the much lower inflections induced by environmental loads [4,8].

To create pre-existing cracks and study their evolution depending on the imposed flexure, moderate impacts have been made with PMMA balls of 4 cm of diameter at a velocity of 6 m/s. In this way, cracks are introduced by an indentation effect. The EL image of four solar cells in the middle of the panel and in the initial flat reference configuration is shown in Fig. 2. Location of impacts can be clearly distinguished by the circular dark spots from where diagonal cracks depart. These patterns are completely invisible by naked eye and can only be determined by the EL technique. Most of them are inclined at

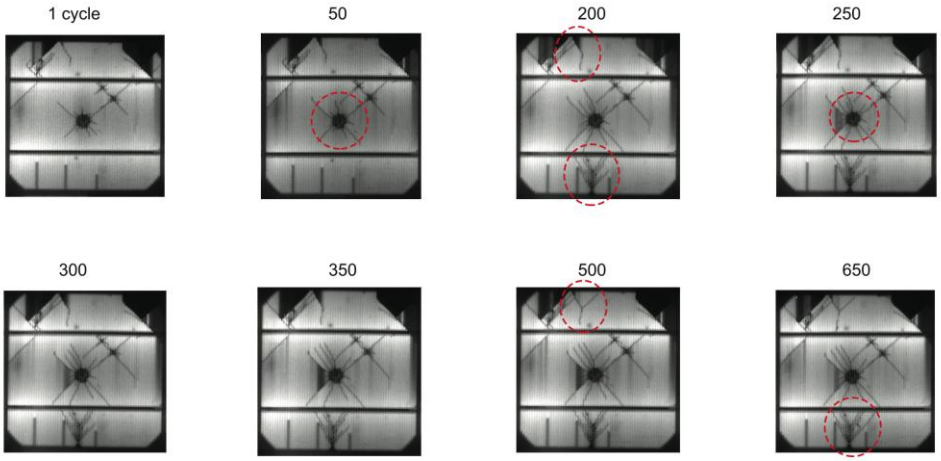


Fig. 3, Fatigue propagation of cracks, EL images at the maximum deflection point of the module (21 cm), for different loading cycles.

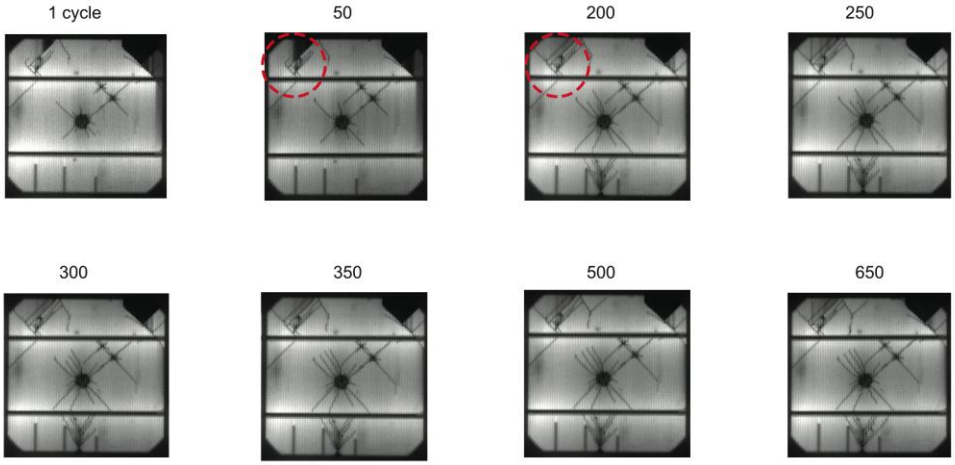


Fig. 4, Fatigue propagation of cracks, EL images of the unloaded configurations (flat, no deflection), for different loading cycles.

±45°, consistently with the 111 crystallographic plane of cubic face centred monocrystalline Silicon. In some cases, large electrically insulated cell portions, appearing as dark areas, are present. Other portions, circled with dashed lines, show regions isolated by cracks but still partially conducting from the electrical point of view. As expected, this can be explained by the fact that their crack opening is smaller than a critical value for complete electric insulation.

Due to a thermo-elastic mismatch between Silicon and the thermoplastic materials composing the module, a compression state in the cell is present in the unloaded configuration. This is the result of cooling down the module from the stress-free condition at the lamination temperature of 150 °C to the ambient temperature [16]. Finite element computations estimate a shortening of each cell of about 0.1 mm. Therefore, a tensile axial displacement of that amount is necessary in bending to open initially compressed cracks.

By increasing the bending deflection (upper row in Fig. 2, loading direction from left to right), cracks open and additional insulated dark areas appear, thus confirming that the electrical behaviour of cracks does depend on the elastic deformation. Very interesting is the unloading stage (lower row in Fig. 2, unloading direction from right to left), where cracks come into contact and conduct again. Hence, although globally the amount of damage is increased during the cycle since the final EL image is darker than before, an electric recovery phenomenon due to crack closure is reported.

III. Fatigue degradation and Electromechanical coupling

The evolution of degradation by increasing the amount of bending cycles (from zero to a maximum deflection of the module of 21 cm per cycle) has also been monitored and it is shown in Fig. 3, where EL images of a significant cell of the module have been recorded at the point of maximum

deflection for 1, 50, 200, 250, 300, 350, 500 and 650 cycles. During the first 50 cycles we report a propagation of cracks around the dark central point of impact. Crack propagation further takes place from 50 to 200 cycles, with the appearance of new cracks from the lower border of the cell, originating from an edge crack. No crack propagation takes place from 200 to 250 cycles, although the EL intensity between diagonal cracks diminishes, thus implying a material degradation and a reduction of electric conductivity across the cracks. This trend continues further by increasing the number of cycles, with also further crack propagation (cycle 300) and coalescence of two propagating cracks (cycle 350). At 500 cycles, a new crack is originated from an initial one on the upper side of the cell as a result of a crack-branching mechanism. After the appearance of this new crack, the dark EL area just on the left of it becomes conductive again. This can be explained as the effect of the phenomenon of strain localization, popular in brittle materials, which corresponds to a stress relief in the surrounding cracks when a major crack propagates [12]. Therefore, with the appearance of a new crack, the previous one in the surrounding is relaxed and it experiences a reduction of its relative opening displacement.

The analysis of the EL images in the undeformed stage at the end of each set of cycles (Fig. 4) shows again the phenomenon of electrical recovery due to crack closure already noticed in the case of the first cycle. In these images, strain-localization can be noted by comparing EL images for 50 and 200 cycles, see the circled areas. The propagation of a diagonal crack to the left relieves the deformation in the Silicon area, the diagonal crack on the right experiences a reduction in crack opening and the EL image of Si above it becomes brighter.

A quantitative analysis of EL images can be performed by associating a value ranging from 0 to 1 to the grey-scale intensity of each pixel of the digital photo. Values of EL close to unity correspond to very bright Silicon areas, whereas values close to the opposite extreme are dark "electrically inactive" zones. Hence we can determine for each EL image the frequency histograms by computing the percentage of pixels with a given EL intensity with respect to the total number of pixels. At maximum deflection (Fig. 3), the histogram is particularly skew-symmetric, with a long tail for EL intensities less than 0.5 (Fig. 5). Moreover, although the difference between EL images by increasing the number of cycles is apparently marginal, i.e., only a few cracks propagate and some regions become darker, the EL histograms quantify the fatigue degradation trend. By increasing the number of cycles, in fact, the EL histogram shifts to the left. For high EL intensities, the pixel percentage is diminished by increasing the number of cycles. Exactly the opposite trend takes place in the low EL intensity range, i.e., more pixels become darker (Fig. 5).

IIII. Discussion and Conclusions

The experimental results herein presented show that, due to the encapsulation of Si cells into a ductile polymeric material, coupling between the electric response measured by EL images and the elastic deformation field takes place due to cracking. Therefore, the worst case scenario of considering all the detected cracks as perfectly insulated ones has to be considered with care, since more complex degradation phenomena occur in reality. In particular, we have shown the appearance of electrical recovery due to crack closure and residual thermo-elastic stresses, fatigue crack propagation, crack branching and crack coalescence, stress relief and recovery in the EL signal in case of strain localization induced by crack propagation of existing cracks or by new crack nucleation. It has to be remarked that all of these phenomena are impossible in stand-alone Si cells which are very brittle and display a different behaviour as compared to Silicon at the micro-scale, like that used for MEMS [17]. Therefore, modelling of Si solar cells behaviour by neglecting their embedding into a composite is meaningless in case of PV applications. Of course, this introduces additional complexity due to the very different size-scales and thicknesses of the various PV layers which require the use of multi-scale computational strategies.

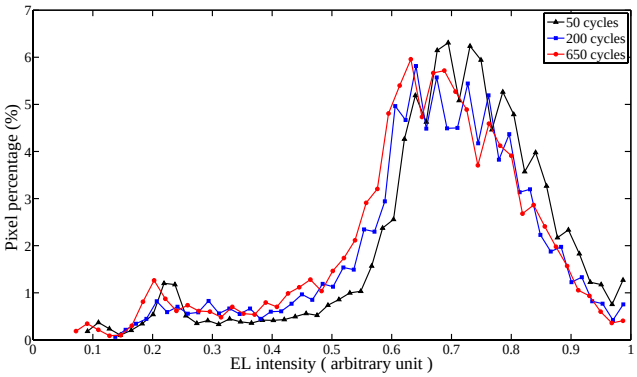


Fig. 5, EL frequency histograms for different number of bending cycles at maximum deflection. Note the shift of the histograms to the left by increasing the number of cycles, a trend consistent with the progression of fatigue damage.

The reported results have general implications for any kind of conductive material embedded into a ductile polymeric one, see, e.g., graphene layers dispersed into a polymeric matrix [18]. In other words, the global behaviour cannot be predicted from the mechanical properties of the individual constituents, but it comes out as the result of the interaction of the system's components, a feature typical of complex systems.

The present results pinpoint the need of modelling thermo-electro-mechanical coupling effects in polycrystalline Silicon solar cells embedded in PV modules, requiring the development of multidisciplinary research involving electronics, materials science and computational mechanics. From the industrial and engineering point of views, the present results imply that the quasi-static application of very high distributed pressures as requested by qualification standards [11] is not enough if we are interested in evaluating the actual degradation rate and possibly infer about the lifetime of produced PV modules. In fact, other unexpected forms of damage, like fatigue crack propagation, are indeed possible due to the composite structure of the module and can be

induced by very common sub-critical loads like vibrations due to transportation or wind gusts, phenomena not yet characterized so far.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013)/ ERC Grant Agreement No. 306622 (ERC Starting Grant ''Multi-field and multi-scale Computational Approach to Design and Durability of PhotoVoltaic Modules''– CA2PVM).

REFERENCES

[1] S. Kajari-Schröder, I. Kunze, U. Eitner, and M. Köntges. "Spatial and orientational distribution of cracks in crystalline photovoltaic modules generated by mechanical load tests". *Sol. Energy Mater. Sol. Cells* **95,** pp. 3054-3059, 2011.

[2] S. Kajari-Schröder, I. Kunze, and M. Köntges, "Criticality of cracks in PV modules". *Energy Procedia* **27,** pp. 658-663, 2012.

[3] M. Köntges, I. Kunze, S. Kajari-Schröder, X. Breitenmoser and B. Bjrneklett. "The risk of power loss in crystalline silicon based photovoltaic modules due to microcracks". *Sol. Energy Mater. Sol. Cells* **95,** pp. 1131-1137, 2011.

[4] M. Sander, S. Dietrich, M. Pander, M. Ebert and J. Bagdahn. "Systematic investigation of cracks in encapsulated solar cells after mechanical loading". *Sol. Energy Mater. Sol. Cells* **111,** pp. 82-89, 2013.

[5] R. Khatri, S. Agarwal, I. Saha, S.K. Singh and B. Kumar. "Study on long term reliability of photo-voltaic modules and analysis of power degradation using accelerated aging tests and electroluminescence technique". *Energy Procedia* **8,** pp. 396-401, 2011.

[6] M.A. Munoz, M.C. Alonso-Garcia, N. Vela and F. Chenlo. "Early degradation of silicon PV modules and guaranty conditions". *Sol. Energy* **85,** pp. 2264-2274, 2011.

[7] B. Weinreich, B. Schauer, M. Zehner and G. Becker. "Validierung der Vermessung gebrochener Zellen im Feld mittels Leistungs PV-Thermographie". *Proc. 27 Symposium Photovoltaische Solarenergie*, pp. 190-196, 2012.

[8] M. Paggi, M. Corrado and M.A. Rodriguez. "A multi-physics and multi-scale numerical approach to microcracking and power-loss in photovoltaic modules". *Compos. Struct.* **95,** pp. 630-638, 2013.

[9] M. Paggi and A. Sapora. "Numerical modelling of microcracking in PV modules induced by thermo-mechanical loads". *Energy Procedia* **38,** pp. 506-515, 2013.

[10] T. Fuyuki, H. Kondo, T. Yamazaki, Y. Takahashi and Y. Uraok. "Photographic surveying of minority carrier diffusion length in polycrystalline silicon solar cells by electroluminescence". *Appl. Phys. Lett.* **86,** pp. 1–3, 2005.

[11] International Standard IEC 61215:2006. *Crystalline Silicon Terrestrial Photovoltaic (PV) Modules – Design Qualification And Type Approval*, 2006.

[12] M. Paggi, and P. Wriggers. "A nonlocal cohesive zone model for finite thickness interfaces – Part II: FE implementation and application to polycrystalline materials". *Comput. Mater. Sci.* **50,** pp. 1634-1643, 2011.

[13] M. Paggi, and P. Wriggers. "Stiffness and strength of hierarchical polycrystalline materials with imperfect interfaces". *J. Mech. Phys. Solids* **60,** pp. 557-571 , 2012.

[14] D. Grote. "Analyses of silicon solar cells and their measurement methods by distributed, circuit simulations and by experiment". *PhD thesis*, University of Konstanz, Germany, 2010.

[15] G. Siyu, M. Fa-Jun, H. Bram, G.A. Armin and P. Marius. "Analysing Solar Cells by Circuit Modelling". *Energy Procedia* **25**, pp. 28 – 33, 2012.

[16] M. Paggi, S. Kajari-Schröder and U. Eitner. "Thermomechanical deformations in photovoltaic laminates". *J. Strain Anal. Eng. Des.* **46,** pp. 772-782, 2011.

[17] C.L. Mulhstein, E.A. Stach and R.O. Ritchie. "A reaction-layer mechanism for the delayed failure of micron-scale polycrystalline silicon structural films subjected to high-cycle fatigue loading". *Acta Materialia* **50,** pp. 3579-3595, 2002.

[18] K.M.F. Shahil, and A.A. Balandin. "Graphene–multilayer graphene nanocomposites as highly efficient thermal interface materials". *Nano Lett.* **12,** pp. 861–867, 2012.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 5

Invariant Classification of Rocks by Rock Drilling Resistance (Procedure)

由岩石钻进过程阻力得来的岩石不变量分类

Tanaino A.S., Usoltseva O.M.

N.A. Chinakal Institute of Mining, Siberian Branch, Russian Academy of Sciences, Krasnyi pr. 54, Novosibirsk, 630091 Russia

tanaino@misd.nsc.ru

Accepted for publication on 2nd August 2014

Abstract - **A great drawback of the existing classifications of rocks by the rock drilling resistance is their dependence on drilling equipment characteristics. The proposed method of rock classification makes no use of the data on rates (rpm) or time (min/m) of drilling by a particular machine or a tool but rests on the information on the structure of a rock mass and physico-mechanical properties of types of rocks the rock mass is composed of. A drilled object, which is a rock mass, is identified by the dimensionless set of the properties and abrasiveness of rocks.**

The rock mass properties B are represented by the functional $B = f\{\varepsilon_1(\sigma),\ \varepsilon_2(\Delta),\ \varepsilon_3(w)\}$ **, where** σ **is the uniaxial compression limit;** Δ **is the characteristic of dimension of structural blocks in the rock mass; is the characteristic of the effect exerted by hydrophysical state of the rock mass on its strength; are dimensionality operators for rock properties, considering their influence on the rock drilling resistance. The functional B is constructed for every layer of rock types composing the rock mass and its weighted mean value is calculated afterwards.**

Potential abrasiveness of lithology layers in the rock mass is represented as the functional $A = \varphi\{\vartheta_1(d),\ \vartheta_2(\Phi),\ \vartheta_3(Sm),\ \vartheta_4(P), \vartheta_5(Sc),\ \vartheta_6(r)\}$ **, where** d and Φ **are the weighted mean dimension of a rock grain and the coefficient of its shape (roundness), respectively;** ***Sm*** **and** ***Sc*** **are the weighted mean hardness of minerals composing a rock, and the hardness of cement (detrital rocks) and glass (magmatic rocks);** ***P*** **is the porosity, %;** ***r*** **is the tensile strength;** ϑ_i **are dimensionality operators for rock properties, considering their influence on the rock abrasiveness in drilling.**

Using the geometrical scale with the common ratio $\sqrt{2}$**, in the real value range of the functional** ***B,*** **we have obtained the rock mass classification by the rock mass strength characteristics, and in the real value range of the functional** ***A*****—the rock mass classification by its abrasiveness. Cumulative values of strength and abrasiveness per classes of rocks are the dimensionless characteristic of the rock drilling resistance. The classification is independent of the drilling equipment performances and is invariant on that score.**

The authors compare the invariant classification of rocks by rock drilling resistance with the rock classifications by drilling rate that are in actual use in mining.

Keywords – **rocks, classifications by drilling, rock resistance, hardness, abrasiveness, physico-mechanical properties**

I. INTRODUCTION

Available classifications of rock drillability are greatly governed by drilling equipment characteristics. Understanding of this disadvantage conditions the search for new approaches to new classifications to be independent of the impact of technological tools. This paper focuses on **the development of a rock drillability classification based on the physico-mechanical properties of rocks.**

It follows from numerous experimental data generalization that the drilling tool penetration resistance of rocks depends on: (1) rock strength; (2) rock mass structure (jointing); and (3) rock abrasivity. The properties (1) and (2) describe the state of a rock mass. The strength of a rock (as a solid) is governed by its genesis, mineralogy, structure etc. Special influence is exerted on the rock mass strength by the hydrophysical environment in the rock mass.

The rock drillability evaluation meets the challenge of aggregate estimate of a set composed of properties stated in various units of measurement. As a way-out alternative, the present authors suggest dimensionless presentation of numerous mechanical properties on a canonical scale for their aggregate estimate later on. The basis for this suggestion is the

uniformity of laws concerned with hierarchy of rock blocks in the structure of rock masses and the clustering of the strength properties of rocks [1].

II. Rock mass strength assessment

2.1. Hydrophysical conditions and rock mass strength

Out of characteristics of hydrophysical environment, we will only analyze moisture. Theoretically, of interest is a relation of the type of $\sigma_w = f(W)$, where the rock mass moisture W ranges from a defined zero saturation (air-dried state) to overall saturation W_w ($0 \leqslant W \leqslant W_w$); σ_w is the limit strength of rocks at the moisture W. Weakening of rocks under influence of moisture is usually assessed using a coefficient of rock mass resistance to water, or water stability, K_w (introduced in [2]):

$$K_w = \sigma_w / \sigma_d;\ \sigma_w \leqslant \sigma_d,\ \ 0 < K_w \leqslant 1,$$

where σ_d is the rock limit strength in the air-dried state.

The authors suggest to account for the influence of hydrophysical conditions on the strength of hard rocks using the following relations:

$$\sigma_w \approx \sigma_d \cdot \exp(-W/P \cdot \lambda),\ K_w = \sigma_w/\sigma_d = \exp(-W/P \cdot \lambda), \qquad (1)$$

where P is the porosity of rocks, %; σ_w is the strength of rocks at the moisture W, MPa; λ is a coefficient of yet unclear nature.

The ratio W/P in (1) is a characteristic of occupation of pores with water ($W/P \leqslant 1$). Evidently, when W = 0 (ideal dry rock), $W/P \cdot \lambda = 0$ and, accordingly, $\exp(-W/P \cdot \lambda) = 1$, $\sigma_w = \sigma_d$, $K_w = 1$ for any P and λ. When rocks are in the condition of overall saturation with water, K_w is governed by the material constitution of rocks and by W/P. For this reason, (1) has the coefficient λ. Presumably λ is in a certain manner exposed to the Rebinder effect, which needs validation, so, the expressions (1) are hypothetical and require respective proof. Some of the deductions are obtained on the static level, by the experimental data on uniaxial strength of rocks in dry and water-saturated state.

It is supposed that λ in (1) can be presented as a functional connection $\lambda = f(x1, x2)$, where x1, x2 are basic properties of material constitution of rock and fluid, respectively. For lack of data, λ is estimated using statistical calculation [3] where for each kind of rocks the moisture W, porosity P and wet rock strength σ_w are known. From (1), for $\lambda = 1$, we calculate initial approximation for the wet rock strength σ_{wc}. The calculated σ_{wc} is compared with the experimentally obtained values of the wet rock strength σ_{we}. On condition that:

$$\delta = (1 - \sigma_{wc} / \sigma_{we}) < (0.1), \qquad (2)$$

it is assumed that λ is equal to 1. If the condition (2) is not fulfilled, we vary λ until (2) holds true and note the so-found value of λ. All in all, 80 kinds of rocks have been analyzed, with the found range of λ as follows: $0.1 < \lambda \leq 1.8$. It has been specified that $\lambda = 1$ for 36 kinds of rocks;

$\lambda < 1$—for 37 kinds of rocks; $\lambda > 1$—for 7 kinds of rocks.

According to the analysis of the calculation results, the values of K_w found from (1) are independent of the values of wet and dry rock strengths, σ_w and σ_d, respectively: K_w may have the same value for different σ_w and σ_d. From the calculation results:

$$K_w = \sigma_w / \sigma_d = 0.89 \approx 69/78 \approx 118/133 \approx 107/120 \approx 121/136 \approx 44/50 \approx \exp(-W/P \cdot \lambda).$$

Such equalities exist for different genesis rocks. Sticking to standardized terminology [2], the water stability coefficient K_w is of deeper meaning. Apparently, $K_w = \exp(-W_i/P_i \cdot \lambda_i)$ is not correlated with the strengths of rocks but relates with the ratio of strengths. Conceptually, K_w characterizes ability of rocks to "weaken" structural bonds at a microlevel under water saturation. The degree of "weakening" may differ in rocks of different material constitution. This phenomenon seems to be of interest and needs further consideration.

Based on the analysis and calculation, we conclude that the hypothetic premise formally presented in (2) is well applicable to predictive estimation of influence exerted by moisture and porosity on strength in rocks. It is found that for real rock masses (in particular, in Kuzbass), K_w is represented by the fourth order canonical scale:

$$K_w = K_{ref} \left(\sqrt[4]{2}\right)^{J-1},\ J = 1, 2, 3, \ldots, r, \qquad (3)$$

where $K_{ref} = 0.3$ is the minimum reference value of the water stability coefficient. Calculating J from (3) at $K_{ref} = 0.3$ and accounting for (1) yields:

$$J_h \approx 5.77 \cdot \mathrm{Ln}(Kw) + 7.95 = 5.77 \cdot \mathrm{Ln}(\exp(-Wz/Pz \cdot \lambda)) + 7.95, \qquad (4)$$

where W_z and P_z are the weighted means of the natural moisture and porosity through borehole section; J_h is the class of hydrophysical condition of a rock.

2.2. Jointing and Rock Mass Strength

Jointing of a rock mass (rock mass block structure) is a very important characteristic of the rock mass state under blasting or drilling. As a measure of a rock mass jointing, we chose an in situ rock block size Δ (cm). By analogy with the linear scale of the rock mass classification by block structure, we represent Δ on the second order canonical scale:

$$\Delta_j = \Delta_0 \cdot \left(\sqrt{2}\right)^{J-1},$$

where Δ_0 is the reference size of an in situ rock block. From the above expression, calculate the class of rock mass jointing:

$$J_j = 2.8854 \cdot \ln\left(\frac{\Delta}{\Delta_0}\right) + 1 \qquad (5)$$

Assuming $\Delta_0 = 9$ cm in (5), obtain:

$$J_j \approx 2.8854 \cdot \mathrm{Ln}(\Delta) - 5.33,\ 9 \leq \Delta \leq 150 \text{ cm} \qquad (6)$$

The range of the natural block size in a rock mass depending on the class of jointing is found:

$$6.36 \cdot \exp(0.3466 \cdot J_j) \leq \Delta \leq 9.0 \cdot \exp(0.3466 \cdot J_j).$$

2.3. Rock Mass Strength

Selecting a type and dimension of a bit, or a drilling regime is based on empirical relations including strength characteristics of rocks, such as hardness factor f; uniaxial compression σ_{com}; shearing strength σ_{shear}; contact strength σ_{cont}; indentation strength σ_{ind}; drilling strength σ_{drill}. These characteristics are also correlated. Therefore, for consistency, the characteristic of rock strength is accepted the uniaxial compression of dry rock, σ_d. Using:

$$\sigma_d = \sigma_0 \cdot \left(\sqrt{2}\right)^{j-1} \quad \text{and } \sigma_w = \sigma_d \cdot K_w \text{ at } \sigma_0 = 5.5 \text{ MPa},$$

obtain:

$$J_\sigma = 2.8854 \cdot \left[\ln(\sigma_d) - \ln(K_w)\right] - 3.92 \ . \tag{7}$$

2.4. Collective estimate ofr Mass condition by strength factors

The collective estimate is summing of (4), (6) and (7):

$$F_{str} = J_h + J_j + J_\sigma \ , \tag{8}$$

where F_{est} is the dimensionless rock mass condition.
After appropriate transformations:

$$F_{str} \approx 2.8854 \cdot [\ln(K_w) + \ln(\Delta) + \ln(\sigma_d)] - 1.3. \tag{9}$$

Placing minimum values (K_w = 0.3; Δ = 9; σ_d = 5.5) and maximum values (K_w = 1; Δ = 150; σ_d = 350) in (9) brings to $F_{str\ min}$ = 6.5 and $F_{str\ max}$ = 30. Based on that, it is accepted that the dimensionless rock mass condition in terms of strength varies from 6 to 30. This range governs the representation of Fstr by the second order canonical scale with base-2 and reference value F_{str} = 6. Quantization of the scale is bounded by a sequence number J_s, at which Fstr from (9) becomes ≥ 30. The quantization results are shown in Fig. 1 as the curves $J_s = f(F_{str})$ that illustrate the offered rock classification by strength. The ranges of Fstr and J_s are:

$$4.24 \cdot \exp(0.3466 \cdot J_s) \leqslant F_{str} \leqslant 6 \cdot \exp(0.3466 \cdot J_s);$$
$$[2.8854 \cdot \ln(F_{str}) - 5.17] \leqslant J_s \leqslant [2.8854 \cdot \ln(F_{str}) - 4.17].$$

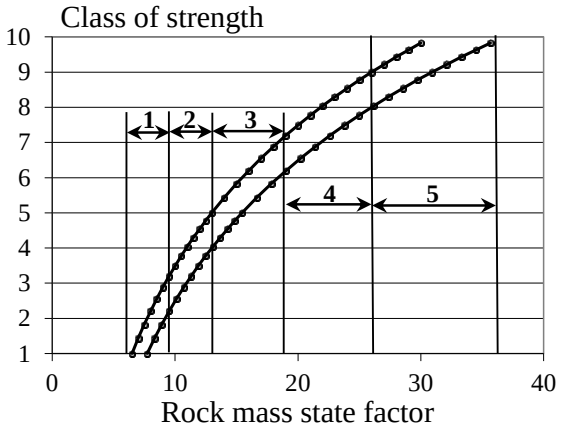


Fig.1. Rock mass classification by strength: 1—very weak rocks; 2—weak rocks; 3—medium strength rocks; 4—strong rocks; 5—very strong rocks

III. Rock Abrasivity Estimate by Physico-Mechanical Properties

To estimate resistance of rocks to drilling, we need to know the rock strength and another very important property that is the rock influence on the drilling tool wear, i.e. the rock abrasivity. This property is not a simple technological characteristic as compared to compression, tension etc. There are a lot of indirect methods for estimating rock abrasivity based on wear of an indenter made of various materials. Incommensurability of the results obtained with these methods impels the present authors to develop an abrasivity estimation procedure based on a set of physico-mechanical properties of rocks.

As the backbone characteristics of rock abrasivity, we select the shape and size of a grain; hardness of a rock-forming mineral and the mineral in whole, bond of grains, and moisture.

We omit description of the influence exerted by the above-listed characteristics on abrasivity of rocks (a more detailed information can be found in [4–10]) and only give the formal description of these characteristics required for the aggregate estimate of rock abrasivity.

3.1. Size and shape of grains, hardness of rock-forming minerals

On a canonical scale these characteristics are collection-wise presented in the relation:

$$\Psi_Z = \left[2.8854 \cdot \ln\left(\frac{D \cdot k_1}{d_0}\right) + 1\right] \cdot s_1; \tag{10}$$

$$s_1 = 0.022 \cdot \exp(0.5465 \cdot R_M), \tag{11}$$

where D and d_0 are the average mean size of a mineral grain and its references value, respectively, mm; k_1 is a coefficient accounting for the grain shape; s_1 is a relative measure of influence exerted by grains of certain hardness on rock abrasivity; R_M is the average mean hardness of rock-forming mineral. So, hardness of the mineral portion of a rock is:

$$R_M = 0.01 \sum_i M_i R_i \ ,$$

where M_i and R_i are an i-th mineral content of the rock (%) and the Mohs hardness, respectively.

The values of k_1 (10) are k_1 = 1 for roundish grains and k_1 = 1.3 for angular grains, given the grains are comparatively equal in size. By taking that this relationship is linear, we have: k_1 = 1 for well-rounded grains; k_1 = 1.11 for rounded grains; k_1 = 1.18 for half-rounded grains; k_1 = 1.25 for half-angular grains; k_1 = 1.30 for angular grains.

Concerning the reference size d_0 in (10), according to [11], grains with minimum size (0.005–0.05 mm) are typical of chemical rocks. On the assumption that grains smaller than 0.05 mm slightly affect abrasivity of rocks, we accept d_0 = 0.05 mm. Insert this value in (9) and obtain the granularity factor of rock abrasivity:

$$\Psi_G \approx [2.8854 \cdot \ln(D \cdot k_1) + 9.64] \cdot s_1;$$

$$0.05 \le D \le 8; \quad 1.0 \le k_1 \le 1.3. \tag{12}$$

3.2. POROSITY AND HARDNESS OF BINDING MATERIAL

Equally in magmatic and sedimentary rocks, pore "surface" may be composed of a material which has either close properties to the properties of the rock grains (crystal), or these properties differ considerably. In the material abrasive tools are made of, the hardness of grains is always higher than the binding material hardness. However, in rocks this relationship is not always well-formed. In sedimentary rocks the roughness and hardness of pore surface is greatly influenced by a cement type (in abrasive tool, a binding material type). In magmatic rocks pores are formed in a non-decrystallyzed material that influences the roughness and hardness of the pores surface, too. The genesis of voidage (no matter pores or fractures) is not important. What important is the hardness of the pore surface.

The porosity factor Ψ_P of rock abrasivity is presented on the canonical scale as follows:

$$\Psi_P = \left(1 - 2.8854 \cdot \ln\left(\frac{P_i}{p_0}\right)\right) \cdot s_2, \tag{13}$$

where P_i and p0 are, respectively, the porosity of a rock and the porosity reference value, %; s_2 takes into account the effect of the binding material hardness on the rock abrasivity.

With the accepted range of P_i from 0.2 to 25% and $p_0 = 25$, and using (13), we obtain:

$$\Psi_P \approx [-2.8854 \cdot \ln(P_i) + 10.29] \cdot s_2 \tag{14}$$

The coefficient s_2 is relative but, unlike s_1 introduced in (11), has the other meaning content: s_2 illustrates relative influence of microhardness of the binding material on rock abrasivity rather than the relative hardness of the rock-forming mineral. The values of s_2 are found from test data on the binding material (cement, glass) microhardness. We make use of ratios of the Mohs microhardness and microhardness found on the PMT-3 microhardness measurement device. We are interested in not the absolute but relative values of hardness, i.e. s_2. The following results were obtained: the PMT-3 measurement microhardness:

$$s_2 = 8.9 \cdot 10^{-5} \cdot T_{PMT\text{-}3},$$

the Mohs hardness:

$$s_2 \approx 2.32 \cdot 10^{-3} \cdot (R_M)^{3.12},$$

where $T_{PMT\text{-}3}$ is the microhardness measured using the PMT-3 device; $R_м$ is the Mohs hardness.

3.3. BONDING STRENGTH OF GRAINS

The influence of the grain bonding strength on rock abrasivity is poorly studied and for sedimentary rocks it is hypothetically based on the strength of a cement. With the strong cement, each grain at its place abrades a cutting tool until the grain is detached and removed from the tool and rock interaction zone. Hence it appears that with the stronger cement, the rock abrasivity increases. This statement is valid for crystal rocks, as well. With the weak bonding strength of grains, abravise refresh gets intensified. The bonding strength of grains is estimated in terms of ultimate tension σ_t (MPa). Assume that rock abrasivity grows with increasing σ_t, then the grain bonding strength factor Ψ_B is:

$$\Psi_B = \left[2.8854 \cdot \ln\left(\frac{\sigma_t}{\sigma_{0t}}\right) + 1\right] \cdot K_w \tag{15}$$

where σ_t is the ultimate tension of dry rock, MPa.

Let the reference value be $\sigma_{0t} = 0.312$ MP, that is typical of weak hard rocks. Placing this value in (15) and appropriate transformation yields:

$$\Psi_B \approx [2.8854 \cdot \ln(\sigma_t) + 4.36] \cdot K_w, \; 0.312 \le \sigma_t \le 40. \tag{16}$$

The bonding strength of grains is affected by rock moisture content, especially in sedimentary rocks. The quantitative estimates of the moisture and abrasivity relationship in rocks are given in [9], where it is shown that abrasivity of water-saturated rocks decreases greatly (to 30%). This effect is included in (15) in terms of the water resistance coefficient K_w discussed above in this article.

3.4. AGGREGATE ABRASIVITY OF ROCKS BY THE COLLECTIVE PHYSICO-MECHANICAL PROPERTIES. CLASSIFICATION

The framework of the rock abrasivity estimate is the set of (12), (14) and (16):

$$Ă = \Psi_G + \Psi_P + \Psi_B,$$

where Ă is the abrasivity of a rock in conventional dimensionless units.

$$Ă \approx 2.8854 \cdot \{\ln(D \cdot k_1) \cdot s_1 - \ln(P) \cdot s_2 + \ln(\sigma_t) \cdot K_w\} + 9.64 \cdot s_1 + 10.29 \cdot s_2 + 4.36 \cdot K_w. \tag{17}$$

The estimate (17) characterizes abrasivity of rocks depending on the rock's physico-mechanical properties. In the conventionally measured abrasive wearing capacity of indenters, abrasivity has unit measures but they are relative, and are only used in classifications the objective of which is to refer this or that type of rocks to a suitable class of abrasivity. It's the abrasivity classes which are of importance in selecting a rock-cutting or rock-drilling tool, and not their quantitative evaluations.

With the accepted minimum values (D = 0.05 mm; $k_1 = 1$; $s_1 = 0.065$; P = 0.2%; $s_2 = 0.02$; $\sigma_t = 0.312$ MPa) and maximum values (D=10 mm; $k_1 = 1.3$; $s_1 = 1$; P = 25%; $s_2 = 1$; $\sigma_t = 40$ MPa) of the parameters included in (17), we obtain the range of Ă: $Ă_{min} = 3.2$; $Ă_{max} = 63$. Quantification of Ă within this range is represented as the second order canonical scale, on the assumption that $Ă_0 = Ă_{min} = 3.0$:

$$Ă_J = Ă_0 \cdot \left(\sqrt{2}\right)^{J_a - 1}, \; J = 1,\ldots,9, \tag{18}$$

Table 1. Classifications of rocks by abrasivity.

Abra-sivity class	Aggregate abrasivity, conventional units (range)	Abrasivity by [12]		Characteristic of rocks
		Initial scale	Canonical scale	
1	2	3	4	
1	3.0–4.2	<5	4.0–5.7	Very low abrasivity
2	4.2–6.0	5–10	5.7–8.0	Low abrasivity
3	6.0–8.5	10–18	8.0–11.0	
4	8.5–12.0	18–30	11.0–16.0	Subaverage abrasivity
5	12.0–17.0	30–45	16.0–23.0	Average abrasivity
6	17.0–24.0	45–65	32.0–45.0	Above-average abrasivity
7	24.0–34.0	65–90	45.0–64.0	Increased abrasivity
8	34.0–48.0	>90	64.0–91.0	High abrasivity
9	48.0–68.0		>91.0	Extremely high abrasivity

Table 2. Rock classification by drilling resistance

Rock mass strength factor (F_{str} by (9))								
6–9	9–13	13–18	18–26	26–37				
Rock strength classes								
1–2	2–3	3–4	4–5	5–6				
Rock abrasivity factor (A by (17))								
3–4	4–6	6–9	9–12	12–17	17–24	24–34	34–48	48–67
Rock abrasivity class (J_a by (19))								
1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10
Aggregate rock drilling resistance (F_{str} + A by (21))								
8–12	12–17	17–24	24–34	34–48	48–67	67–95	95–135	
Rock drilling resistance classes								
0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8	

where J_a is the class of abrasive capacity of the rock.

The results of the calculations by (18) are compared with the rock abrasivity classification by Baron and Kuznetsov [12] in Table 1.

By the data from the second column in Table 2, we obtain the abrasivity ranges within a class:

$$2.12\cdot\exp(0.3466\cdot J_a) \le Ă \le 3\cdot\exp(0.3466\cdot J_a),\ J_a=1,...,9. \quad (19)$$

It follows from (19) that:

$$2.8854\cdot\ln(Ă) - 3.17 \le J_a \le 2.8854\cdot\ln(Ă)+2.17.$$

The side-by-side comparison of the calculations by (19) and the data of the Baron and Kuznetsov classification (column 4 in Table 1) is possible using the plots in Fig. 2a, with the pre-found:

$$2.1\cdot\exp(0.496\cdot Y) \le B \le 4.12\cdot\exp(0.45\cdot Y);\ Y=1\div7, \quad (20)$$

where B and Y are, respectively, the value of abrasivity index (wear of indenter, mg) and the class of abrasivity by Baron and Kuznetsov.

From the comparison of (19) with (20), the abrasivity indexes in both classifications exponentially depend on their classes though have different units of measurement and different amount of classes. The difference of the coefficients attached to the exponents in (19) and (20) is conditioned by different approaches to ranging the classification indexes in classes. As for the difference of the coefficients in front of the exponents in (19) and (20), this is because of the different units of measurement used in the classifications.

For the purpose of comparison, we present the classification by Baron and Kuznetsov on the canonical scale with the chosen reference value B = 3 mg (Fig. 2b).

The classifications are fairly close, and we can relate the aggregate rock abrasivity index obtained based on physico-mechanical properties of rocks, Ă, with the rock abrasivity based on steel indenter wear by the Baron and Kuznetsov method, B:

$$2.89\cdot Ă-6 \le B \le 2.75\cdot Ă-4.46.$$

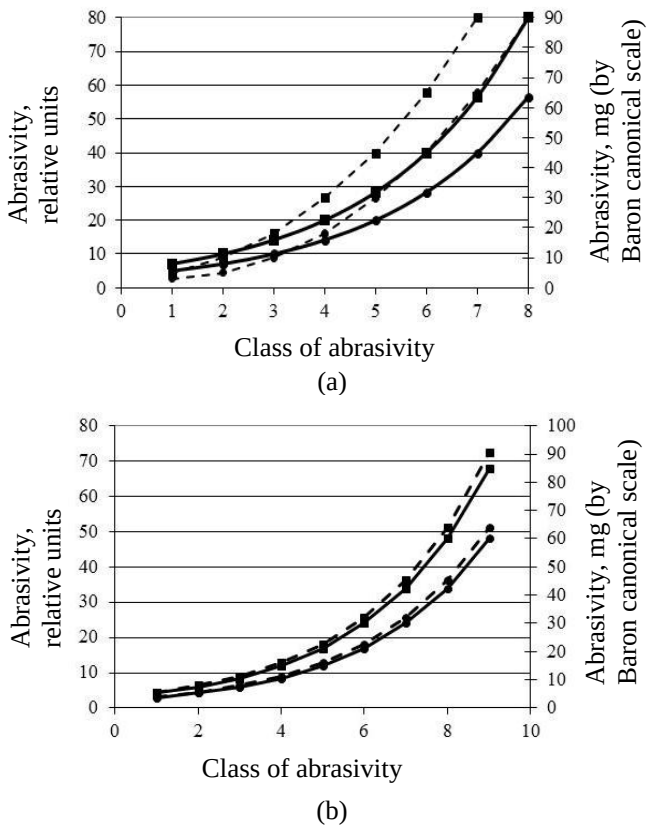


Fig. 2. Classifications of rocks by abrasivity: (a) classification by abrasivity (solid lines) versus the classification by indenter wear (dashed lines); (b) classification by abrasivity (solid lines) versus the classification by indenter wear on canonical scale (dashed lines).

We have compared the proposed method of rock abrasivity assessment by the aggregate physico-mechanical properties of rocks with the method of rock abrasivity assessment by Baron and Kuznetsov because the latter is widely used in mining industry.

The idea to apply aggregate physico-mechanical properties of rocks for estimating rock abrasivity appeared productive. This is confirmed by the fact of coherence of the classifications of rocks by abrasivity based on their physico-mechanical properties and based on steel indenter wear.

IV. Classification of Rocks by Drilling Resistance

Rock drilling resistance Є is found as $Є = F_{str} + Â, ,$ i.e., we sum up the right-hand sides of (9) and (17) and, after transformations, have:

$$Є = 2.8854\cdot\{\ln(K_w) + \ln(\Delta) + \ln(\sigma_d) + \ln(D\cdot k_1)\cdot s_1 - \ln(P)\cdot s_2 + \ln(\sigma_t)\cdot K_w\} + 9.64\cdot s_1 + 10.29\cdot s_2 + 4.36\cdot K_w - 1.3$$

$$\ln(\sigma_t)\cdot K_w\} + 9.64\cdot s_1 + 10.29\cdot s_2 + 4.36\cdot K_w - 1.3. \qquad (21)$$

With the minimum and maximum values of the parameters in (9) and (17), we obtain from (21) that Є falls in the following range: $8.4 \le Є \ge 95$, accounting for rounding off. Based on that we set a clustering procedure for Є in the canonical series with the denominator. The results of clustering Є on the canonical scale are given in Table 2 that is the classification of rocks by drilling resistance based on physico-mechanical properties rocks. Using the data from Table 2, we have related the rock drilling resistance index Є and J_σ:

$$5.93\cdot\exp(0.3466\cdot J_\sigma) \le Є \le 8.4\cdot\exp(0.3466\cdot J_\sigma);$$
$$\{2.8854\cdot\ln(Є) - 6.14\} \le J_\sigma \le \{2.8854\cdot\ln(Є) - 5.14\}.$$
$$J_\sigma = 1,\ldots, 8.$$

From the comparison of the proposed classification of rocks by drilling resistance and the rock drillability classification suggested by the Central Research Institute for Geological Exploration in Figure 3, it is evident that the classifications are equitype.

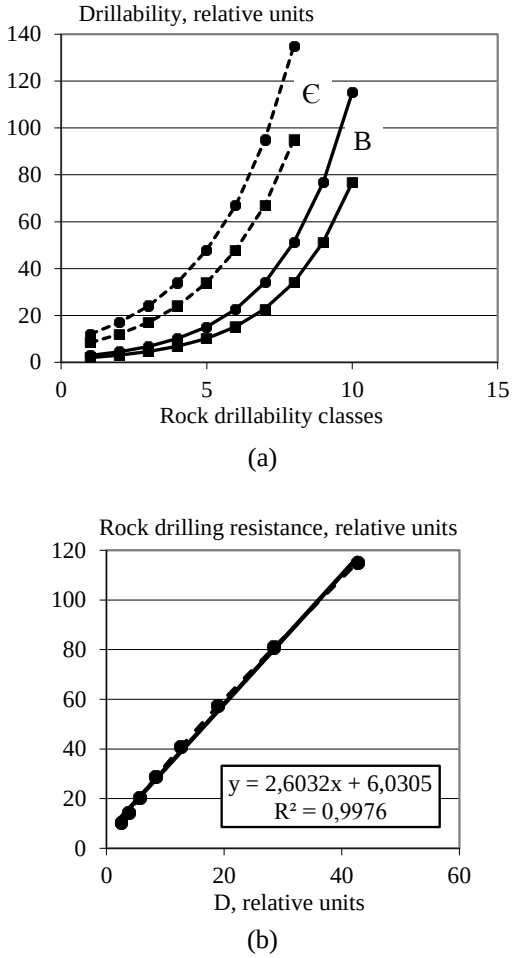


Fig. 3. Comparison of the rock drilling resistance classification based on aggregate physico-mechanical properties (Є) with the rock drillability classification offered by the Central Research Institute for Geological Exploration (D).

The deduction above is confirmed by the functional relationship Є = f(D) with the correlation ratio $R^2 = 0.99$ in Figure 3b. The method of estimation of rock drilling resistance by the aggregate of physico-mechanical properties of rocks

offered in this article is quite appropriate to describe the essence of drilling.

V. Conclusion

The authors have proposed the new method of rock classification by drilling resistance, that is invariant regarding drilling equipment and uses borehole drilling data on physico-mechanical properties of rocks.

ACKNOWLEDGMENTS

The work was supported by partly the Russian Academy of Sciences, project ONZ-RAN 3.1, Siberian Branch of the Russian Academy of Sciences, integration project no. 100. The equipment is kindly provided by the Shared Geomechanical, Geophysical and Geodynamic Measurement Center of the Siberian Branch of the Russian Academy of Sciences.

References

[1] Oparin V.N., Tanaino A.S., Canonical Scale for Representation of Hierarchies in Science on Rocks. Novosibirsk: Nauka, 2011.

[2] Russian Federation Standard R-50544-93. Rocks. Terms and Definitions. Moscow, 1994

[3] Solodukhin M.I., Arkhangelsky, I.V., Geotechnic and Hydrological Survey Reference Guide for GeoTechnicians. Moscow: Nedra, 1982.

[4] Abrasmon M.G., Baidyuk B.M., Zavaretsky V.S. et al., Reference Guide on Mechanical and Abrasive Properties of Oil and Gas Host Rocks. Moscow: Nedra, 1985.

[5] Vozdvizhensky V.I., Melnichuk I.P., Peshalov Yu.A., Physico-Mechanical Properties of Rocks and Their Influence on Drilling. Moscow: Nedra, 1973.

[6] Baron L.I., Glatman L.B., Contact Strength of Rocks. Moscow: Nedra, 1966.

[7] Shreiner L.A., Petrova O.P., Yakushev V.P., Portnova A.T. et al., Mechanical and Abrasive Properties of Rocks. Moscow: Gostoptekhizdat, 1958.

[8] Golubintsev O.N., Mechanical and Abrasive Properties of Rocks and the Rock Drillability, Moscow: Nedra, 1968.

[9] Spivak A.I., Popov A.N., Destruction of Rocks in Borehole Drilling. Moscow: Nedra, 1985.

[10] Mavlyutov M.R., Rock Failure in Borehole Drilling. Moscow: Nedra, 1978.

[11] Pettijohn F.J., Sedimentary Rocks. Harper, 1949.

[12] Baron L.I., Kuznetsov A.V., Abrasivity of Rocks in Mining. Moscow: AN SSSR, 1961.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 6

Drilling Mud Loss in a Natural Fracture - A PKN Fracture Geometry Model Based Analysis

天然裂缝中钻井液的漏失-基于PKN裂缝几何模型的分析

Yixuan Sun（孙义轩）, Haiying Huang*（黄海鹰）

School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

haiying.huang@ce.gatech.edu

Accepted for publication on 5th August 2014

Abstract - **A mathematical model is formulated in this work to determine the drilling mud loss in a natural fracture intersected by a wellbore. The formulation is based on a PKN fracture geometry model, assuming the natural fracture intersects the wellbore with the fracture plane more or less parallel to the wellbore axis. The drilling mud is treated as a single-phase fluid. The rheology of the drilling mud is assumed to be non-Newtonian of Herschel-Buckley type. The reservoir formation is assumed to be permeable with the leak off following Carter's model. Fluid flow along the fracture is considered as one-dimensional lubrication flow. Deformation of the fracture is governed by local elasticity and the wellbore pressure and the pore pressure in the reservoir are assumed to be constant. The problem defined above is solved numerically using an explicit moving mesh algorithm. Effects of the yield stress and the overbalance pressure on the drilling mud loss are evaluated.**

Keywords – **drilling mud loss, PKN fracture model, natural fracture**

摘要 **－ 天然裂缝性地层容易出现严重的井漏。在过平衡钻井时，大量钻井液的漏失不仅影响钻井操作，而且还会对随后的完井及油藏的生产产生不良影响。为了能够合理地采取措施减少井漏，我们迫切需要对钻井液在天然裂缝的漏失进行定量预测。本文将建立一个基于PKN裂缝几何形状的数学模型对天然裂缝中钻井液的漏失进行模拟计算。我们假设天然裂缝与井筒相交，裂缝平面大致平行于井孔轴。钻井液将被视为单相流体。钻井液的流变特性采用非牛顿流体的赫谢尔-巴克利模型。钻井液在地层中的渗透将由卡特模型来描述,在裂隙中的流动是一维的，可以由流体润滑方程来描述。地层的变形假设是局部弹性的。井中压力和油藏孔隙压力假设是常值。 显式移动网格算法将被用于求解上述问题。本文将对流变参数和过平衡压力对钻井液漏失的影响进行评估。**

关键词 **－ 钻井液漏失, PKN裂缝模型, 天然裂缝**

I. INTRODUCTION

Naturally fractured reservoirs are prone to severe circulation losses. In overbalance drilling, since the wellbore pressure is larger than the reservoir pore pressure, large volume loss of the drilling mud in fractured formations could be a significant problem not only in the drilling operation but also during subsequent well completion and reservoir production. Quantitative prediction of the drilling mud loss is therefore critically needed for the control of lost circulation. In general, drilling mud may be treated as a single-phase incompressible fluid displaying non-Newtonian rheological characteristics such as yield stress and shear thinning. The yield stress is the most critical rheological element responsible for the eventual cessation of mud invasion. Loss of the drilling mud in the natural fracture is driven by the pressure drop between the wellbore and the reservoir. The pressure gradient near the wellbore is the largest at the onset of drilling mud invasion and decreases as the mud flows further into the fracture. For a yield stress fluid, when the pressure gradient falls below a threshold, mud invasion ceases and the fracture can be considered sealed. The Herschel-Buckley model with a yield stress and power law for shear thinning has been considered an accurate description for the drilling mud rheology, especially in the low shear rate range [1].

Drilling mud loss in a natural fracture has been previously modeled in the literature assuming radial flow through two parallel disks with or without the consideration of reservoir deformation, e.g. [2][3]. The assumption of radial flow is most applicable when the natural fracture is nearly perpendicular to the wellbore axis.

In this work, a mathematical model is formulated to determine the drilling mud loss for the scenario when the natural fracture intersects the wellbore with the fracture plane aligned more or less parallel to the wellbore axis. The natural fracture is initially closed. Fluid flow inside the fracture is coupled with the mechanical deformation of the permeable reservoir. The fracture opening developed as a result of mud invasion is assumed to be elliptical with a constant height, following the Perkins-Kern-Nordgen (PKN) geometry model [4][5].

An explicit moving mesh algorithm modified from those for solving the PKN hydraulic fracturing model for power law [6] and piecewise power law rheology [7] is employed in this study. The mathematical formulation and the numerical implementation are first presented. Effects of the yield stress

and the overbalance pressure on the drilling mud loss are then analyzed.

II. MATHEMATICAL FORMULATION

Schematic of a PKN geometry model of length L and height H is shown in Fig. 1. A key assumption in the PKN hydraulic fracture model is that the elastic response of the rock is local, namely, the opening of the hydraulic fracture at a given position x depends on the local overpressure only, but not on the pressure everywhere else inside the fracture, i.e.,

$$p = \frac{w}{M_c} \tag{1}$$

Eq. (1) states that the average fracture width w of a cross section is related to the net local pressure p through the fracture compliance $M_c = \pi(1-\nu)H/4G$, where G and ν are the shear modulus and the Poisson's ratio of the rock, respectively. The net pressure p, i.e., the overpressure above the reservoir stress, is assumed to vary only along the $x-$axis, namely, the direction of fracture propagation, and is uniform at a given cross section. The average width is defined according to $w = A/H$, where A is the cross sectional area. For an elliptical cross section, if the width at the mid-height is w_0 (see Fig. 1), $w = \pi w_0/4$. The local elasticity assumption implies each cross section is deforming independently under the condition of plane strain, which is most applicable when the fracture length is much larger than the fracture height and the fracture width is much smaller than the height [8]. The local elasticity assumption relieves us of solving the integral form pressure-width relation required in other hydraulic fracture geometry models such as the KGD [9][10] and the penny-shaped fracture models [11]. The PKN hydraulic fracture model therefore offers unique simplicity that allows incorporation of complex fluid rheology.

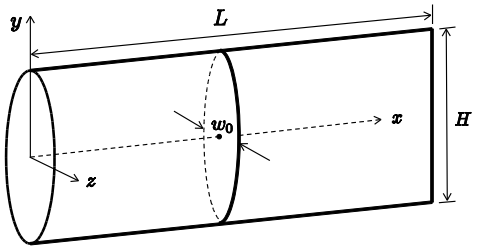


Fig. 1, A PKN fracture of length L and height H; at a given cross section, the fracture width at the mid-height is denoted as w_0.

We may assume that fluid flow along the fracture is one-dimensional. Consequently, the local continuity equation can be written as,

$$\frac{\partial q_x}{\partial x} + \frac{\partial w}{\partial t} + u = 0 \tag{2}$$

where q_x is the flow rate per unit height of a cross section; t is the elapsed time from the onset of drilling mud invasion and u is the fluid leak off velocity from the fracture faces. For a relatively low permeability formation, fluid leak off may be assumed to be one-dimensional obeying Carter's leak off model,

$$u = \frac{2C_L}{\sqrt{t - t_a(x)}} \tag{3}$$

where C_L is the leak off coefficient and $t_a(x)$ is the fracture tip arrival time at location x. Carter's leak off model, Eq. (3), relates the leak off velocity u to the time of exposure for a given position. The velocity u accounts for the leak off from both sides of the fracture faces.

Fluid rheology is only involved in the equation of balance of fluid momentum for the PKN hydraulic fracture model. With the assumption of one dimensional lubrication flow, the equation of balance of momentum can be approximated by integrating the Poiseuille slot flow solution over the height of the elliptical cross section. The slot flow solution for a Herschel-Buckley fluid therefore needs to be derived first.

For a Herschel-Buckley fluid, the relationship between the shear stress and the shear strain rate can be expressed as,

$$\tau = \begin{cases} \tau_0 & \dot{\gamma} = 0 \\ \tau_0 + K\dot{\gamma}^n & \dot{\gamma} > 0 \end{cases} \tag{4}$$

where τ_0, K and n are the yield stress, the consistency parameter and the power index. Fluid flow occurs only if the shear stress τ exceeds the yield stress τ_0.

After combining the rheology model with the equilibrium equation, the geometrical equation, and the no-slip condition at the slot walls, the steady state velocity profile for slot flow can be readily obtained,

$$V_x = \begin{cases} \dfrac{bn\left(\dfrac{\bar{p}b}{K}\right)^{\frac{1}{n}}}{1+n}\left[(1-s_*)^{1+\frac{1}{n}} - (|s|-s_*)^{1+\frac{1}{n}}\right] & |s| > s_* \\ \dfrac{bn\left(\dfrac{\bar{p}b}{K}\right)^{\frac{1}{n}}}{1+n}(1-s_*)^{1+\frac{1}{n}} & |s| \le s_* \end{cases} \tag{5}$$

where V_x is the fluid flow velocity in the flow direction; b is the half width of the slot; $s = y/b$ is the scaled coordinate $(-1 \le s \le 1)$ in the slot width direction and $\bar{p} = |dp/dx|$ is the magnitude of the pressure gradient in the fluid flow direction. Due to the existence of the yield stress, within the region of $|s| \le s_*$, the velocity is constant and fluid flows in a plug form. The critical pressure gradient to initiate fluid flow is $\bar{p}_c = \tau_0/b$ and the half width s_* of the plug flow region can be determined from $s_* = \tau_0 b/\bar{p}$.

The total flow rate Q_x^s at a given pressure gradient $\bar{p}$ for slot flow can be obtained by integrating the velocity across the slot width,

$$Q_x^s = \frac{2b^{2+\frac{1}{n}}\left(\dfrac{\bar{p}}{K}\right)^{\frac{1}{n}}}{1+\dfrac{1}{n}}(1-s_*)^{1+\frac{1}{n}}\left(\frac{n+1+s_*}{2n+1}\right) \tag{6}$$

Fig. 2 shows the variation of the velocity profiles with respect to the applied pressure gradient $\bar{p} = 0.05, 0.1$ and 0.2 MPa/m for the following rheological parameters: $\tau_0 = 10$ Pa, $n = 0.7$, and $K = 0.3$ Pa·s^n. As can be seen from Fig. 2, the plug zone size reduces as the pressure gradient $\bar{p}$ increases. The pressure gradient and the flow rate relationship as shown

in Fig. 3 indicates that the flow rate is zero before the yield stress is overcome by the pressure gradient along the slot.

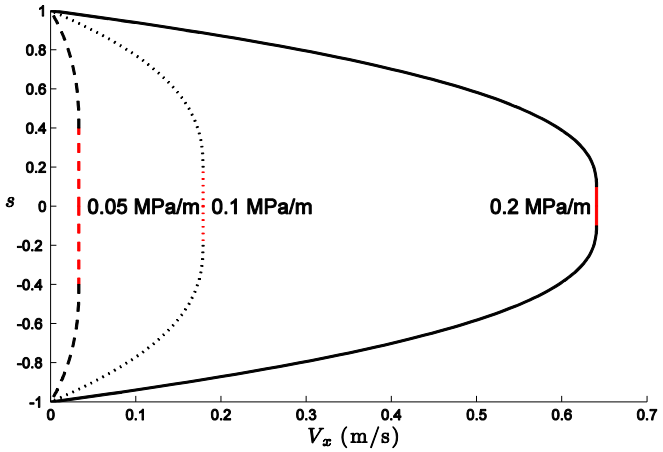


Fig. 2, Velocity profiles under different pressure gradients; the red lines denote the plug flow zones.

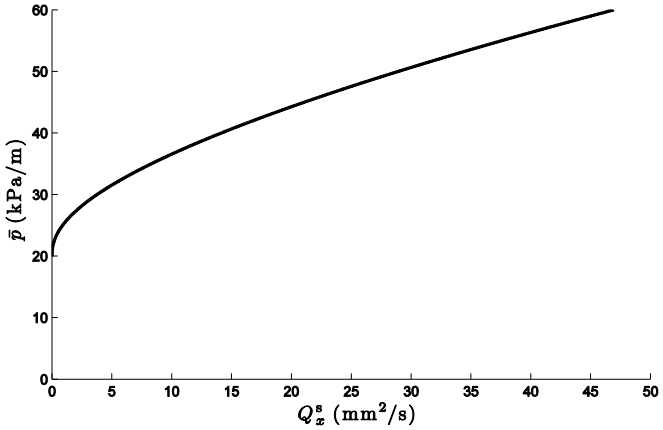


Fig. 3, Variation of the pressure gradient with respect to the total flow rate for slot flow.

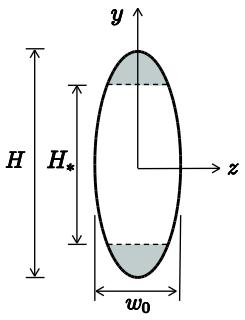


Fig. 4, A PKN model cross section for a Herschel-Buckley fluid.

With the assumption of one-dimensional lubrication flow inside the fracture, the existence of the yield stress is manifested through the stagnation zones near the two ends of the fracture cross section, as shown in the shaded areas in Fig. 4. The ratio between the height H_* of the moving fluid region and the total height H can be determined from,

$$\frac{H_*}{H} = \sqrt{1 - \left(\frac{2\tau_0}{\bar{p} w_0}\right)^2} \tag{7}$$

At a given height position, the local flow rate Q_x can be obtained based on the slot flow solution, Eq. (6),

$$Q_x = \begin{cases} \dfrac{2\left(\frac{w_y}{2}\right)^{2+\frac{1}{n}}\left(\frac{\bar{p}}{K}\right)^{\frac{1}{n}}}{1+\frac{1}{n}}(1-s_*)^{1+\frac{1}{n}}\left(\dfrac{n+1+s_*}{2n+1}\right) & |y| < \dfrac{H_*}{2} \\ 0 & \dfrac{H_*}{2} \le |y| < \dfrac{H}{2} \end{cases} \tag{8}$$

where w_y is the fracture width at a given height y, $w_y = w_0\sqrt{1-(2y/H)^2}$.

The balance of momentum equation relating the total flow rate or the mud loss rate per unit height q_x to the local pressure gradient $\bar{p}$ can be obtained by integrating Eq. (8) over the fracture height,

$$q_x = \frac{C_1}{\bar{p}^2}\left[M_1 - \frac{1}{3}\left(\frac{H_*}{H}\right)^2 M_2\right]\left[\left(\frac{w_0}{2}\right)\bar{p} - \tau_0\right]^{1+1/n}\left[\left(1+\frac{1}{n}\right)\left(\frac{w_0}{2}\right)\bar{p} + \tau_0\right] \tag{9}$$

where,

$$\begin{aligned} C_1 &= \frac{2}{\left(1+\frac{1}{n}\right)\left(2+\frac{1}{n}\right)K^{\frac{1}{n}}}\left(\frac{H_*}{H}\right) \\ M_1 &= {}_2F_1\left[\frac{1}{2}, -\frac{1}{2n}, \frac{3}{2}, \left(\frac{H_*}{H}\right)^2\right] \\ M_2 &= {}_2F_1\left[\frac{3}{2}, -\frac{1}{2n}, \frac{5}{2}, \left(\frac{H_*}{H}\right)^2\right] \end{aligned} \tag{10}$$

where ${}_2F_1[\ \ ]$ denotes the hypergeometric function.

To complete the formulation, boundary conditions both at the wellbore and at the fracture tip need to be supplied. We may assume that the wellbore overpressure P_{in} is constant and the initial reservoir pore pressure and the reservoir stress are both constant, which serve only as references. In addition, no flow condition can be imposed at the fracture tip $x = L$, i.e.,

$$\begin{aligned} p &= P_{\text{in}} \quad \text{at } x = 0 \\ w = 0,\ q_x &= 0 \quad \text{at } x = L \end{aligned} \tag{11}$$

III. Numerical Implementation

The system of governing equations, Eqs. (1) – (3) and Eq. (9), is first transformed by introducing a moving coordinate $\theta = x/L(t)$, $\theta \in [0,1]$. An equation for the fracture width can be obtained as follows,

$$\dot{w}_0 = \frac{\theta \dot{L}}{L}\frac{\partial w_0}{\partial \theta} + c_w \frac{\partial^2 w_0}{\partial \theta^2} + c_s\left(\frac{\partial w_0}{\partial \theta}\right)^2 \tag{12}$$

where $\dot{(\ )} = d(\ )/dt$ is the material time derivative and c_w and c_s are functions of the material parameters and the fracture width w_0 and the gradient $\partial w_0/\partial\theta$. For brevity, expressions for c_w and c_s are omitted here.

Such a "diffusion-type" of equation for the width w_0 can be solved by using an explicit finite-difference scheme [6]. After

discretizing the fracture length into N nodes, at a given time step, say, $t = t^{k+1}$, the critical time step to ensure numerical stability can be computed from,

$$\Delta t_c^{k+1} = \min\left[\frac{\Delta\theta_i^2}{2(c_w)_i^k}\right];\ \ i = 1,\cdots,N-1 \qquad (13)$$

Here the nodal number is denoted by the subscript i and the time step is denoted by the superscript k for simplicity, e.g., $(c_w)_i^k = c_w(\theta_i, t^k)$. In the time-stepping algorithm, after knowing all the nodal quantities in the previous time step, a critical time step is first chosen. The fracture width w, the net fluid pressure p in the fracture, the local flow rate q_x, the leak off velocity u, and the fracture length L, are computed sequentially.

Since the numerical scheme is explicit, an initial guess is required to start the calculation. The analytical solution for the PKN model in an impermeable formation with a power law fluid [12] is used as the initial guess for the fracture length, width and the flow rate at an arbitrarily small starting time. As shown in [7], with the proper choice of the critical time step, the numerical scheme is robust and the solutions are not sensitive to the initial guesses obtained from different rheology.

IV. Numerical Results

Effects of the yield stress and the overbalance pressure on the drilling mud loss behaviors are investigated using the set of input parameters as listed in Table I.

TABLE I, INPUT PARAMETERS FOR THE SIMULATION CASES

Consistence parameter, K	0.3 Pa·s$^{0.7}$
Power law index, n	0.7
Poisson's ratio, ν	0.2
Shear modulus, G	10 GPa
Fracture height, H	10 m
Leak-off coefficient, C_L	6.3×10^{-5} m/s$^{1/2}$

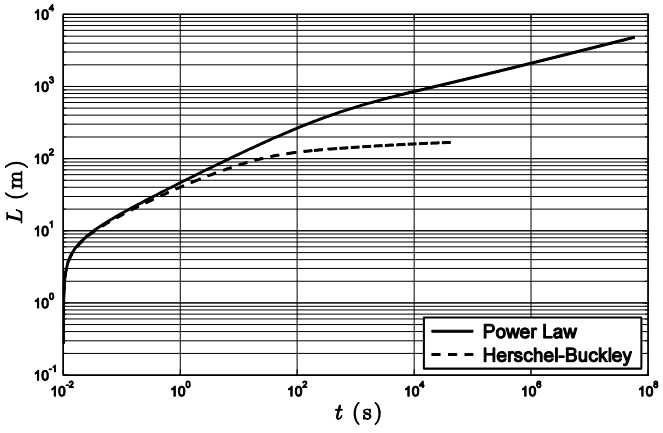


Fig. 5, Fracture length as a function of time from the simulation cases with a power law fluid and a Herschel-Buckley fluid ($\tau_0 = 100$ Pa) with an overpressure $P_{\text{in}} = 10$ MPa.

Variations of the fracture length or the mud invasion length with time for a power law fluid and a Herschel-Buckley fluid are compared in Fig. 5. The two cases have the same power law parameters and the wellbore overpressure is $P_{\text{in}} = 10$ MPa. The yield stress for the Herschel-Buckley fluid is $\tau_0 = 100$ Pa. The Herschel-Buckley solution coincides with the power-law solution at early time, but gradually deviates from the power law solution at late time. This means that fluid flow is governed by the high shear rate rheology at early time, but by the low shear rate rheology at late time. Eventually, the Herschel-Buckley solution reaches a plateau, indicating that the fracture is no longer propagating. In other words, there is no further mud invasion into the fracture. Meanwhile, for the power law case, the absence of the yield stress results in continuous fracture growth and the fracture cannot be sealed.

Evolution of the fracture half-width is plotted against the fracture length at different times for a case with $\tau_0 = 10$ Pa and $P_{\text{in}} = 5$ MPa. Since the fracture width at the inlet is governed by the inlet pressure and the fracture compliance, it remains constant as the fracture propagates. Meanwhile, the gradient of the fracture width, and thus the pressure gradient, decreases and the rate of increase in the fracture length or the mud invasion length decreases with time.

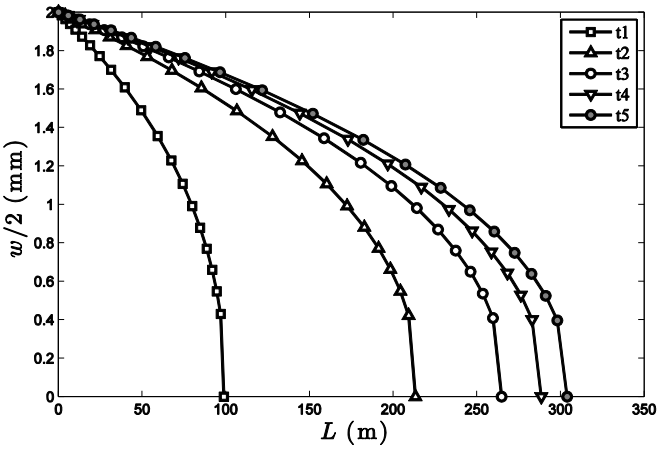


Fig. 6, Evolution of the fracture width profile with time with $P_{\text{in}} = 5$ MPa and $\tau_0 = 10$ Pa, and $t_1 = 506.6$ s, $t_2 = 8.558e4$ s, $t_3 = 4.709e5$ s, $t_4 = 9.949e5$ s, $t_5 = 1.608e6$ s.

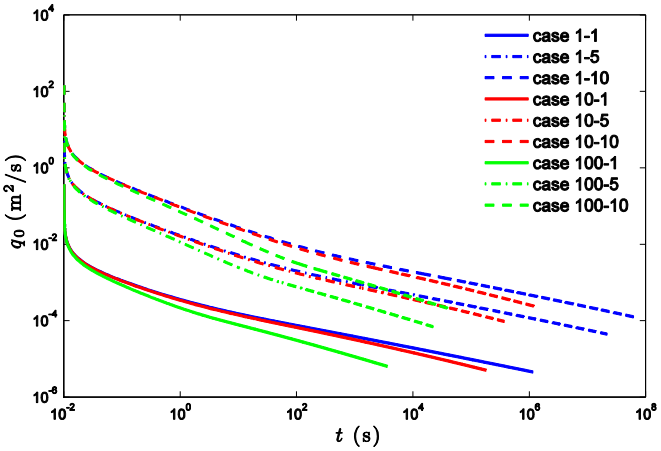


Fig. 7, Evolution of the inlet flow rate per unit height q_0 with time t; the case numbers refer to the yield stress in Pa and the overpressure in MPa; namely, case 10-1 is a test case with $\tau_0 = 10$ Pa and $P_{\text{in}} = 1$ MPa.

A series of simulations are conducted by varying the yield stress and the inlet overpressure, $\tau_0 = 1,\ 10,\ 100$ Pa and

$P_{in} = 1, 5, 10$ MPa, to investigate their influences on the mud loss rate at the inlet q_0 (per unit height) and the fracturing efficiency η. The fracturing efficiency is defined according to $\eta = (V - V_\ell)/V$, where V is the total volume of mud loss and V_ℓ is the volume of fluid leaked into the formation. The efficiency measures the percent of fluid volume stored inside the fracture. Though, in the range of parameters explored, the drilling mud loss behavior is predominantly affected by the prescribed inlet pressure, increasing the yield stress to $\tau_0 = 100$ Pa has a considerable effect on the mud loss rate. As can be seen from Fig. 7, the mud loss rate for $\tau_0 = 100$ Pa and $P_{in} = 10$ MPa indeed becomes smaller than that in the cases with $\tau_0 = 1$ and 10 Pa and $P_{in} = 5$ MPa at late time.

As shown in Fig. 8, the fracturing efficiency decreases with time, suggesting that at the early stage, a higher percentage of the drill mud loss is in opening up the natural fracture. Matrix leak off becomes relatively important only at the later stage.

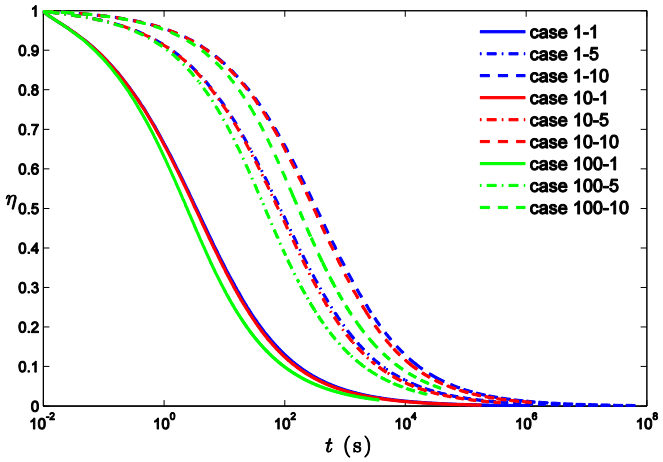


Fig. 8, Variation of the fracturing efficiency η with time t; the case numbers refer to the yield stress in Pa and the overpressure in MPa; namely, case 10-1 is a test case with $\tau_0 = 1$ Pa and $P_{in} = 1$ MPa.

V. Conclusion

In this paper, a mathematical model for predicting drilling mud loss in a natural fracture is developed based on the classical PKN hydraulic fracture geometry model. Rheology of the drilling mud is assumed to be non-Newtonian of Herschel-Buckley type. The numerical scheme using an explicit moving mesh algorithm provides a robust tool that allows not only systematic investigation of the effects of the formation and the fluid characteristics on the drilling mud loss, but also improvement in the fluid design to control lost circulation.

Acknowledgements

Support from the National Science Foundation through grant NSF/CMMI-1055882 is gratefully acknowledged.

References

[1] T. Hemphill, W. Campos, and A. Pilehvari. "Yield-power law model more accurately predicts mud rheology," *Oil and Gas Journal*, **91**, pp. 1993.

[2] A. Lavrov. "Modeling flow of a biviscous fluid from borehole into rock fracture," *Journal of Applied Mechanics*, **73**, pp. 171-173, 2006.

[3] R. Majidi, S. Miska, R. Ahmed, M. Yu, and L. Thompson. "Radial flow of yield-power-law fluids: Numerical analysis, experimental study and the application for drilling fluid losses in fractured formations," *Journal of Petroleum Science and Engineering*, **70**, pp. 334-343, 2010.

[4] T. Perkins and L. Kern. "Widths of hydraulic fractures," *Journal of Petroleum Technology*, **13**, pp. 937-949, 1961.

[5] R. Nordgren. "Propagation of a vertical hydraulic fracture," *SPE Journal*, **12**, pp. 306-314, 1972.

[6] E. Detournay, A. H-D. Cheng, and J. McLennan, "A poroelastic PKN hydraulic fracture model based on an explicit moving mesh algorithm," *Journal of Energy Resources Technology*, **112**, pp. 224-230, 1990.

[7] H. Huang and J. Desroches. "A PKN hydraulic fracturing model with piecewise fluid rheology," *Gulf Rocks 2004, Proceedings of the 6th North America Rock Mechanics Symposium (NARMS)*, Houston, Texas, United States, June 5--9, 2004.

[8] J. Geertsma and R. Haafkens. "A comparison of the theories for predicting width and extent of vertical hydraulically induced fractures," *Journal of Energy Resources Technology*, **101**, pp. 8-19, 1979.

[9] S. Khristianovic and Y. Zheltov. "Formation of vertical fractures by means of highly viscous fluids," Vol. 2, *Proceedings of the 4th World Petroleum Congress*, Rome, Italy, June 6--15, 1955.

[10] J. Geertsma and F. De Klerk. "A rapid method of predicting width and extent of hydraulically induced fractures," *Journal of Petroleum Technology*, **21**, pp. 1571-1581, 1969.

[11] H. Abe, T. Mura and L. Keer. "Growth rate of a penny-shaped crack in hydraulic fracturing of rocks," *Journal of Geophysical Research*, **81**, pp. 5335-5340, 1976.

[12] K. Nolte and M. Smith, "Interpretation of fracturing pressures," *Journal of Petroleum Technology*, **33**, pp. 1767-1775, 1981.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 7

Combined Heat and Power (CHP) as a possible method for reduction of the CO_2 Footprint of Greenhouse Horticulture

热电联产作为减少温室园艺二氧化碳足迹的一种可能方法

Peter Vermeulen

Wageningen UR Greenhouse Horticulture, PO box 20, 2665 ZG Bleiswijk, The Netherlands

Peter.vermeulen@wur.nl

Accepted for publication on 6th August 2014

Abstract - **In recent years, the horticultural sector has been confronted with questions about the carbon footprint of its products. In greenhouse cultures, energy consumption is the main component of the CO_2 emission. To save energy, many Dutch greenhouse companies use CHP to heat their greenhouses. These growers may sell the superfluous electricity produced by the CHP to the national grid, thereby generating two products; the horticultural product, e.g. a tomato, and the electricity. The CO_2 emission of the electricity production should be deducted from the total CO_2 production of the CHP, in order to calculate the CO_2 emission that should be assigned to the production of the crop.**

To investigate the carbon footprint of a greenhouse production system to cases are compared: with and without a CHP system to heat the greenhouse. An example for grown tomatoes in The Netherlands is worked out. It shows the specific input factors and their impact on the CO_2 footprint. The functional unit used is kg CO_2 per 1000 kg product, and the system boundary is from seedling production until the delivery of product at the distribution centre of wholesalers or supermarkets.

The CO_2 footprint of the tomato crop grown without cogeneration is 55% higher than that of the crop grown without cogeneration and more than triple that of the conventional crop grown with CHP. The use of CHP is a way to reduce the CO_2 footprint for tomato growers.

Keywords - **CHP, sustainable greenhouse, greenhouse horticulture, carbon footprint, electricity production**

I. INTRODUCTION

Global heating as a result of greenhouse gasses (GHG) is a hot topic. The environmental impact of the modern horticulture sector is the subject of an increasing interest to the community. Wholesalers, supermarkets and consumer organisations therefore want insight into the GHG emission of their products, for both organic and conventional cropping methods. They plan to show the CO_2 footprint on their products, as an indicator of the impact on global heating by the production of their products. As a result, Carbon Trust, the UK Department for Environment Food and Rural Affairs (DEFRA) and the British Standards Institution (BSI) have developed a protocol for the calculations of the CO_2 footprint, the so-called PAS 2050 [2] [3]. This protocol is based on the methodology of the life cycle assessment (LCA) of the International Reference Life Cycle Data System [7]. In 2008, the Dutch Horticultural Board and the Ministry of Agriculture, Nature and Food Quality decided to start a pilot project to build a model to calculate the CO_2 footprint, so the sector can anticipate the answer to the potential question "What is the CO2 footprint of Dutch greenhouse production?" [1]. This model may be used by the members of the Dutch Horticultural Board to calculate the CO_2 footprint of their own production plant and is able to calculate the effects of changes in the production method (http://www.tuinbouw.nl/artikel/co2-footprint-berekenen).

During this study, it became clear that the use of cogeneration for the production of heat and electricity reduced CO_2 emissions, and consequently, the CO_2 footprint. Growers use co-generation to save costs and energy. In Dutch greenhouse horticulture in 2011, combined heat and power (CHP) systems which generated approximately 3000 MW of electric power were installed in a total area of 10,300 ha. Their annual electricity production is about 11,8 TWh. This electricity is partly used for artificial lighting, but most is sold to the national grid. The heat generated is used for heating the greenhouses. This decentralised cogeneration of electricity at greenhouses has benefits compared with central electricity production at normal power stations, where most of the heat is cooled

and thus wasted. Organic crop production has to compete with this modern way of cropping with the use of CHP. This article calculates the CO_2 footprint of organic and conventional tomato cropping systems. Examples of the allocation methods for CHP are described and the impact on the CO2 footprint is shown.

II. MATERIALS AND METHODS

The CO_2 footprint of a greenhouse plant with and without combined heat and power generation (CHP) were compared. An overview of data for a tomato crop produced between mid-December and late November was used. Data were obtained from "Kwantitatieve Informatie voor de Glastuinbouw 2012-2013" (Quantitative information on greenhouse horticulture), a report that frequently contains overviews of the actual inputs, cost and yields for the main crops of the Dutch greenhouse horticulture [10]. Cogeneration is used to save energy, by avoiding energy waste, especially heat, at the central electricity plants. The relevant data are shown in Table 1.

TABLE 1. INPUT DATA OF TOMATO CROP PRODUCTION IN THE NETHERLANDS, 2013.

Input	Unit	Tomato	Tomato CHP
Production	$kg \cdot m^{-2} \cdot year^{-1}$	65.5	65.5
Electric power CHP	$MW \cdot ha^{-1}$	0.0	0.6
Cogeneration	$hours \cdot year^{-1}$	0	3560
Natural gas boiler	$m^3 \cdot m^{-2} \cdot year^{-1}$	39.8	13.1
Natural gas CHP	$m^3 \cdot m^{-2} \cdot year^{-1}$	0	58.5
Electricity	$kWh \cdot m^{-2} \cdot year^{-1}$	10.00	10.00
Electricity production	$kWh \cdot m^{-2} \cdot year^{-1}$	0.00	213.00
PE/PVC/PS	$kg \cdot ha^{-1} \cdot year^{-1}$	940	940
Pesticides	$kg \cdot ha^{-1} \cdot year^{-1}$	8.5	8.5
K_2O	$kg \cdot ha^{-1} \cdot year^{-1}$	1593	1593
N	$kg \cdot ha^{-1} \cdot year^{-1}$	1593	1593
P_2O_5	$kg \cdot ha^{-1} \cdot year^{-1}$	360	360

In the situation with the CHP system, the grower produced two different products; tomatoes and electricity. For assigning CO_2 emission from a central source to multiple objectives, three ranked allocation methods can be distinguished [2] [3]:

1. System reduction. The CHP production process was broken down into sub-processes: the electricity production and the heat production, and the allocation was based on energetic output. In the case of 40% electric and 50% thermal return of power, 1 m^3 natural gas (31.65 $MJ \cdot m^{-3}$) produced 3.52 kWh (31.65/3.6*40%) electricity (Eq. 1). With a total return of 90%, 1 kWh of electricity was produced with (1/3.52) 0.284 m^3 of gas (Eq. 2). In practice, the electric return varied between 38% and 42% and the thermal return between 50% and 55%. So the CO_2 emission of the electricity was based on (40% / (40% + 50%) * 0.284) = 0.126 m^3 natural gas per kWh (Eq. 3).

$$P_e = E * C * S_e \quad (1)$$

$$G_t = 1/P_e \quad (2)$$

$$G_e = (S_e / (S_e + S_h) * P_e \quad (3)$$

P_e= Electricity production kWh
E = Energy value natural gas (heat) $MJ \cdot m^{-3}$
C= Conversion factor $kWh \cdot MJ^{-1}$
S_e= Share electricity output %
S_h= Share heat output %
G_t= Total gas input CHP to produce 1 kWh electricity $m^3 \cdot kWh^{-1}$
G_e= Share electricity of gas input CHP $m^3 \cdot kWh^{-1}$

In horticulture, the CO_2 produced was also used in the crop production process. The electricity produced by the CHP in the greenhouse plant was used outside the greenhouse system and had an impact on the national electricity production. Because the electricity sold to the national grid is not recognized as a reduction in CO2 output by this allocation method, makes the system reduction allocation method a poor choice.

2. System expansion. This method is based on expanding the system to include the impact of displaced products. In the case of cogeneration, the electricity that would have been produced by the national grid (i.e. the avoided electricity) was displaced by the electricity that was produced by the CHP system and sold back to the grid (i.e. the replacement electricity). This allocation method was useful in the cogeneration cropping case. The system included the production of tomatoes and the production of replacement electricity. The emission of the replacement electricity was deducted from the total emission of the tomato crop and electricity production at the greenhouse plant, to calculate the emission level of the tomatoes.

3. Economic allocation. This allocation method was based on the economic return of the electricity and the crop. If, for example, in a tomato crop, the yearly returns are €50.00 per m^2 and the electricity returns are €12.50 per m^2, the share of the electricity in the gas consumption of the CHP will be 12.5/(50+12.5) = 20%. If you need 0.284 m^3 gas to produce 1 kWh, the electricity part will be 0.0568 (20% * 0.284) m^3. So for each kWh electricity supplied to the national grid, the CO_2 emission of the CHP can be reduced by the emission of 0.0568 m^3 gas. This method is very unstable and will give different CO_2 footprints throughout and over the years with a comparable input of energy. Because system expansion can be used, PAS 2050 doesn't allow use of the economic allocation method.

TABLE 2. CO_2 EMISSION OF ELECTRICITY PRODUCTION IN THE NETHERLANDS. (Based on [6] [8] [9]).

Electricity source	kg $CO_2 \cdot kWh^{-1}$ excl. pre combustion
Nuclear	0
Natural gas average	450
Oil	660
Coal	870
Import in Holland 2006	586
Production average Holland 2006	543

Looking at the replacement electricity production by CHP, the time of production is important. In The Netherlands, the electricity source is different at different times of the day and on different days of the week. There is a base load of electricity production that is supplied by long-lasting power plants such as those fuelled using coal or nuclear power. However, the daily fluctuation of electricity consumption is supplied mainly by gas combustion power plants. All these production methods have their own CO_2 emissions (Table 2).

In the case of tomato, the CHP is used for two purposes: 1) production of heat and CO_2 for crop production and 2) electricity as a co-product not used for the production of tomatoes. The electricity produced is sold to the national electricity grid. The electricity market in The Netherlands is divided into two main parts: base and peak hours. The peak hours Monday to Friday from 07:00 to 23:00, the hours with the highest electricity consumption. The base hours are from 23:01 to 06:59 weekdays and the 48 hours of the weekend. The

peak hours have a high rate paid and the base hours have a low rate paid. Because the CO_2 demand by the (tomato) crop is also during the day, most of the growers use the CHP during daytime hours, both during the week peak time hours and the weekend base time hours with the low rate paid. The heat produced is used in the greenhouse directly or stored for the night in heat water storage tanks, except for the summer period when a portion of the heat cannot be used because the heat water storage tank is to small and is wasted.

Back to the question 'what is the amount of avoided electricity?' This question was answered by a panel of experts. Participants were a grower with a CHP, a PhD researcher on the energy market, a CHP specialist, an horticultural economist, a seller of electricity and two energy production specialists. The panel concluded that in The Netherlands, electricity delivered in the peak hours reduced electricity produced by gas-combusted electricity plants, and in the base hours, that produced by coal-combusted plants. In this case, it was simplified by calculating with 5/7 by gas- and 2/7 by coal-produced electricity, based on the number of days with and without peak hours, respectively. The so calculated avoided CO_2 emission was offset against the CO_2 emission of the gas used by the CHP. In situations where the amount of electricity that is delivered during peak and base hours is known, the real distribution can be used.

In the tomato crops without CHP, the allocation will be simply that all the CO_2 emissions will be due to the tomato production.

The CO_2 footprint looks at the effect on the GHG of all materials used during the whole production cycle. For all cases, the emission will be calculated for 1000 kg tomatoes. The system boundary of the life cycle assessment started with seedling production and the growth of the young plants, included their transport to the greenhouse and the fruit production at the greenhouse, and ended with the transport of the fruit to the gate of the distribution centre of the wholesaler or supermarket.

The main materials used during the seedling, young plant and fruit production periods were energy (gas and electricity), fertilizers, pesticides, plastics, rock wool, peat, etc. In these cases, the emission of the seedling and young plant production and transport is estimated at 10% of the emission for the fruit production.

An inventory in 2008 at a new tomato production greenhouse (11) gave the amount of materials used for greenhouse construction as shown in Table 3, with the average annual depreciation and percentage of recyclable materials at the end of its lifetime. PAS 2050 excluded the emissions of the production of these capital goods.

TABLE 3. MATERIALS USED IN GREENHOUSE CONSTRUCTION (TON·HA-1), THE AVERAGE DEPRECIATION (%) PER YEAR AND AMOUNT OF MATERIAL THAT MAY BE RECYCLED (%). (11)

	ton·ha^{-1}	Average % depreciation	% recyclable
Concrete	109	7.0 %	90 %
Aluminium	37	8.1 %	90 %
Glass	119	7.0 %	90 %
Steel	196	8.2 %	90 %

III. RESULTS

The CO_2 emission of a tomato crop are compared in two cases: without and with the use of a CHP for heating of the greenhouse. The results are shown in Table 4 and Fig. 1.

TABLE 4. RESULTS OF CO_2 FOOTPRINT THE CALCULATIONS (KG CO_2 EQ·TON^{-1})

	Tomato	Tomato with CHP
Young plants	130	70
Gas boiler	1136	374
CHP[1)]	0	165
Electricity	98	98
Fertilizer	43	43
Materials	18	18
Transport	8	8
Total	1430	775

[1)] Gas CHP – electricity to national grid.

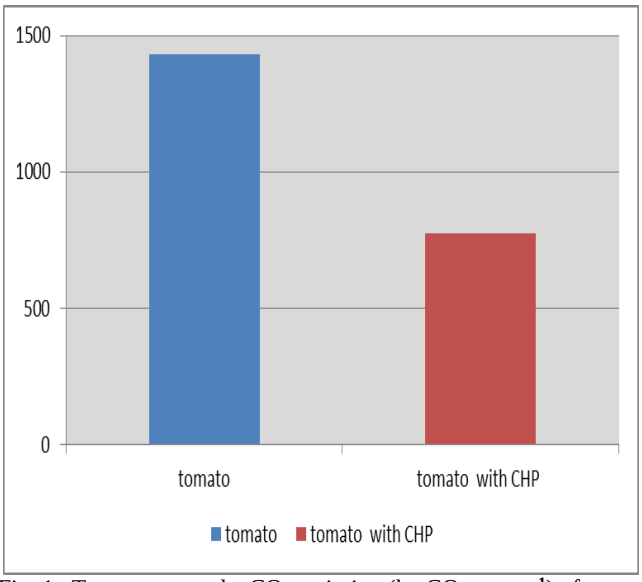


Fig. 1, Tomato crop: the CO_2 emission (kg CO_2 eq·ton^{-1}) of a crop with and without heating using a CHP system.

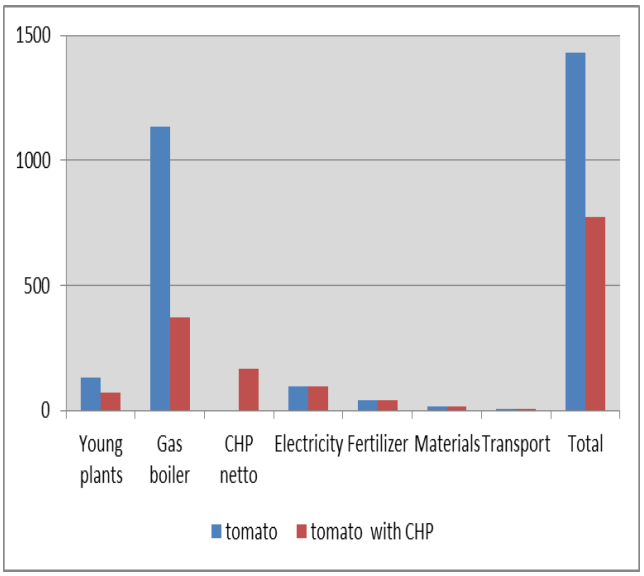


Fig. 2, Tomato crop: total and components of the CO_2 emission (kg CO_2 eq·ton^{-1}) of a crop with and without heating using a CHP system. (CHP netto = Gas CHP – electricity to national grid.)

The use of a CHP system lowers the CO_2 footprint of the crop by about 55%, due to the avoided production of electricity by power plants, for the tomato crop. So, the use of cogeneration has a positive impact on reducing the CO_2 emission of the community. The consumption of gas with CHP will be almost 50% higher than without CHP, due to the production of electricity for the national grid. However, because heat and CO_2 are used in the greenhouse production process, cogeneration results in an overall energy savings by avoiding electricity production in a central electrical production plant that generally wastes the generated heat. The final impact depends on the kind of electricity plant that the CHP-produced electricity replaces. Consequently, a greenhouse grower may decide to use cogeneration to lower his CO_2 footprint.

As shown in Fig. 2, the gas consumption is the greatest CO_2 emission component of greenhouse tomato production; without CHP in the tomato growing system it is 79%; and with CHP, it is 70%.

Energy savings and the use of green energy are the major components in the reduction of the CO_2 footprint of protected horticulture. The other factors that can be considered for a further reduction of the CO_2 eq. emission of a tomato are the use of fertilizers and the transport of the product to the distribution centre.

IV. DISCUSSION AND CONCLUSIONS

Greenhouse horticulture has to innovate their production system to lower the CO_2 footprint. Actually the greenhouse horticulture is quickly adapting new technologies, such as CHP, for its use. Other new energy systems already have been developed or will be developed, such as:

- Heat from CHP delivery by greenhouse growers to other companies and/or non-greenhouse partners, such as schools, swimming pools, etc.;
- Heat or CO_2 delivery by electricity or industrial plants to greenhouses;
- Use of geothermal heat; August 2014 10 wells in the Netherlands;
- Bio energy;
- Fermentation.

Greenhouse growers can choose from these options and look at the effects of the chosen option(s) on the CO_2 footprint for their production system.

Growers have to become aware that the community and wholesalers want insight into the production method of their suppliers and the impact of the production method on global warming and environmental burdens. The CO_2 footprint is said to be the indicator that wholesalers and supermarkets will use, explaining only part of the overall environmental impact of the production method used. Abiotic resource depletion, human, aquatic and terrestrial toxicity, acidification, eutrophication, deduction of stratospheric ozone depletion and photo-oxidants formation, etc., which are the other (sub)indicators of the LCA methodology, are not considered. For comparison studies, however, these other indicators should be considered to avoid any misinterpretation of the environmental effects of a specific growing system according to the International Reference Life Cycle Data System Handbook (ILCD, 2008).

There are a lot of databases with elements of the CO_2 eq emissions of materials that use different emission figures on the same materials. Using these different figures can have a high impact on the level of the CO_2 footprint. A widely accepted database which explains local differences in data will be necessary. In this study, the database Ecoinvent was used. (www.ecoinvent.ch) [4] [5]

The CHP case is one of many possible ways to use cogeneration in greenhouse horticulture. The potential CO_2 emission reduction depends on many specific factors. In this study, the most important factors were: electric and heat return of the CHP, number of hours with cogeneration, type of electricity production avoided (i.e. coal or nuclear vs. gas), amount of generated power in relation to the area of the greenhouse, and heat and CO_2 demand of the greenhouse. This CO_2 footprint method is an easy tool for growers to use to calculate the CO_2 emission of their own crop and production method. In this case study, the use of the CHP is based on the heat and CO_2 demand of the crop, to ensure the least possible heat wastage at the greenhouse plant. To achieve the illustrated reduction of CO_2 emissions using CHP, the investment and extra gas consumption have to be recouped by the returns from the electricity sales. In 2008, which had high prices for both base and peak time electricity delivery, growers let the CHP run extra hours to generate extra income. In 2010-2014, which had low electricity prices, growers stopped this extra use of cogeneration because the extra gas consumption would not have been recouped by the sale of electricity. Therefore, to realise a reduction of CO_2 emissions with cogeneration in horticulture, there needs to be a stable electricity market with fair prices.

REFERENCES

[1] Blonk, J, Ponsioen, T and Scholten, J. *Carbon footprint berekening voor Agro ketens. Een nadere uitwerking van PAS 2050 [Carbon footprint calculation in agricultural chains. A further detail development of Pas 2050]* (in Dutch). Productschap Tuinbouw, Zoetermeer. The Netherlands 2009

[2] BSI. *Guide to PAS 2050. How to assess the carbon footprint of goods and services.* British Standards Institution (BSI), UK. 2008a.

[3] BSI. PAS 2050:2008. *Specification for the assessment of the life cycle greenhouse gas emissions of goods and services.* British Standards Institution (BSI), UK. 2008b.

[4] Dones et al. *Ecoinvent data v2.0.* Swiss Centre for Life Cycle Inventories. 2007.

[5] Frischknecht, R., Jungbluth, N., Althaus, H.-J., Bauer, C., Doka, G., Dones, R., Hischier, R., Hellweg, S., Humbert, S., Köllner, T., Loerincik, Y., Margni, M. and Nemecek, T. *Implementation of life cycle impact assessment methods. Ecoinvent report No. 3, v2.0.* Swiss Centre for Life Cycle Inventories, Dübendorf, Switzerland. 2007.

[6] Groot, M.I. and van de Vreede, G.J. *Achtergrondgegevens stroometikettering 2006 [Background data power labels]* (in Dutch). CE, Delft, The Netherlands. 2007.

[7] Hienrich, A.B. *International reference life cycle data system handbook. Supporting business and public authorities towards sustainable production and consumption.* The International Journal of Life Cycle Assessment 15:524-525 2010.

[8] Seebregts, A.J. and Volkers, C.H. *Monitoring Nederlandse elektriciteitscentrales [Monitoring Dutch power stations] 2000-2004 (in Dutch).* ECN Beleidsstudies, Petten, The Netherlands. 2005.

[9] Sevenster, M.N., Croezen, H.J., Blom, M. and Rooijers F. 2007. Nieuwe elektriciteitscentrale in Nederland [New power plants in the Netherlands] (in Dutch). CE, Delft, The Netherlands.

[10] Vermeulen, P.C.M. *Kwantitatieve Informatie voor de Glastuinbouw [Quantitative information on greenhouse horticulture] 2012 - 2013* (in Dutch). Wageningen UR Glastuinbouw, Wageningen, The Netherlands. pp.107, G52 and G53, 2013.

[11] Vermeulen, P.C.M. *Materiaalinzet kas en teeltsysteem versie 10-03-09 [Material use greenhouses and growing systems]* Internal database, Wageningen UR Glastuinbouw, Wageningen, The Netherlands.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 8

The new oil to natural gas price ratio paradigm
石油对天然气价格比的新范式

Douglas B. Reynolds

School of Management, University of Alaska Fairbanks, PO box 756080, Fairbanks, Alaska, 99775-6080, U.S.A.

DBReynolds@Alaska.Edu

Accepted for publication on 6th August 2014

Abstract - **Before 2005, the ratio of the U.S. price of oil (dollars per barrel) to that of natural gas (dollars per thousand cubic feet (MCF)) stayed close to seven to one (i.e. close to 1 to 1.2 in dollars per giga joule (GJ) equivalent). The ratio was dependent on the fact that U.S. electric utilities were able to switch between oil derived fuel and natural gas for peak and intermediate electric power generation needs around that ratio. That means technology drove the oil to natural gas price ratio relationship. However, now that most generators use cheap natural gas in America and no longer switch to expensive oil, a new, volatile oil to natural gas price ratio paradigm has emerged in the U.S. where no one technology drives the relationship. However, that could change. For example in 2007, there was a massive switch from gasoline-fueled vehicles to natural gas-fueled vehicles in Utah, U.S.A. The costs and benefits of this switch showed a price ratio between oil and natural gas that could approach 22 to 1 (3.7 to 1 in dollars per GJ). This suggests that the major natural gas markets in Europe and America may be able to drive the world oil market to some degree.**

Keywords – **Natural gas, Energy prices, Oil Substitution**

I. INTRODUCTION

As the price of oil has risen, there is concern for how the world's economy will be able to operate with less oil. Ideas for oil substitution range from using coal-to-liquids to gas-to-liquids to electric cars. However, since natural gas is amenable to being used in internal combustion engines right now, then a much simpler idea is to use natural gas directly to replace oil derived fuels in automobiles. In the United States the price of natural gas has been low for several years and yet not many U.S. regions have used compressed natural gas (CNG) automobiles as a significant portion of their fuel use strategies. This analysis suggests that the costs of using natural gas may be too high yet to be able to make such a switch happen.

Looking at the U.S. natural gas market, it is clear that for many years, the price of natural gas in the United States and the price of oil were linked by power generation technology. Modern electric power plants are able to switch from residual fuel oil (or condensate fuel oil) to natural gas and back again as needed. Anytime the price of natural gas per thousand cubic feet (roughly the same price per giga joule), dipped below a ratio of 1/6th or 1/8th the price of oil per barrel (about 1 to 1.2 times the price of oil, in terms of dollars per GJ), power generation switched to natural gas. Whenever the price of natural gas increased above that rough ratio "rule of thumb," then power generation switched to oil. See [1], [2], [3], [4] and [5] and see Figure 1 for a history. Switching power generation feedstock back and forth meant that the price ratio stayed within the same bounds because any time the price of natural gas went above or below the ratio for any length of time, the demand for natural gas quickly altered to compensate. The oil market was the deeper, more stable energy market. If natural gas prices went too high, demand for natural gas quickly plummeted, if natural gas prices went too low, demand for natural gas quickly increased. The oil dog wagged the natural gas tail.

Fig.1, The roughly 6 to 1 or 8 to 1 price ratio of natural gas to oil. U.S. market, from EIA yearly data to 2002 before natural gas volatility occurred.

However, now that natural gas supplies in the United Sates have increased well above power and space heating demand, with very low natural gas prices, and new supplies look to be large, there is no more reason to switch to oil for electric power generation in the U.S. Although, as with the rest of the world, the U.S. still looks to use some coal fired generation in place

of oil and natural gas as well. Therefore, oil does not look to be a factor in world electricity production. This is especially true since oil prices are forecast to remain high, and natural gas prices are forecast to remain low for several years if not decades, and there looks to be plenty of natural gas for all projected space heating and electric needs in the U.S. and possibly around the world due to shale gas. We could also look at Hotelling [6] rent as a cause for price rises for oil and natural gas, as Gordon [7] and Holland [8] suggest, nevertheless, Reynolds [9] shows that the Hotelling rule cannot be used.

Although, the price of natural gas could go back up if demand picks up and the price of oil could go down if market forces change, nevertheless, as of 2014 the economics of using natural gas are so much better than the economics of using oil for power generation that there does not look to be any more switching in the U.S. and very little in the rest of the world. However, if supplies of natural gas continue to increase substantially, due to shale natural gas—as the Potential Gas Committee [10] and the EIA [11] show—and if oil prices go up—as UKERC [12], Fantazzini [13], Kumhof et. al [14], Lutz et. al [15], Mirchi et. al [16], Hallock et. al [17] and Reynolds [18] show due to "peak oil"—then a new substitution paradigm may well emerge. The technology around the natural gas switching may be with automobiles rather than electric power production.

According to Ramberg and Parsons [19], the historic cointegration relationship between oil and gas prices "may not be a very reliable predictor of future natural gas prices." Ramberg and Parsons do give one option for the future of natural gas prices which is that gas-to-liquids technology may determine the upward bound of the oil price to natural gas price relationship. However, there is a more feasible option. Natural gas and oil may in the future substitute directly for one another as a transportation fuel for automobiles seeing as natural gas looks to be a cheap energy source for a long time and has characteristics similar to a liquid fuel, although Flynn [20](2002) shows many of the drawbacks of using natural gas for an automotive fuel.

Nevertheless, Krause [21] shows that in Utah, starting in 2007, a new fuel switching paradigm took place where consumers switched from gasoline-driven cars to compressed natural gas (CNG) cars. Boyle [22] also shows that many developing countries could start using CNG vehicle technology. Typically, such conversions require that a large CNG tank be installed in the car and slight alterations done to the engine. Large cars such as SUVs and pickup trucks are amenable to conversion and there are plenty of those vehicles in the U.S.

In this paper, we will analyze a case study in Utah in order to understand what kind of rate of return consumers need in order to make the switch from oil to natural gas. In section III a simple cost estimate model for the relationship between the price of oil and the price of natural gas for automotive consumers is derived. In section IV, a competitive market analysis is done. In section V the implications for such an oil-to-natural gas price ratio paradigm are explored. If we assume natural gas supplies will be much larger than future oil supplies and will be at a reliable low cost, then the natural gas market could dictate the price of oil in the oil market. In the last section, I give concluding remarks.

II. The Utah Case

In Utah, in 2007 and 2008, a substantial number of motorists switched their transportation vehicles from using gasoline to using CNG because of the low price for natural gas and because there were relatively many utility and government sponsored natural gas fueling-stations. See Krause [23]. Flynn [24] shows that using CNG is difficult due to many factors including consumer cost markups, natural gas fuel-station availability and the perception that CNG vehicles won't work as well.

Turrentine and Sperling [25] suggest that perceived risk by the consumer will create resistance to a changeover, presumably the risk of a lack of natural gas fuel availability when needed, but also the risk of an explosion of the CNG tank itself. However, those concerns were overcome in Utah in 2008 when a surge in interest in purchasing CNG vehicles or converting to CNG vehicles occurred. Although there were many problems with fueling which were not emphasized, particularly the problem of having too few natural gas fueling-stations available, which we will look at below, nevertheless, it did work to a degree. Importantly, Aslam et. al. [26] show that CNG vehicles are comparable to gasoline vehicles in many performance characteristics. So, while some motorists in Utah simply bought automobiles built to use CNG, such as the Honda civic GX, many actually paid to have their existing gasoline automobiles converted to CNG vehicles. A close inspection of the costs and benefits for the switching option of turning an existing gasoline vehicle into a CNG vehicle shows the following economics.

The typical cost of conversion to switch from a conventional car to a compressed natural gas car was $12,000 in 2008. Assuming a typical car in Utah got 20 miles per gallon and drove 10,000 miles a year and that there was no appreciable change in mileage after CNG conversion, then a typical car used roughly 500 gallons of gasoline equivalent per year (63 GJ). The price of gasoline at the time consumer switching took place in Utah was $3.95 per gallon ($32/GJ) in 2008 with oil at about a $120 per barrel price. The city gate price of natural gas in Utah was about $8 per GJ which is equivalent to $0.87 per gallon of gasoline equivalent. Since natural gas retail stations were government run, the city gate prices were roughly the same price as the natural gas fueling-stations. Therefore the price difference of the oil based gasoline fuel compared to the natural gas fuel was $3.08 per gallon of gasoline equivalent. If you multiply that by 500 gallons, the cost savings come out to be $1,540 per year or roughly 13% of the cost of conversion. In other words consumers were saying that to pay for a $12,000 conversion, they needed a 13% return on their capital investment.

However, one of the problems that Utah had was that people were waiting in line to be able to fill up their CNG vehicles, or

they were only allowed to fill up half way, or they had to come to fill their CNG car at night. That signifies an economic constraint on their options, which is a further cost of CNG vehicles. We will go into that constraint in section IV.

III. A Switching Cost Example

What we would like to do is to first understand what the costs and benefits are of switching from oil to natural gas as a fuel source using simple engineering economic concepts. Once we have the engineering costs and benefits, we can expand later into market implications. First, consider a simple case of converting from gasoline to CNG. If the price of natural gas is low, relative to the price of liquid fuels, than people who already own gasoline or diesel cars and trucks will switch to natural gas by converting their existing automobiles into CNG vehicles. People who are considering buying new vehicles, will simply buy pre-equipped CNG vehicles, although probably both would be bought on new car markets as not all U.S. regions are amenable to using natural gas.

In the U.S., 150,000 natural gas cars have been sold, and if that number grows to one million, as expected, it still does not put a dent in the 250 million, U.S.-car fleet. Therefore, for the next ten or twenty years, millions of used cars in North America can be converted from using liquid-fuel to using natural gas. It is these used cars that will determine a switching market. There will be an increase in used car conversions if oil prices rise very high relative to natural gas and a decrease in used car conversions if oil prices are particularly low relative to natural gas, even as much of the rest of the world, and even some regions in the U.S., continue to purchase oil fueled cars due to having no other option. Such natural gas free regions could provide used cars that could be brought to the U.S. for switching. However, used, liquid-fueled cars will not be able to be sold easily in order to buy new natural gas cars, as no one will want a used oil fueled car when they can buy a new natural gas fueled car, and the used car market will become a buyers' market with discounted prices. So, the only option for millions of car owners would be technological conversion of used cars.

If the price of natural gas is high relative to the price of oil, people who already own gasoline and diesel vehicles will refrain from switching their car to use CNG. This means that there may emerge a switching relationship between oil and natural gas. To understand the potential relationship when natural gas starts from a high price and moves low, we need to do an engineering analysis of the costs of a switch from a gasoline vehicle to a CNG vehicle.

Assume consumers must pay a capital cost in order to convert their automobile from gasoline to CNG similar to what happened in Utah, and assume the natural gas fuel stations will emerge in only a few regions near dense cities or near well-developed natural gas infrastructure. Since the U.S. has a very deep natural gas market, all that is needed is for some U.S. cities or regions to develop CNG stations in order for the switching to occur, i.e. it is not necessary for a national CNG grid to have some conversions occur. Cars can also be bought to run only on natural gas and probably at a similar cost as gasoline driven cars; however at the margin, where switching will occur, it is the people that convert from gasoline to natural gas that could drive the price ratio of oil to natural gas.

Assume consumers who pay for CNG conversions consider the car itself is a sunk cost, then every year the consumer must be able to pay off the conversion investment with the following payment:

$$yearly\ cost = C \cdot r$$

where C is the cost of converting a gasoline vehicle to a CNG vehicle, and r is the interest rate paid by the consumer for the loan to pay for the conversion.

We also need to understand the advantage in fuel savings. Let:

$$G_r = \text{the retail price of gasoline (\$ per gallon)}$$
$$N_r = \text{the retail price of natural gas (in \$ per MCF)}$$

where \$ per MCF is roughly equal to \$ per GJ.

In order to make the above retail prices comparable, a conversion factor is needed. Assume N_r is dollars per MCF (\$ per GJ) and G_r is dollars per gallon of gasoline equivalent, the conversion factor is:

1 gallon of gasoline = 0.125 MCF

or

The gasoline equivalent MCF = $8G_r$.

The consumer pays for the capital cost of conversion by saving money on driving, which means the difference in fuel cost multiplied by fuel use per year has to pay the conversion costs or:

$$(8G_r - N_r)F = C \cdot r \qquad (1)$$

where F = fuel use per year in MCF.

Right now the average U.S. light vehicle driver drives 10,500 miles per car per year, and that of a light truck drives 12,400 miles per year, but may drive more or less depending on their sensitivity to fuel prices (EPA [27]). The EPA assumed the average passenger vehicle drove 12,000 per year in 2005, so by now, it may be 11,000 due to higher gasoline prices. The natural gas fuel use for a typical CNG car, assuming similar characteristics for a CNG vehicle as for a gasoline fueled vehicle, is:

$$\text{Fuel use per year } (F) = \left({}^{M_i}/_{8MPG}\right)$$

M_i = miles driven per year

MPG = miles per gallon of gasoline equivalent.

Therefore the consumer needs the price difference to be the following:

(Price difference) x (fuel use) = investment cost

$$(8G_r - N_r)\left({}^{M_i}/_{8MPG}\right) = C \cdot r \qquad (2)$$

(note: 25 *MPG* = 9.4 liters per 100 km; or 10.63 km per litter).

The retail natural gas price in the United State is usually just above a city gate price, particularly for a fueling-station buying the natural gas in bulk. The actual cost for the fueling-station will be explained more in section IV. For now, assume the retail price of natural gas for a residence is roughly the same as for a fueling-station. Therefore the natural gas price to the retailer and to the customer is assumed to be roughly 50% higher than the Henry Hub price in North America, due to taxes, costs of distribution and returns:

$$N_r = 1.5 \cdot P_{HH}$$

where P_{HH} = the Henry Hub natural gas price ($/MCF).

By multiplying the gasoline price by 46 gallons (the GJ equivalent value of gasoline compared to the GJ equivalent value of crude oil per barrel rather than the normal 42 gallons per barrel), we get a per barrel of oil equivalent price of gasoline. Including taxes and the cost of refining, which all increase as the price of oil increases, due to the entropy subsidy problem (Georgescu-Roegen [28]), i.e. the problem where an initial cost estimate of energy technology inflates as the cost of energy used to make or run that technology increases. Therefore, the retail gasoline price is roughly twice that of the crude oil price or:

$$46 \cdot G_r = 2 \cdot P_{WTI}$$

where P_{WTI} = the West Texas Intermediate price of oil per barrel.

In the future, most of the natural gas in the United States will be produced from shale natural gas, although that shale natural gas is highly capital intensive relative to conventional natural gas. Currently, U.S. natural gas prices are $4.50 per thousand cubic feet (MCF) (€3.5/GJ) partly because many shale oil plays are producing both oil and natural gas simultaneously and producers are being forced to dump their natural gas on the market for whatever price they can get seeing as shale oil production provides all the profits and pays for all the costs of drilling. The rough cost of obtaining shale natural gas, where it is developed from scratch and where there is no associated oil, is about $7 to $9 per MCF (€6.5/GJ) in 2014 dollars depending on the shale natural gas play, technology and the price of metals used in the natural gas drilling capital machinery. See Rogers [29], Berman and Pittinger [30] and Mazur [31]. For now assume the future cost of shale natural gas developed in the U.S. from scratch will be $7 per MCF ($7 per GJ), and assume that eventually that greenfield cost will dictate the U.S. Henry Hub price of natural gas. In the long run, the price of shale natural gas will increase due to those real capital costs but is not expected to increase due to supply constraints. Therefore the long run price of shale natural gas should decrease due to technology and increase due to steel and metal costs and thus can be used as a real inflation index for capital and technology costs in general. That index would be:

$$I = {}^{P_{HH}}/_{A}$$

where P_{HH} is a long run average Henry Hub price, and where *A* is the average base year expected cost in dollars per GJ of extracting shale natural gas from a greenfield and I is a unit-less inflation index that should equal unity in the long run.

Note, the inflation index is not expected to change over the long run, but it is a convenient way to integrate costs of capital in both CNG conversion and costs of extraction. The inflation index can be multiplied with the capital cost of converting a car to CNG to get an indexed real capital cost of automotive conversions, which can be added to Equation 2:

$$(8 \cdot G_r - N_r)\left[{}^{M_i}/_{8MPG}\right] = C \cdot r \cdot I$$

Since G_r, N_r and I have the P_{WTI} and P_{HH} variables subsumed, we can solve for the P_{WTI} /P_{HH} ratio, which gives us an oil to natural gas price ratio based on automotive fuel use of the two options (see appendix 1):

$${}^{P_{WTI}}/_{P_{HH}} = 4.31 + \left[{}^{23(C \cdot r \cdot MPG)}/_{(A \cdot M_i)}\right] \quad (3)$$

Equation 3 is the future oil to natural gas price ratio paradigm. In Section II, it was seen that consumers look to require a rate of return of about 13% in order to invest in conversion capital. Also, retail borrowers often pay 13% for the cost of borrowing for automotive purchases, although we might expect a range of interest rates depending on market conditions. Assume the base cost of shale natural gas, (*A*), is $7 per MCF ($7 per GJ), the capital cost, (*C*), required to change a car from gasoline to CNG is $12,000. The typical car in America drives, 11,000 miles, (M_i), a year, and in the U.S., car mileage, (*MPG*), is inching up to 25 mpg. In that case, using equation 3, the price ratio of oil, in dollars per barrel, to natural gas, in dollars per MCF, will be 16 to 1 (see appendix 1). That means if natural gas at the U.S. Henry Hub is $7/MCF, then the P_{WTI} will have to be at $172 per barrel before switching occurs. However, right now oil is priced at $100 per barrel and natural gas is at $5 in the U.S. which is a ratio of 20 to 1. Such a high ratio with such a low natural gas price suggests the U.S. should already be switching to natural gas vehicles right now, but very little switching is occurring in the U.S. Part of this may be Dixit and Pindyck's [32] concept of sunk cost externalities, although there may be other barriers.

We need to look at the cost and complexity of having natural gas fueling-stations as that looks to be the reason no one is switching. In the Utah case, the government owned and ran natural gas fueling-stations, but the Utah government could not satiate demand and there were queues. So if we want market-driven, fueling-station retailers, then we need to include the cost of the retail fueling-stations themselves in our analysis. Indeed, as the Utah case shows, there tends to be limited availability of natural gas compressor fueling-stations, undoubtedly due to the expense of building them.

IV. Non-competitive market factors

One of the problems with using natural gas as a fuel for automotive transportation is the high expense of a compressor-fueling-station in order to fill up CNG tanks. Also Flynn [33] shows that market barriers to entry and monopolistic competition create a number of high costs for CNG cars

including getting replacement parts. Therefore not all of the costs are taken account of in equation (3). Here, let us concentrate on the fueling-station costs, later we can look at geographical, monopolistic-competition aspects. Assume the high compressor-fueling-station costs preclude anyone from having a home compressor, but that a simple fueling-station is very expensive.

Consider again an engineering economic analysis for understanding the cost of a natural gas fueling-station market. First, we can add a fueling-station cost to equation 3. Right now a typical compressor-fueling-station would cost as much as $1 million and would operate 2 hoses. Typically, a customer can fill up a CNG tank in about 3 minutes, however, most customers may need additional time to organize themselves, put their credit card in and clean their windows. Therefore it could take 6 minutes per fill-up of an 8 gallon of gasoline equivalent CNG tank, that is 1 million British thermal units (mmBtu) or 1 GJ. If we assume that such a high cost facility requires about 35% of initial cost per year to pay for maintenance, operation, taxes and a return on capital investment, then that means typically there needs to be $350,000 per year of revenue to cover the yearly cost of a $1 million compressor-station, or roughly $1,000 per day.

If we assume most cars will be filling up during rush hour, 2 hours in the morning and 2 hours in the evening, and maybe one more hour in addition, then the fueling-station must achieve the requisite $1,000 per day in revenue, above the cost of natural gas, in 5 hours with 2 hoses, or in 600 combined minutes per day. The 600 minutes must be divided by the 6 minutes needed to fill up, which gives the number of fill ups per day at 100 that must pay for the fueling-station's capital, taxes and operational costs of $1000 per day. Thus each fill up must charge an extra $10 per fill up, i.e. $10 per GJ or roughly $1.25 per gallon of gasoline equivalent, just to pay for the compressor station. If competition is needed to make sure that one station does not have a geographical monopoly power, then you basically need two fueling-stations just to maintain competition. In that case, you would have half as many fill-ups with the same costs, which comes to $2.50 per gallon of gasoline equivalent or $20 per fill-up ($20 per GJ) just to pay for a competitive fueling-stations. Even then you could still have geographic monopoly power and barriers to entry that could well keep the price of natural gas above $3 or $4 per gallon of gasoline equivalent. This would make natural gas unusable under most assumptions until oil is well above the $200 per barrel range and even if natural gas is below $5 per MCF ($5 per GJ).

However, let us assume that compressor technology, or liquefied natural gas (LNG) technology, can be used to reduce costs of a fill-up, and let us assume that marketing strategies can be used to spread out the time when people fill-up their cars, then maybe the cost of a fueling-station can be reduced to a more reasonable $15 per fill-up ($1.85 per gallon of gasoline equivalent or $15 per GJ). Assume this creates enough spare fueling capacity to keep competition available. Using an inflation index similar to *I* above, let *b* be a technology of fill-up factor, where *b* = $15 per GJ, which is the cost of a fill up. Let us add the cost of a fill up to the equation 3.

First:

$$N_r = I \cdot b + 1.5 \cdot P_{HH}$$

So equation 2 becomes:

$$\left(8G_r - {}^{P_{HH} \cdot b}/_{A} - 1.5P_{HH}\right)\left({}^{M_i}/_{8MPG}\right) = C \cdot r \cdot I$$

And solving for P_{WTI} / P_{HH}, see appendix 2, we get:

$${}^{P_{WTI}}/_{P_{HH}} = 4.31 + \left({}^{23(C \cdot r \cdot MPG)}/_{(A \cdot M_i)}\right) + {}^{2.90 \cdot b}/_{A} \quad (4)$$

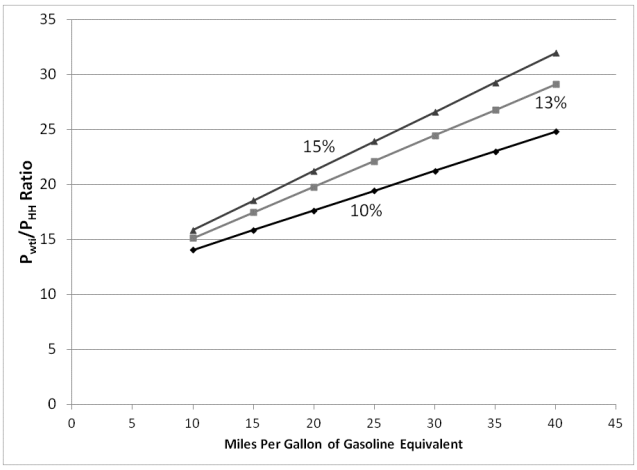


Fig 2, Oil price to natural gas price ratio, driving 11,000 miles per year and shale costs of $7 (various interest rates).

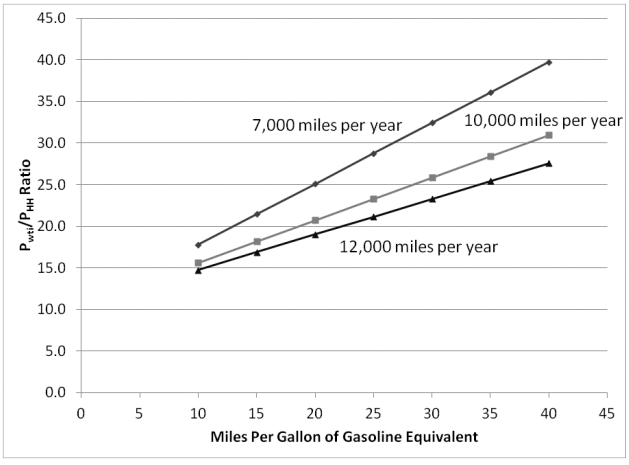


Fig 3, Oil price to natural gas price ratio with a 13% interest rate and shale costs of $7 (various miles driven per year).

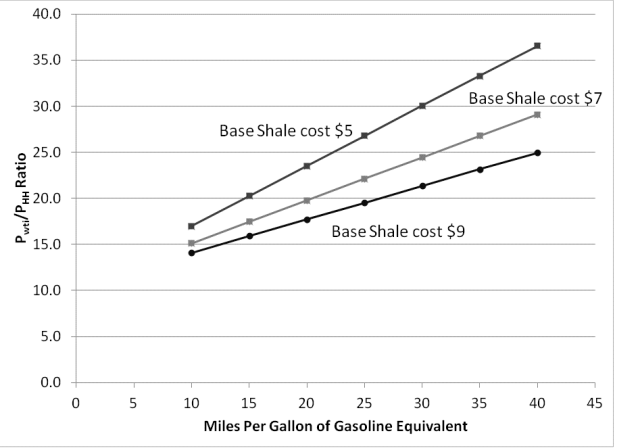


Figure 4. Oil price to natural gas price ratio with a 13% interest rate and 11,000 miles per year (various base shale costs).

Assume the base cost of shale natural gas, (*A*), is $7 per GJ, the capital cost, (*C*), required to change a car from gasoline to CNG is $12,000. The typical car in America drives, 11,000 miles, (M_i), a year, and in the U.S., car mileage, (MPG), is inching up to 25 mpg, and b = $15 per GJ. In that case, using equation 4, the price ratio of oil, in dollars per barrel, to natural gas, in dollars per MCF, will be 22 to 1. This suggests that the price of oil has to be $154 per barrel before such switching would happen. Figures 2, 3 and 4 show different scenarios for different assumptions on M_i, r and A.

V. The Natural Gas to Oil Markets

While there are many factors that could play into the automotive fuel market, such as electric cars, hybrid cars and mass transit, nevertheless, these options look to have low consumer utility. For example electric cars and hybrids look to be more expensive than liquid-fueled or natural gas cars due to the specialized materials required and mass transit typically requires two or three times the amount of passenger travel time compared to automotive travel time. So it will still be the case that in most places around the world, oil derived liquid-fueled cars will continue to be used extensively, and that as countries are able to expand natural gas options to more cities and regions, that natural gas will be able slowly to take the place of liquid-fueled cars. Therefore, we need to look closely at the economics of the switch.

Equation (4) depends on the inflation index in order to solve for P_{HH}. However, the P_{HH} variable used for the inflation index should be considered a long run average price rather than a short run volatile price. If real costs of capital increase in the long run for capital intensive shale natural gas production, then we would expect real capital costs to increase for the capital costs of converting an automobile to natural gas. Technology can cause the real costs relative to general inflation to decline, but steel and metal costs, which increased remarkably in 2007 and 2008, can cause real costs of conversions to increase relative to general inflation. The inflation index is also based on the assumption that the cost of extracting natural gas is not affected by supply constraints because natural gas supplies are considered so vast that the price of natural gas will undulate around the long run supply cost, which will be the average cost of extracting shale natural gas. Therefore this is a long run ratio.

In the future, if for any reason oil prices go above the roughly 22 to 1 ratio, everyone will begin to switch their cars to CNG vehicles, at least at the margin. If we assume that there is a deep market of existing gasoline and diesel vehicles that can be switched to CNG, then this market switch at the margin will continue to be available for many years. That means that there will be a reduction in oil demand as soon as oil prices go above the ratio and the demand for oil and the price of oil will then decline. If, though, oil prices go below the 22 to 1 ratio, then everyone who has converted their vehicles to CNG vehicles will not switch back to gasoline vehicles because the investment into CNG is a sunk cost and the price of natural gas will still be lower than oil fuels. Once the investment is made, then the consumer simply uses whichever fuel, gasoline or natural gas is cheaper per GJ and that will undoubtedly always be a natural gas option. This implies hysteresis on the part of the consumer switching, so that when oil prices go well below the roughly 22 to 1 ratio, the CNG cars will not switch back into oil.

However, another factor for the natural gas to oil price ratio is the supply of oil. Since UKERC [12] claims that the world is at peak oil and Reynolds [34] shows that shale oil may reach a peak, then if a decline in demand for oil causes the price of oil to go below the 22 to 1 ratio, the supply of oil will still be declining worldwide due to peak oil which will eventually push the price of oil back up. Considering that many underdeveloped countries will not have easy access to natural gas, but will continue to increase their automotive usage due to higher incomes, then there will always be a substantial demand increase for gasoline driven vehicles, as the world GDP increases, that will keep oil prices pushing upward. There will not be quick reaction to oil prices that go below the 22 to 1 ratio, so oil and natural gas prices can remain quite volatile, nevertheless, the overall market forces will tend to push back to the 22 to 1 ratio.

The implication of this ratio is that if natural gas is say $7 per GJ, then at a 22 to 1 ratio, oil should be $150 per barrel. If natural gas costs are higher, at say $10 per GJ in real terms, and the inflation index base price of natural gas is $10, i.e. where A = $10, then the oil to natural gas price ratio will be 17 to 1 and the price of oil will be $170. Such high oil prices would normally create speculation that oil sands, oil shale or other alternative oil technologies can create a vast and viable substitute for oil, assuming there are no supply constraints for those options such as capital, pipelines and labor, and that those substitutes force the price of oil down below such a ratio. However, many of these oil alternatives have yet to take off in substantial quantities that can take care of a significant share of the world's 80 to 90 million barrel a day thirst for liquid oil-type fuels. When oil prices rose after 2004 and the oil sands industry was primed to pick up its output, environmental problems, labor constraints and high royalties all conspired to hinder that expansion, and so oil sands bitumen supply hasn't

been nearly as abundant as originally touted. See Reynolds [35].

Assuming shale natural gas supplies are vast and the CNG switching occurs, then that implies the price of natural gas could drive the price of oil worldwide rather than the reverse. In the past, oil drove natural gas prices since natural gas supplies were limited by pipeline access. Certainly natural gas has been cheaper than oil in regional natural gas markets, where supplies could not leave the region due to bottlenecks of natural gas exports. However, natural gas was never more expensive than oil since oil could relatively easily and cheaply be transported to wherever it was needed worldwide at close to the same oil price as oil was in Texas or the Middle East.

One issue with natural gas though is being able to get enough natural gas to the various markets at one time. The North American and European natural gas markets are relatively flush with natural gas pipelines from producers to consumers (see Tussing and Tippee [36]), but much of the shale natural gas occurs in new regions that may not have enough exit pipeline capacity, and much of the CNG switching will be in far flung consumer regions where natural gas entry supply capacity is still limited. Therefore, as the price of oil increases and the demand for natural gas increases in parallel, the speed with which natural gas infrastructure develops could slow down the availability of shale natural gas and cause local natural gas prices to be too high to give consumers the incentive to switch into natural gas. Pipeline bottlenecks could still hinder the price ratio relationship. However if natural gas prices manage to stay low in a give region for a year or two, then that may instill enough confidence in consumers to make the switch to CNG vehicles.

Another issue with CNG vehicle switching is the true size of the shale natural gas supply. As of now, shale natural gas potential reserves look vast, but we also saw how oil sands reserves in Canada looked vast in the past, but did not expand to that potential. Even if the shale natural gas reserves turn out to be as large as stated, it remains to be seen if the capacity of the shale natural gas production can increase in parallel to the high reserves especially since shale reserves often have high initial production rates but are followed by very low outputs for each natural gas well. Nevertheless, the world may be looking for shale natural gas to replace oil on a grand scale, and while shale natural gas may be able to replace conventional natural gas in America, it remains to be seen if shale natural gas reserves are vast enough to replace oil in the world. Clearly, oil-sands has proven to be limited in its role as an oil replacement, the same could be true for shale natural gas.

VI. Conclusion

In the past, the ratio of the price of natural gas to the price of oil was driven by electric power feedstock needs, where natural gas could easily substitute for oil fuels as an input into generating power and back again. In the future with plentiful natural gas supplies forecast, most intermittent electric power needs will be supplied from natural gas and not oil and therefore there may not be any new electric power feedstock switching. However, there still will be a relationship between the price of oil and the price of natural gas which could be driven by the switching of automotive fuels. World shale natural gas supplies may or may not be vast depending on technology, on how many shale sweet spots there are, and on political considerations. However, assuming shale natural gas is vast, then we can analyze how switching between natural gas and oil fuels will work.

Oil derived gasoline used to be considered the only low cost and dependable fuel source for automobiles, especially when compared to the problems of using and storing natural gas for CNG vehicles. Now, CNG automobile conversion is relatively easy and natural gas supplies can be relatively cheap. This suggests that switching to CNG for vehicle fuel needs will be common place in the future. If oil prices ($/Bbl) increase above a roughly 22 to 1 ratio compared to natural gas prices ($/GJ), then many vehicle owners will switch to CNG vehicles which will keep oil prices from rising farther. If the price of oil declines below the 22 to 1 ratio, then CNG vehicles will not switch back to gasoline, but new car buyers especially in countries that do not have access to natural gas will continue to buy gasoline automobiles rather than CNG cars and increase the demand and price for oil. Therefore some of the reaction to oil and natural gas price changes will be slow.

The switching of automotive fuels, at the margin, rather like the switching of electric power generation fuels at the margin in the past, will moderate the price of oil worldwide assuming that the supply of shale natural gas is as vast as speculated. That means if natural gas is say $7 per GJ, then at a 22 to 1 ratio, oil should be $150 per barrel. Currently shale oil and shale natural gas are produced simultaneously with the shale natural gas being dumped on the market which has kept U.S. natural gas prices low, but if compressed natural gas vehicles demand increases, the price would up to $7 per GJ. If natural gas costs are closer to $10 per GJ, then the oil to natural gas price ratio will be 17 to 1 and the price of oil will be $170. Volatility of oil and natural gas prices should remain high due to the lack of incentive to switch back to oil from CNG when oil prices go low; due to the possibility of bottlenecks for natural gas delivery when oil prices go high; and due to peak oil. However, in general the 22 to 1 ratio may become the new oil and gas price paradigm.

The interaction between consumers switching from oil to natural gas, LNG markets taking supplies from the U.S. to Europe and Asia, OPEC making agreements, and the Gas Exporting Countries Forum (GECF) including Russia making natural gas market moves should make for an exciting future oil and gas market and could further create bottlenecks for natural gas. The entire world oil market, and LNG market, may well be waged from the North American shale natural gas supply market and the emerging East European shale natural gas market if one emerged there. However, due to the hysteresis of switching from oil to natural gas and the possibility of short run limits to shale natural gas supplies, there should be a lot of volatility in both the oil and gas market in the future even more so than in the past.

REFERENCES

[1] Brown, Stephen P. A. and Mine Kuban Yucel. "What Drives Natural Gas Prices?" *The Energy Journal*, **29**, pp. 45-60, 2008.

[2] Bachmeir, Lance J. and James M. Griffin. "Testing for Market Integration: Crude Oil, Coal and Natural Gas," *The Energy Journal*, **27** pp. 55-72, 2006.

[3] Asche, Frank, Petter Osmundsen and Maria Sandsmark. "The UK Market for Natural Gas, Oil and Electricity: Are the Prices Decoupled?," *The Energy Journal*, International Association for Energy Economics, **27**, pp. 27-40, 2006.

[4] Hartley, Peter R., Kenneth B. Medlock III and Jennifer E. Rosthal. "The Relationship of Natural Gas to Oil Prices," *The Energy Journal*, **29**, pp. 47-65, 2008.

[5] Ramberg, David J. and John E. Parsons. "The Weak Tie Between Natural Gas and Oil Prices," *The Energy Journal*, **33**, pp. 45 -60, 2012.

[6] Hotelling, Harold. "The Economics of Exhaustible Resources," *Journal of Political Economy*, **39**, pp. 137 – 175, 1931.

[7] Gordon, Richard L. "Hicks, Hayek, Hotelling, Hubbert, and Hysteria or Energy, Exhaustion, Environmentalism, and Etatism in the 21st Century." *The Energy Journal*, **30**, pp. 1-16, 2009.

[8] Holland, Stephen P. "Modelling Peak Oil", *The Energy Journal*, **29**, pp. 61-79, 2008.

[9] Reynolds, Douglas B. "Uncertainty in Exhaustible Natural Resource Economics: The Irreversible Sunk Costs of Hotelling," *Resources Policy*, **38**, pp. 532-541, 2013.

[10] The Potential Gas Committee. *Potential Supply of Natural Gas in the United States*, December 31, 2008, the Potential Gas Agency at the Colorado School of Mines, 2009.

[11] EIA, United States Energy Information Administration. *Annual Energy Outlook 2012*, Washington DC. http://www.eia.gov/forecasts/aeo , 2012.

[12] UK ERC, United Kingdom Energy Research Center. *Global Oil Depletion: An assessment of the evidence for a near term peak in global oil production*, UK ERC, ISBN number 1-903144-0-35, 2009.

[13] Fantazzini, F., Hook, M., and A. Angelantoni. "Global oil risks in the early 21st century," *Energy Policy*, **39**, pp. 7865-7873, 2011.

[14] Kumhof, M. and Benes, J. and Kamenik, O. and Mursula, S., Chauvet M., and Selody J., and D. Laxton. "The Future of Oil: Geology versus Technology," *IMF Working Papers* 12/109, International Monetary Fund, 2012.

[15] Lutz, C., Lehr, U., and K. Wiebe. "Economic effects of peak oil," *Energy Policy*, **48**, pp. 829-834, 2012.

[16] Mirchi, A., Hadian, S., Madani, K., Rouhani, O.M., Rouhani, A.M. "World Energy Balance Outlook and OPEC Production Capacity: Implications for Global Oil Security," Energies, **5**, pp. 2626-2651, 2012.

[17] Hallock, John L. Jr. , Wei Wu, Charles A.S. Hall, Michael Jefferson. "Forecasting the limits to the availability and diversity of global conventional oil supply: Validation," *Energy*, **64**, pp. 130-153, 2014.

[18] Reynolds, Douglas B. "World oil production trend: Comparing Hubbert multi-cycle curves," *Ecological Economics*, **98,** pp. 62–71, 2014.

[19] ibid. Ramberg and Parsons, 2012.

[20] Flynn, Peter C. "Commercializing an alternate vehicle fuel: lessons learned from natural gas for vehicles," *Energy Policy*, **30**, pp. 613-619, 2002.

[21] Krause, Clifford. "Surge in Natural Gas Has Utah Driving Cheaply," *New York Times*, August 29, 2008.

[22] Boyle, Grant. "Hydrogen Fuel Cells and Alternatives in the Transport Sector: Issues for Developing Countries," United Nations University International Conference, *An Overview of Alternative Transport Fuels in Developing Countries: Drivers, Status, and Factors Influencing Market Deployment*, Maastricht, The Netherlands, 7-9 November, 2005.

[23] ibid. Krause, 2008.

[24] ibid. Flynn, 2002.

[25] Turrentine, Thomas S. and Daniel Sperling. "A Decision Process Model for Analyzing Consumer Purchases of Alternative-Fueled Vehicles: The Case of CNG Vehicle Conversions in British Columbia," *University of California, Davis, Institute of Transportation Studies*, Reference Number: UCD-ITS-RP-89-01, Research Report, 1989.

[26] Aslam, M.U., H.H. Masjuki, M.A. Kalam, H. Abdesselam, T.M.I. Mahlia and M.A. Amalina. "An experimental investigation of CNG as an alternative fuel for a retrofitted gasoline vehicle," *Fuel*, **85**, pp. 717-724, 2006.

[27] EPA, United States Environmental Protection Agency. *Emission Facts: Greenhouse Gas Emissions from a Typical Passenger Vehicle*, EPA420-F-05-004, February 2005, http://www.epa.gov/otaq/climate/420f05004.htm, 2005.

[28] Georgescu-Roegen, Nicolas. *The Entropy Law and the Economic Process*, Boston, Harvard University Press, 1971.

[29] Rogers, Deborah. "Shale and Wall Street: was the decline in natural gas prices orchestrated?," *Energy Policy Forum*, February, 2013.

[30] Berman, Arthur E. and Lynn F. Pittinger. "U.S. Shale Gas: Less Abundance, Higher Cost," http://www.theoildrum.com/node/8212, Posted by Aeberman on August 5, 2011 - 10:15am, *The Oil Drum*, supply, 2011.

[31] Mazur, Karol. "Economics of Shale Gas," http://www.energybiz.com/article/12/10/economics-shale-gas, *Energybiz*, October 3, 2012.

[32] Dixit, Avinash K. and Robert S. Pindyck. *Investment under Uncertainty*, Princeton, NJ, Princeton University Press, 1994.

[33] ibid. Flynn (2002).

[34] ibid. Reynolds (2014).

[35] Reynolds, Douglas B. "The Economics of Oil Definitions: The Case of Canada's Oil Sands" *OPEC Review*, **29**, pp. 51 – 73, 2005.

[36] Tussing, Arlon and Bob Tipee. *The Natural Gas Industry: Evolution, Structure, and Economics*, Tulsa, Pennwell Publishing Company, 1995.

APPENDIX 1

Variables:

G_r = the retail price of gasoline
N_r = the retail price of natural gas
C = the cost of converting a gasoline vehicle to a CNG vehicle
r = interest rate paid
M_i = miles driven per year
MPG = miles per gallon of gasoline equivalent
P_{WTI} = the West Texas Intermediate price of oil per barrel
P_{HH} = the Henry Hub natural gas price

Assume:

$$H_r = 1.5P_{HH}$$
$$G_r = 2\left({}^{P_{WTI}}/_{46}\right)$$

and an inflation index is:

$$I = {}^{P_{HH}}/_{A}$$

where, A = the average base year expected cost in dollars per GJ of extracting shale (the gasoline equivalent price per GJ = $8 \bullet G_r$.

The cost of natural gas with the inflation index set at 1 is:

$$\left(8G_r - N_r\right)\left[{}^{M_i}/_{8MPG}\right] = C \cdot r \cdot I \qquad (2)$$

Solving for ${}^{P_{WTI}}/_{P_{HH}}$

$$\left[\left({}^{8 \cdot 2 \cdot P_{WTI}}/_{46}\right) - 1.5P_{HH}\right]\left[{}^{M_i}/_{8MPG}\right] = C \cdot r \cdot \left({}^{P_{HH}}/_{A}\right)$$
$$\left({}^{8 \cdot 2 \cdot P_{WTI}}/_{46}\right) - 1.5P_{HH} = {}^{8MPG \cdot C \cdot r \cdot P_{HH}}/_{A \cdot M_i}$$
$$0.3478\, P_{WTI} = \left[{}^{(8MPG \cdot C \cdot r)}/_{(A \cdot M_i)} + 1.5\right] P_{HH}$$
$${}^{P_{WTI}}/_{P_{HH}} = \left[{}^{1.5}/_{0.3478}\right] + \left({}^{8}/_{0.3478}\right)\left[{}^{(MPG \cdot C \cdot r)}/_{A \cdot M_i}\right]$$

$${}^{P_{WTI}}/_{P_{HH}} = 4.31 + {}^{23(MPG \cdot C \cdot r)}/_{A \cdot M_i} \quad (3)$$

When C =\$12,000; r = 13%; MPG = 25 mpg; A = \$7.00 per MCF, and M_i = 11,000 then:

$${}^{P_{WTI}}/_{P_{HH}} = 16$$

If P_{HH} = \$7, then P_{WTI} = \$112 per barrel.

APPENDIX 2

Competitive fueling-station cost of natural gas scenario:

$$N_r = I \cdot b + 1.5 \cdot P_{HH}$$

where b is the expected cost per MCF (equivalent to the cost per GJ) of the fueling-station.

Start with equation (2)

$$\left(8G_r - N_r\right)\left[{}^{M_i}/_{8MPG}\right] = C \cdot r \cdot I \qquad (2)$$

Substitute for Nr gives:

$$\left(8G_r - b \cdot I - 1.5P_{HH}\right)\left[{}^{M_i}/_{8MPG}\right] = C \cdot r \cdot I$$

Then, solving for $\left({}^{P_{WTI}}/_{P_{HH}}\right)$ we have:

$$\left[\left({}^{8 \cdot 2 \cdot P_{WTI}}/_{46}\right) - \left({}^{b \cdot P_{HH}}/_{A}\right) - 1.5P_{HH}\right]\left[{}^{M_i}/_{8MPG}\right] = \left({}^{C \cdot r \cdot P_{HH}}/_{A}\right)$$
$$\left({}^{8 \cdot 2 \cdot P_{WTI}}/_{46}\right) - \left({}^{b \cdot P_{HH}}/_{A}\right) - 1.5P_{HH} = \left({}^{8MPG \cdot C \cdot r \cdot P_{HH}}/_{A \cdot M_i}\right)$$
$$0.3478 \cdot P_{WTI} = \left[{}^{(8MPG \cdot C \cdot r)}/_{(A \cdot M_i)} + \left({}^{b}/_{A}\right) + 1.5\right] P_{HH}$$
$${}^{P_{WTI}}/_{P_{HH}} = \left[{}^{1.5}/_{0.3478}\right] + \left({}^{8}/_{0.3478}\right)\left[{}^{(MPG \cdot C \cdot r)}/_{A \cdot M_i}\right] + \left({}^{1}/_{0.3478}\right)\left({}^{b}/_{A}\right)$$

$${}^{P_{WTI}}/_{P_{HH}} = 4.31 + \left[{}^{(23 \cdot MPG \cdot C \cdot r)}/_{(A \cdot M_i)}\right] + 2.875\left({}^{b}/_{A}\right) \qquad (4)$$

When C =\$12,000; r = 13%; MPG = 25 mpg; A = \$7.00 per MCF; b = \$15 per MCF, and M_i = 11,000 then

$$P_{WTI} / P_{HH} = 22$$

When P_{HH} = \$7 per MCF (i.e. per GJ), then P_{WTI} = \$154 per barrel.

Journal of Energy Challenges and Mechanics
http://www.nscj.co.uk/JECM/
ISSN 2056-9386
Volume 1 (2014) issue 1, article 9

Grid integration of renewables – challenges & technology adoption

可再生能源并网发电 - 挑战与技术采纳

Himadri Endow[1*], Jayant Kumar[2]

[1]*Head, Smart Grid Business Development, Alstom T&D India Ltd*
[2]*Global Smart Grid Program Director, Alstom Grid Inc, USA*

himadri.endow@alstom.com

Accepted for publication on 13th August 2014

Abstract - **As per conservative estimates (BAU framework), India is set to add, 19.2 GW of Solar and 22.4 GW of Wind in the 12th and 13th plan periods with sustained growth of more than 100%. This is expected to push the installed base of grid connected Wind and Solar to 38.5 GW and 20 GW respectively by the year 2022 with a combined capacity penetration of 12.5% up from the present 9.25%. However, the Wind and Solar capacity is mostly concentrated (~90% at present) in the five renewable-rich states in the south and west of the country where local penetration levels vary between 10% and 34%. To promote the use of green power among all states, the renewable purchase obligations (RPO) and Renewable Energy Certificates (REC) have been put in place. Inter-state, as well as Intra-state transmission corridors, are also being strengthened to maximize evacuation and use of the green power across state boundaries. Yet utilisation factors remain woefully low.**

The growing penetration levels and the wide diurnal and seasonal variation of wind generation, introduces complex power flows in the regional grid networks leading to issues in power balance and grid stability. It also impacts the way reserves are managed. The central and state transmission utilities are in the process of defining the requirements of dedicated renewable energy control desks throughout all the dispatching centres. This paper suggests methods fortify the systems in place at the Grid control centres to deal with the variability of renewables, yet maximize its capacity utilisation. It will enumerate the adoption and integration of new generation technology and tools with traditional Energy Management Systems to aid the operators at the grid control rooms. In the process, the basic expectations from the renewable asset owners are also defined.

Keywords – **Renewable integration; decision support; Power balance; Renewable forecasting**

I. INTRODUCTION

Around 90% of renewable generation capacity in India is today concentrated in the five renewable rich states in the west and the south. The combined penetration levels in these states are much higher (at 10% to 34%) than the national average (9.25%). The projected growths of RE capacity in these states mean that the penetration levels will steadily rise and surpass the requirements of renewable power obligations (RPO) of the states, by the end of the current 12th five year plan. Further the mechanism of renewable power certificates [1] will ensure a growing volume of renewable power being evacuated through the state as well as the inter-state transmission corridors to fulfill the RPOs of the renewable deficient states.

The network of modernized state and five regional level load dispatch facilities have been implemented by Power Grid in a phased manner between 1994 and 1996. The NLDC commenced commercial operation in April 2009. These load dispatch centres are equipped with advanced EMS centered scheduling and load dispatch solutions and together have provided the means to improve grid reliability and facilitate the effective implementation of availability based tariff.

The energy deficit situation coupled with the large inertia of the synchronous system in India is responsible for the low levels of correlation between variations in demand and that of grid integrated renewable generation. However continued integration of large amounts of highly variable resources, such as wind and solar, is going to change the grid operating paradigm to a great extent. Maintaining demand-supply balance and ensuring continued system reliability will become a real problem over time for grid operators. The fundamental differences arise from the fact that unlike other types of generation, wind and solar generation cannot be predicted with absolute certainty and cannot be dispatched up and down like a conventional generator. The unpredictable fast changing renewable output unless mitigated will create unexpected power flow situations leading to transmission security issues and also severely stress the conventional generators which will be expected to provide the power balance. A set of dedicated functions and operating plans will be required to be added to the present generation EMS systems to deal with the new reality. The present paper summarises the main technology and tools that need to be inducted and merged with the currently operating systems at the load dispatch facilities.

II. Modelling of Renewables

The renewable energy generating stations are connected to the grid normally at 33 kV, 66 kV, 110 kV, and 132 kV. The EHV transmission system beyond the first point of connection is at 132 kV, 220 kV or 400 kV. Only the large capacity generating stations are connected to the grid directly at 132 kV or 220 kV levels. Thus the renewable injections are more like distributed energy resources, not always within the direct purview of the transmission system operators (Fig. 1). So to integrate these DERs into the traditional EMS of the grid operators, the first step would be to model them so as to make them appear as clusters of equivalent generating units at different voltage levels with the net injection being dynamically computed based on all available information of RE generation. Not all of these DERs, especially the ones at the sub transmission or distribution voltage levels can be expected to be telemetered. So the modeling characteristics and rules should enable estimations of the RE clusters active power production, even if not telemetered, based on historical data and metered data from other RE sources. The estimations are typically done for a defined set of RE clusters by calculating the utilisation factors (UFs), which is the weighted average of actual use over total capacity (Refer Eq. (1)).

$$UF = \sum_{i_{UFREF}\, in UF} \left[WEIGHT(i_{UFREF}) . \frac{MW(i_{UFREF})}{MWMAX(i_{UFREF})} \right] \qquad (1)$$

where,

$MW(i_{UFREF})$= MW value of DER pointed by i_{UFREF}
$MWMAX(i_{UFREF})$= Max capacity of DER pointed by i_{UFREF}
$WEIGHT(i_{UFREF})$= Weight of i_{UFREF}

Typically each DER can be assigned multiple sources for MW value. These are field telemetry through SCADA as the primary source, one or more estimations as the secondary source and split external forecast value as a tertiary source with splitting rules defined from top level entities down to individual DERs. To take care of fluctuating telemetry, filtering rules can be applied to smoothen its effect on the calculation of UF.

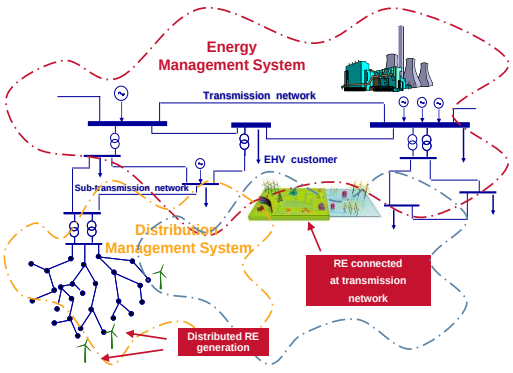


Fig.1, Renewable injections at different voltage levels

The modeling environment shall have to meet two other criteria that of time activated changes in the data model and generic grouping or aggregation of these resources. For example data model with RE sources that will be commissioned or retired in the future or taking into account that a particular resource will change ownership over time. Such criteria assumes importance because within the same RE clusters individual units may belong to different market participants with period ownerships and may ultimately impact the way power balance and settlements are handled. The generic aggregation types can be used to represent topological data such as electrical connectivity, control area or organizational data such as ownership.

III. Forecasting Renewables

There has been a lot of debate of late on how much forecasting accuracy band the wind generator shall be held responsible for in order that the transmission system operator can produce reasonable day-ahead schedules which include renewables. On the other hand, there have not been enough discussions on how to lay down a model which can improve these accuracies over a period of time and ensure that the forecasts are managed such that it can be exploited not only by the real time operating system at the grid control centres but also by other utility wide enterprise applications.

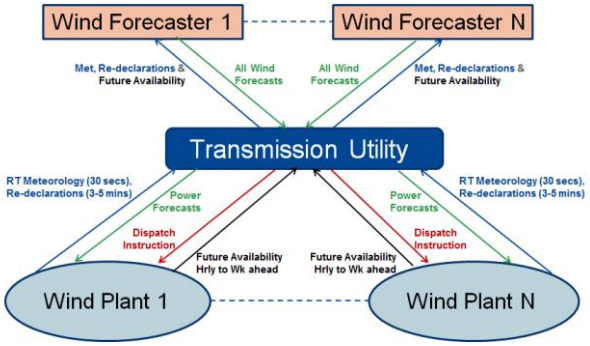


Fig.2, Data flow between RE, Transco & Forecaster

Wind power forecasting has made significant progress in the recent past. Such wind power forecasting is being offered as a service by several companies worldwide. Forecasting is usually based on Numerical Weather Prediction (NWP) models and other complex statistical techniques. Still, results have not reached the same accuracy level as in other forecast domains, often requiring use of multiple forecast sources, and blending to compose a final wind generation schedule. Studies have also shown that implementing a centralized renewable power forecast system is essential to the future of grid operations. Instead of seeking generation forecast from each renewable generator, these forecast schedules can be carried out at the level of the state transmission utility (or the regional grid operator) after ensuring that each renewable generator open up their assets to full on-line monitoring (Fig. 2). This

approach will ensure the historical plant data will be always available with the transmission utility so that when the need arises switching or tuning in an additional forecast service provider as well as best case forecast compositing would be feasible. This will also facilitate sanitisation and quality checks of the raw data of the wind plants before being passed on to the forecast service provider, besides enabling analysis and audits necessary for settlement of RECs and raw forecast performance calculations. Hence the forecast management system shall have to bridge the space between the forecast provider and wind plant data stream on the one hand (top layer in Fig. 3) and the real-time grid dispatching applications on the other (bottom layer in Fig. 3).

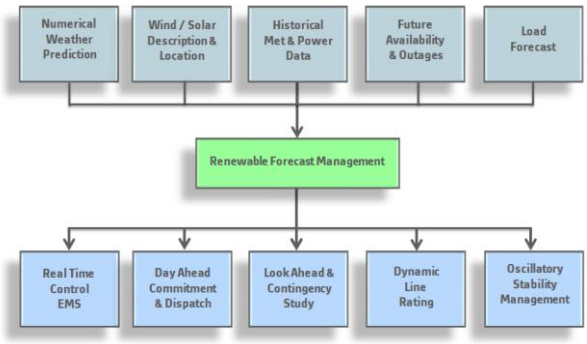


Fig.3, Centralized forecast management

The typical data the wind plant is required to share are real-time meteorology, and on-line schedule re-declarations like turbine active count, max achievable MW, Economic Max MW, Plant generation and Plant lead/lag data (Ref Fig. 4).

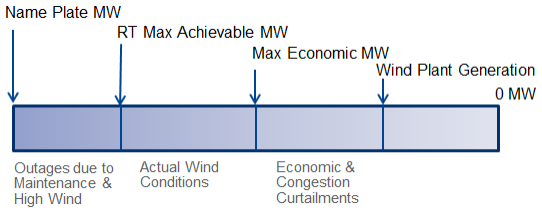


Fig.4, Renewable plant generation build

Following is a high level summary of the basic requirements of the forecast management platform,

- Get wind plant data as mentioned above through interface with SCADA
- Capable of receiving updates from multiple sources, multiple types of time series forecasts
- Monitoring timely reception of incoming forecast schedules and raising alarms for missing data
- Create contiguous composite forecasts using weighting factors (or profiles relative to current time up to a certain end time) that goes from short-term, via mid-term to long-term with the best accuracy of all three forecasts combined
- Create combination of composite schedules like schedule of net-load required to be supported by conventional generation
- Provide for granular schedule over-rides with traceability
- Schedule aggregation along with confidence intervals as per established RE models
- Alarm processing of actual and schedule data like, absolute high/low, rate of change or ramp, maximum change in forecast update and maximum deviation between forecast and actual
- Geospatial schedule visualization and temporal navigation
- Sharing relevant forecast schedules with plant operators in accordance with entity area of responsibility (EAOR), using web services to enable them to re-declare plant data
- Exporting forecast schedules and alarms thereof with the real-time EMS of grid control centre

IV. Situational Awareness

Integrating high levels of renewable generation adds a significant amount of uncertainty combined with higher probability of experiencing large scale system events. Weather related events such as storms or absence of wind/solar are likely to have a direct impact on available generation at a very large scale. Therefore, Situation Awareness becomes a key requirement to operate a system with significant renewable penetration. The goal here is to provide clear and concise information presented in a manner such that operators can quickly understand the current and near future state of the system in order to take appropriate operational decisions.

The user interface (UI) can be built as an extension of existing systems but more appropriately as a separate renewable operator desk. The UI can run within a web browser with several types of views built around a map viewer (WMS, WFS compliant) that provides advanced zooming and panning. Capability to add weather related overlays would be invaluable to depict how system events evolve. The example as in Fig. 5, show transmission lines, individual wind farms (triangles) colored by current utilization factor and total wind generation by balancing areas (green & yellow shaded areas). Details of entities on the map viewer window can be seen in the entity detail pane on the right top of the screen. The charts section below, provide a temporal view of data for the selected entity. This typically includes actual generation, short-term and medium-term forecast for a wind farm or an aggregated area and the corresponding confidence intervals. The Alarm tab is reserved for alarm information of the selected entity and the children tab, for entities that are immediate children of the selected entity in the system hierarchy, e.g., wind farms belonging to an area. By right-clicking on the Chart tab, the user can select the dual view which will simultaneously show the chart and data grid. The slide bar is used to switch from history to current mode back and forth. When in history mode,

Fig.5, User Interface built to SA principles

the map viewer is surrounded by a thick yellow border to warn the user that displayed information is not current.

V. Integration With RT Control System

The conventional real time control systems at the various load dispatch centres have two broad suites of applications,

- Network security analysis and optimization
- Generation / Reserves scheduling, monitoring and control

These are equipped to provide steady state solutions for specific constrained objective functions of the power system. Typically in this case generation and load schedules along with regulating reserves are planned a day in advance. Fig. 6 is a schematic of this real time system showing standard EMS functions (in light blue colour) with specific application modules (in orange colour) added to manage renewables in the mix.

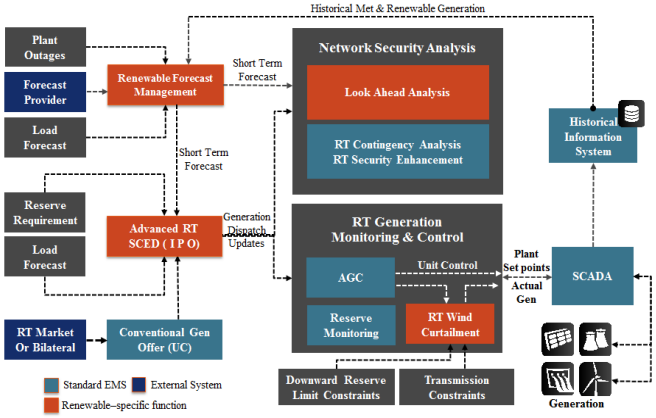


Fig.6, Renewable management within real time EMS system

Generation Dispatching

With RE added to the portfolio, the task is to set up a base load plan considering a reasonable variability (confidence level) of RE generation forecast. This is done by performing a modified and advanced Economic Dispatch (named Intraday Plant Optimization or IPO) to calculate unit base points for current time and for times in the future. To accommodate evolving RE conditions this is repeated so as to further update and implement the short term base load plan every 5 minutes or so. IPO receives the current operating plans for the conventional units, submitted unit offers of currently committed units and the latest load and renewable forecast results. It then produces least cost, minute based unit generation schedules for the next several hours, respecting their minimum and maximum generation limits and also taking into account ramp up/down rate limit and unit forbidden zones. As constraints are modelled in a generic way in the optimization problem, hence additional constraints can be defined, such as network security constraints where IPO is coupled with the contingency application that verifies network security.

Optimal unit generation trajectories computed by IPO are transferred to AGC for real time implementation. Typically it is executed periodically (configurable) and at the occurrence of events like changes in RE production forecast, generation cost curve etc. If the renewable assets and storage devices are clubbed into virtual power plant (VPP), the IPO can be implemented to determine optimal schedules of both the conventional units as well as the VPP to provide instantaneous power demand for each time point and distribute the VPP schedules over the individual assets. For the units, these correspond to not only the required generation MW level, but also to the reserve and regulation requirements. If the total RE generation is too high, the system may not have the required down regulating reserve to balance. In such cases RE curtail limit commands in the form of "output not to exceed" shall be

issued to the VPP or RE. As the value of down regulating reserve recovers, curtail limits are withdrawn that is limit values are raised. A similar dispatching may take place to alleviate network congestion if that is identified as the mitigating source.

Look Ahead Network Analysis

Network analysis relies on well-established processes, such as Power Flow, Contingency Analysis, Security Enhancement, Short-Circuit Analysis, Volt-VAR dispatch and dynamic stability. These applications are made to embrace the RE generation through proper modelling (Section II) so that these RE appear as generator clusters feeding into different injection points with defined composite schedules and confidence levels (Section III). Renewable area forecasts are applied to the model for each time point in the future and the suite of network analysis functions is run. The output results lists all security issues reported for each time point in the future as well as the net transfer capacities of a pre-defined list of corridors.

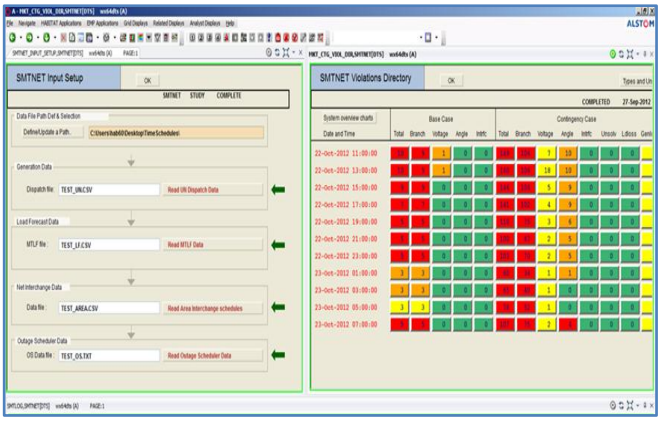


Fig.7, Look Ahead study results within real time EMS

In Fig. 7, the navigation pane on the left is the area where the inputs like load forecast, RE forecast, generation schedules, network analysis function etc. are configured. When the Look Ahead function runs it provides the various base case and contingent case node and branch violations as shown on the right side pane of Fig. 7. The results show the network security over the next twenty two hours in eleven runs. The violations are depicted in various coloured cells where green means normal conditions and yellow, orange, red colours indicate increasing levels of severity of the violations.

Dynamic Transfer Limits

As part of the renewable integration targets, India has plans to strengthen the intra-state as well as inter-state transmission system to facilitate transfer of RE power from RE rich states to other states. This build up is taking place between 132 kV through to 765 kV levels. However the capacity build-up is always faster than building transmission infrastructure leaving gaps in the extent to which the demand will be actually served.

In this context calculating the Total Transfer Capacity (TTC) and the Available Transfer Capability (ATC) for defined transfer paths becomes important. In a conventional EMS system these are calculated following a rigorous process taking into account a variety of factors that can limit the MW transfer. These include branch flow limits, voltage limits, voltage collapse and MW/MVAR interface limits. As a part of the Look Ahead Network Analysis, the RE forecast estimations (Section III) can be factored in to tune the standard TTC/ATC calculations to conform to RE generation forecast. However this still would remain a static calculation as it does not account for the dynamic impacts on the transfer corridor from factors like ageing, current loading of feeder, ambient temperature, solar heating, wind speed and direction, humidity etc. which has a bearing on the conductor's dynamic thermal capacity and consequently capacity to carry load. Field trials have already proven that transfer capacities (or ampacity) can be improved upon by 20% to more than 50% (at times even 70% for shorter periods) for defined intervals of time if impacts of weather and current loading conditions are taken into account. This is a substantial amount which cannot be ignored if injection of bulk renewable power in the grid network is to be maximized. Evolutions in the EMS today provide an alternative to conventional TTC/ATC calculations by providing dynamic line rating calculation module with forecasts (based on historical observations) made available that can be effectively used alongside the renewable forecasts. To calculate ampacity figures on-line, the feeder's sag has to be measured and kept updated. For this three broad methods are employed. These are measurement of weather data, measurement of tension on the conductor or measurement of vibration of the conductor.

VI. Conclusion

Seamless integration of renewable plant visibility and output forecasting, into utility control room operations is fundamental to accommodating large penetrations of wind energy in a huge power system network as exists in India. The forecast need to be updated on an hourly or sub-hourly basis to factor in the dynamically changing conditions of the renewable generation. This is then used to reschedule reserves and re-dispatch conventional generation as often as necessary, as a last resort curtailing RE if enough down regulating reserves are not available. To be able to do all this, all distributed renewable plants have to be modeled in the power system network irrespective of injection voltage levels and 100% visibility of these generations obtained through either real-time telemetry or estimations. Forecast algorithms will mature over time. For use in real-time operations, best-in-class accuracies can be obtained by blending outputs of multiple forecast sources as well as combining forecasts of different time horizons (short term, medium term and long term) together with their varying accuracy levels. Ideally forecasts should be managed centrally at the utility level and shared with renewable generators for the purpose of re-declaration of the renewable generation.

Present generation EMS applications shall have to embrace the embedded renewable generation and carry out intra-day generation schedule optimization to meet active, reserve and

regulation requirements. With time series renewable forecasts, the real-time network analysis functions would be able to depict base case and contingent network and branch node violations from the present time to time points in the future. This look-ahead capability would enable the renewable desk operator to modify his dispatching actions as frequently as required. Total transfer capacity of major network corridors could be substantially improved by taking advantage of the dynamic line rating calculation module. The operator user interface should make the operator aware at all times, the evolving weather and system events with a geographical perspective showing balancing area alarms and temporal trends for actual generation, forecasts and forecast errors.

REFERENCES

[1] Central Electricity Regulatory Commission "Regulation on Renewable Energy Certificates", India, 14th Jan, 2010.

[2] Power Grid Corporation of India Ltd. "Transmission Plan for Envisaged Renewable Capacity," 1st July 2012.

Journal of Energy Challenges and Mechanics

ISSN 2056-9386

http://www.nscj.co.uk/JECM/

Editor:

Dr. Henry Tan
University of Aberdeen, Scotland, United Kingdom

Scope:

Since James Watt, a Scottish inventor, improved efficiency of the steam engine, human civilization relies more and more on a steady supply of energy. Today we are at a transitional age. On the one hand, we see technology advances in the exploration and development of oil and gas, a depleting resource; we see growth in handling aging and decommissioning. On the other hand, we see ideas and plans for new energy infrastructure. This journal is about energy challenges and the underlying mechanics, involving multiple disciplines in science, technology, management and policy-making. Mechanics, fundamentally, is about force and the related behaviours, where force is about relationships, including those physical, human and social. For mechanics, the journal covers interactive boundaries with many other disciplines. For energy, topics include both fossil fuels and many different forms of renewable energy; also, issues related to energy economy, energy policy, efficiency, safety, environment and ecology will also be covered.

North Sea

Cove Bay, Aberdeen, Scotland, United Kingdom