ISSN 1913-1844(Print)
ISSN 1913-1852(Online)

MODERN APPLIED SCIENCE

Vol. 7, No. 2 February 2013

CANADIAN CENTER OF SCIENCE AND EDUCATION

Editorial Board

Contents

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The Investigation of the Barriers in Developing Green Building in Malaysia

Milad Samari[1], Nariman Godrati[1], Reza Esmaeilifar[1], Parnaz Olfat[1] & Mohd Wira Mohd Shafiei[1]

[1] School of Housing, Planning and Building, Universiti sains Malaysia, Malaysia

Correspondence: Milad Samari, School of Housing, Building and Planning, University Sains Malaysia (USM), Penang 11800, Malaysia. E-mail: milad.s1366@gmail.com

Received: December 6, 2012 Accepted: January 1, 2013 Online Published: January 15, 2013

doi:10.5539/mas.v7n2p1 URL: http://dx.doi.org/10.5539/mas.v7n2p1

Abstract

Green building is the foundation of the sustainable construction development. Construction industry with the high contributes with gross domestic product, has undeniable impacts on the economy. Although Green buildings provide a wide range of benefits for the society, green building development suffers from different kinds of market barriers in developing countries including Malaysia. In order to meet green building development in Malaysia, this study aims to investigate the level of developing green building in the current situation, to find important key players and to identify, and to eliminate the important obstacles to green building development. In this research, the respondents were randomly selected from the professionals of Malaysian construction industry across the country and the method applied for collecting data is questionnaire survey. All the questionnaires were sent out to the respondents manually or through e-mail. A total of 673 sets of questionnaire were sent out and 167 (24.81%) questionnaires were received. The quantitative method was used for analysing data through SPSS version 19. Based on the results, the level of developing green building in Malaysia is not satisfied and government has a key role in the development of green buildings in Malaysia. The main barriers can be listed as: lack of credit resources to cover up front cost, risk of investment, lack of demand as well as higher final price.

Keywords: green building, barriers, sustainable construction, Malaysia

1. Introduction

Recently, climate-change, energy crisis and increasing environmental pollution have made the sustainable development issue receive a great attention from the world (EPA, 2008). With growing concern of community about negative affect of human life on the environment, United Nation (1992) lunched the sustainable development document in Rio de Janerio in order to protect the environment (Parkin, 2000). Sustainable development tries to improve a quality of life for current people and future generation (Bossel, 1999). Many definitions of sustainable development are represented in the different contexts and scopes. Sustainable development can be defined "as growing natural and industrial resources which meet the energy need of the present times without settling the ability of next generations for meeting their needs in the same manner" (Hill, 2003). In addition, the United Nations (1987) explains that "sustainable Word development is a collection of methods in order to relieve poverty, create the equitable standards of living, satisfy the basic needs of all peoples, and set up sustainable political practices all while taking the steps essential to avoid irreversible damages to be natural environment in the long-term". One of the enormous and most important industries known as the largest polluters on the environment is the construction industry (Horvath, 1999). In addition, construction yields an annual output of US $4.6 trillion, contributing to 8-10% of the global gross domestic product (GDP) encompassing a workforce of 120 million people and billions of transactions each day (McGraw-Hill, 2008). In US, construction comprises 13.4% of the $13.2 trillion US GDP, in which commercial and residential building construction constitutes 6.1% of the GDP (AHKS, 2011). Therefore, it was revolved that the construction industry has direct and indirect important connections with the various aspects of sustainable development (Bourdeau, 1999).

2. Implementation of Sustainable Development in the Construction Industry of Malaysia

The construction industry of Malaysia is separated into two areas. The first area is general building, which includes residential building, non-residential building and civil engineering building. The second area is special

trade works related to the construction include the activities of metal works, electrical works, and etc (AHKS, 2011). The estimated portion of construction industry in Malaysia is 5% to 6% of the GDP at the end of 2012, there is job opportunities for almost 1.03 million people that represent 8% of total workforce (CIDB, 2006). Estimation demand for construction under 9th Malaysia Plan is projected at RM 280 billion in the average of RM 56 billion per year. The projection is based on the estimation of RM 180 billion of government funded projects, RM 140 billion of private funded and RM 20 billion Public Finance Initiatives (PFI) in this stipulated time frame (CIDB, 2008). The construction industry makes up an important part of the Malaysian economy due to the interaction with other industry branches. The construction industry could be described as a substantial economic driver for Malaysia to achieve sustainable economy (Abdullah, 2004). According to the Master plan and 10th Malaysian plan, the government should apply sustainable development's goals in the construction industry.

Consequently, the various activities of the construction area have to be judged and examined when considering sustainable development. As a significant element in order to implement of sustainable development's goals in the construction industry, green building has an important role in achieving the aim of sustainable development (Fisk, 1988).

3. Green Building

Green building is an important area where cities can implement sustainability objectives. Green buildings are designed to reduce negative impacts on the environment while increasing the occupant health, by addressing these five categories:

- Sustainable site planning
- Safeguarding water and water efficiency
- Energy efficiency, renewable energy and lower greenhouse gas emissions
- Conservation and the reuse of materials and resources, and
- Improved health and indoor environmental quality

The environmental impact of buildings is often underestimated, while the perceived costs of green buildings are overestimated. Kats et al. (2003) comprehensively examined the costs and benefits of green buildings for the state of California in the United State. According to Kats, the average cost premium over just building to code is less than 2%. The Kats report finds that "minimal increases in upfront costs of about 2% to support green design would, on average, result in life cycle savings of 20% of total construction costs more than ten times the initial investment". The majority of savings from green building are in the maintenance part and utility costs (CEA, 2011).

Table 1. Financial benefits of green buildings (per $ft._2$)

Category	20-year Net Present Value
Energy savings	$5.80
Emissions savings	$1.20
Water savings	$0.50
Operations and maintenance savings	$8.50
Productivity and health value	$36.90 to $55.30
Subtotal	$52.90 to %71.30
Average extra cost of building green	(-$3.00 to $5.00)
Total 20-year net benefit	$49.90 to $66.30

Source: Capital E analysis, www.cap-e.com

Considering the statistics, reducing the amount of natural resources buildings consume and the amount of pollution given off are considered crucial for future sustainability.

4. Green Building Development in Malaysia

In January 2009, Malaysia Green Building Index (GBI) was started at the Green Design Forum and organized by the Architectural Association of Malaysia (PAM). The Malaysian construction industry identified the necessity of green rating tool to improve and adapt itself to the tropical climate. GBI has been designed based on another international rating system such as BREEAM (Building Research Establishment Environmental Assessment Method); USA's LEED (Leadership in Energy and Environmental Design) and has been evaluated to be adapted to Malaysian climate conditions. It is an extensive rating system and environmental assessment used for appraising the environmental design and the performance of Malaysian buildings (ACEM, 2012).

Many experts (Atsusaka, 2003; Samari, 2012) believe that the role of governments in promoting green building is undeniable and effective. Rules and regulations should be replaced with enforcing new ones to support green building development. Governments can facilitate green building development by a variety of instruments. However, there is argument about the most effective and efficient instruments among the specialists and researchers. Some studies stated that market base intensives are both effective and efficient tools to address market failure together with non-market problems to improve the situation for green buildings development (Dennis, 2006). Shafii et al. (2005) point out that there are many impediments to developing sustainable development in Asia such as: Lack of awareness (people), Lack of training and education about sustainable design, higher cost, special materials, rules and regulation, lack of demonstration, lack of technology and lack of demand. Davis (2001) believed that most important barriers to green building development can be divided in three groups:

- Builder Incentives: Energy saving and worker productivity are popular benefit of green building. These benefits have positive effect for final owners and impose extra cost for builder. Hence, cost-effective is the main obstacle to green building development.
- Product Information and Sourcing: The common obstacle to green building development in developing countries is lack of green product's information for high-performance building systems. This obstacle leads to the developers to hire specialized consultants. Client Knowledge: the effective ways to remove this barrier are introduce a credible evidence of the advantages of green building and long-term studies to prove the benefits of green building.

5. Methodology

The construction industry in Asia has grown dramatically (Raftery et al., 2004; Bon et al., 2000). Global concerns on the environment and sustaining the world resources for the future generations have increased. Hence, Malaysia has launched initiatives for sustainable development in region. Malaysia's framework for sustainable development started when the following policies were had been formed (Chua et al., 2011):

- National Energy Policy 1979 (NEP79)
- National Depletion Policy 1980 (NDP80)
- Four Fuel Diversification Policy 1981 (4FDP81)
- Fifth Fuel Policy 2000 (5FP2000)

Base on the 10th Malaysia Plan, the government established the AFFIRM framework (Awareness, Faculty, Finance, Infrastructure, Research and Marketing) in order to promote the implementation of sustainable development criteria in the construction industry. Green building as part of sustainable development is under government's consideration to achieve better future for next generations (Sood et al., 2011). Thus, according to the plan adopted, this paper tries to investigate the level of developing green building, to find important key player and major barriers to green building development in Malaysia.

The methodology for this study includes collecting and analyzing data, namely initial data compiled by the questionnaires of the recognized sample. In an effort to eliminate barriers, to develop, and to promote green building in Malaysia, This paper seeks to address the following questions:

1) What is the level of developing green building in Malaysia?

2) Who plays a major role in developing green building in Malaysia?

3) What are important barriers against developing green building in Malaysia?

In this research, the respondents are randomly selected from the professionals of Malaysian construction industry across the country and the method used for collecting data is questionnaire survey. All the questionnaires were sent out to the respondents manually and through e-mail. A total of 673 sets of questionnaire were sent out and

167 (24.81%) questionnaires were received. Quantitative method has been used for analysing data through SPSS version 19. The questionnaire form is designed into the three following sections:

i. Part A: respondent's background.

ii. Part B: the awareness and perception about the green building.

iii. Part C: Green building development barriers

Reliability statistics used in this research is Cronbach Alpha. It is "determining the internal consistency or average correlation of items in a survey instrument to gauge its reliability" (Cronbach, 1951). The formula of Cronbach alpha:

$$\alpha = \frac{\mathrm{N}.\bar{} \; C}{\bar{} \; \mathrm{v} + (\mathrm{n}-1).\bar{} \; \mathrm{c}}$$

N = number of items

$\bar{C}$ =average inter-item covariance among the items

$\bar{v}$= average variance

The reliability of each part of the questionnaire was investigated to ensure that date is reliable. α-Cronbach is between 0 and 1. If α-Cronbach is less than 0.5, data are not reliable; therefore, the results which were gained through the data analysis are not applicable. In opposite, when α-Cronbach is near 1, data are reliable. In this research, α-Cronbach for part B is 0.909 and part C is 0.993. The results were achieved through the SPSS software version 19.

Table 2. Reliability Statistics Part B

Cronbach's Alpha	N of Items
0.909	4

Table 3. Reliability Statistics Part C

Cronbach's Alpha	N of Items
0.993	15

6. Data Analysis and Findings

Understanding the obstacles to green building development will help to find ways to promote the green building market (Chan et al., 2009). The lack of expertise's knowledge in green building development creates an environment that lengthens development time frames (Choi, 2009). In addition, expert's knowledge is a key factor to promote sustainable building (Miyatake, 1996). The level of general awareness about sustainable buildings and their benefits among the construction professionals is low (bellow moderate). According to Table 4 construction industry suffers lack of expertise's knowledge in terms of green building; consequently it will lead to a low level of implementation of green building concept in construction projects.

Table 4. Professional's awareness, Company commitment and level of green building

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
						Statistic	Std. Error
Awareness of professionals about the green building	167	1	5	2.75	1.112	.140	.188
Company commitment to green building	167	1	4	2.23	1.045	.332	.188
The level of developing green building in Malaysia	167	1	5	2.36	1.131	.772	.188

Government's involvement in developing green building is considered as one of the fundamental part and efficient ways (Varone et al., 2000; Qian et al., 2007; Atsusaka, 2003). According to 10th Malaysia Plan, the

government should be promoting green building in residential and non-residential sectors. But professionals' points of view in terms of the level of developing green building in both sectors in Malaysia (Mean: 2.36) are not satisfactory. It is clear that the current strategy for developing green building in Malaysia is not effective.

This also accords with our earlier observations which showed that developing green building by current strategy is not satisfied. Table 8 shows that construction companies would not like to enter green market. This cause is related to the insecure market and high risk for investment. Table 9 illustrates that constructions companies (1: contractor, 2: Consultant, 3: Developer) have low interest in taking part in green projects. Forasmuch as construction companies are important deriver for developing green building, it can be conclude that construction companies are not willing to participate in green projects. This is regarded as another major obstacle facing the development of green buildings in Malaysia.

Tables 5 and 6 show the relationship between professionals's awareness and their education levels. This result (R Square: 0.859, sig: 0.000) indicates that by increasing the respondents' education level their awareness about green homes rises. Therefore, raising the education level among the experts will lead to increasing construction companies' interest in the green building market. It also helps them practice more environmentally friendly methods in their future projects as part of their responsibility to the society. By applying this strategy, level of green features applied by professionals in the buildings will increase dramatically due to high concern on projects' environmental impacts. This aim can be achieved by establishing training courses and periodical seminars for experts in order to promote their knowledge of green building (Samari et al., 2012).

Table 5. Correlation model

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.927[a]	.859	.859	.418

a. Predictors: (Constant), Education level

Dependent Variable: What is the awareness of professionals about the green building?

Table 6. Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.605	.075		8.080	.000
	Education level	1.081	.034	.927	31.754	.000

a. Dependent Variable: What is the awareness of professionals about the green building?

Table 7. Cross tabulation between firm categories and company commitment

			How do you describe your company commitment to green building?					
			very Low	Low	Moderate	High	Total	Mean
Firm category	Contractor	Count	39	13	0	0	52	1.25
		% within Firm category	75.0%	25.0%	.0%	.0%	100.0%	
	Consultant	Count	0	39	39	0	78	2.50
		% within Firm category	.0%	50.0%	50.0%	.0%	100.0%	
	Developer	Count	12	0	0	25	37	3.02
		% within Firm category	32.4%	.0%	.0%	67.6%	100.0%	
Total		Count	51	52	39	25	167	
		% within Firm category	30.5%	31.1%	23.4%	15.0%	100.0%	

According to Tables 7 and 8, Chi-square= 210.924, $P = 0.00$, there is statistically significant association between firm categories and company commitment. Based on the firms' category, developers (Mean= 3.02) have more commitment to green building concept. Although, this commitment is slightly higher than moderate, it might be effective to stimulate consultants and contractors to be more obsessed with environment as they can ask consultants and contractors to practice environmentally friendly in their projects.

Table 8. Chi-square tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	210.924[a]	6	.000
Likelihood Ratio	237.493	6	.000
Linear-by-Linear Association	67.526	1	.000
N of Valid Cases	167		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.54.

According to the previous research, experts believed that government has important role to promote green building (Varone et al., 2000; Fisher et al., 1989; Sutherland, 1991; Golove et al., 1996; Ofori, 2006). In Malaysian construction context, government plays a significant role to promote green building. Government is the key player in term of promoting green building in the construction industry. Government can affect the construction industry by a variety of instruments. Regulatory instruments and incentive instruments are the main tools for governments to develop green building (Yung et al., 2002).

A combination of legislations to enforce companies and market to sustainable development and incentive package for construction firms that practice sustainability in their projects is the best approach that can be applied by governments. Contractors as next major players undoubtedly can promote sustainability and minimize environmental impacts in construction field by using new technologies and environmentally friendly products, and applying waste management Life Cycle Assessment method in construction stage.

Table 9. Major role in developing green building in Malaysia

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Contractor	26	15.5	15.6	15.6
	Consultant	52	31.1	31.1	46.7
	Developer	13	7.7	7.8	54.5
	Government	76	45.5	45.5	100.0
	Total	167	100.0	100.0	

Table 10. Cross tabulation between work experience and major role in developing green building in Malaysia

			Who plays a major role in developing green building in Malaysia?				Total
			Contractor	Consultant	Developer	Government	
Work experience in construction project	x< 5 years	Count	26	52	13	0	91
		% within Work experience	28.6%	57.1%	14.3%	.0%	100.0%
	5 ≤x<10 years	Count	0	0	0	52	52
		% within Work experience	.0%	.0%	.0%	100.0%	100.0%
	x ≥ 10 years	Count	0	0	0	24	24
		% within Work experience	.0%	.0%	.0%	100.0%	100.0%
Total		Count	26	52	13	76	167
		% within Work experience in construction project	15.6%	31.1%	7.8%	45.5%	100.0%

Base on Master Plan and 10th Malaysia Plan, government has to develop green building concept in order to preserve natural resources and enhance urban life quality for the residents. Figure below shows three steps which government can take to develop green concept in construction industry by identifying the barriers that hinder green building development in the country.

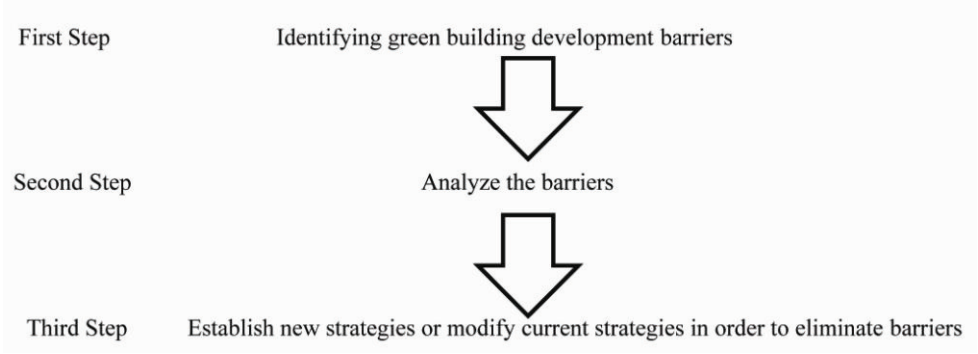


Table 11. Barriers

Code	Barriers
B1	Lack of building codes and regulation
B2	Lack of incentives
B3	Higher investment cost
B4	Risk of investment
B5	Higher final price
B6	Lack of credit resources to cover up front cost
B7	Lack of Public awareness
B8	Lack of demand
B9	Lack of strategy to promote green building
B10	Lack of design and construction team
B11	Lack of expertise
B12	Lack of professional knowledge
B13	Lake of database and information (case study)
B14	Lack of technology
B15	Lack of government support

To explore which barriers to green building are the most important compared to other ones in construction industry in Malaysia, respondents were asked to rate their importance in different levels for each item. Data were analyzed based on the Mean and Median (Table 13). These measures of dispersion are used to assess the homogenous or heterogeneous nature of the collected data (Bernard, 2000).

Table 12. Barrier's descriptive statistics

Code	Rank	Range	Minimum	Maximum	Mean	Medan	Variance	Skewness	
		Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
B1	5	4	1	5	3.44	4	1.477	-.416	.188
B2	11	4	1	5	3.05	3	1.280	.171	.188
B3	10	4	1	5	3.06	3	.984	-.159	.188
B4	2	4	1	5	3.68	4	1.449	-.717	.188
B5	4	4	1	5	3.58	4	1.471	-.521	.188
B6	1	4	1	5	3.80	4	1.762	-.606	.188
B7	13	4	1	5	2.59	3	1.134	.451	.188
B8	3	4	1	5	3.61	4	1.784	-.829	.188
B9	9	4	1	5	3.10	3	1.657	-.283	.188
B10	8	4	1	5	3.13	3	1.344	-.002	.188
B11	14	4	1	5	2.67	3	1.114	.257	.188
B12	6	4	1	5	3.21	3	1.239	-.132	.188
B13	15	4	1	5	2.44	2	1.609	.439	.188
B14	12	4	1	5	2.83	3	1.506	-.198	.188
B15	7	4	1	5	3.14	3	1.047	-.314	.188

In Malaysia, for developing green building government has just introduced two incentives such as tax exemption and stamp duty (ACEM, 2012). According to the finding the current incentives are not effective enough to encourage construction firms to enter green building development. Financial incentives are also not able to recoup the high upfront cost of green buildings and make it more affordable for construction companies. In 2007, Bandy et al. have shown that that higher upfront cost (new design, technology and construction method) is the main impediment to green building development. Governmental financial incentives have essential rules to provide low risk and affordable financial resources for green developers in both commercial and residential sectors. Public awareness about green building has been an important component that led to high demand (Toronto green development standard, 2006). Improving public awareness about green building leads to better informed consumers who will demand better products from companies and encourage more green building development. In addition, cost savings can potentially increase a consumer's willingness to pay extra. To achieve sustainable green home development and to make balance between green home owners' benefits and construction companies' profits are critical issues.

7. Conclusion

This paper reports the results of a questionnaire survey conducted in Malaysia on the barriers of the green building development. The level of green building development from professionals' point of view has been investigated and the most important barriers have been identified. The findings suggest that government roles especially incentive instruments such as structural incentives, subsidy and rebate program, tax incentive scheme, low interest mortgage loan, voluntary rating system and market and technology assistance are the significant drives for eliminating barriers to green building development.

Acknowledgement

This work was supported in part by the Fellowship Scheme of Universiti Sains Malaysia.

References

Abdulllah, F. M., Chiet, V. C., Anuar, K., & Shen, T. T. (2004). An Overview on the Growth and Development of the Malaysian Construction Industry. Workshop on Construction Contract Management, University Teknologi, Malaysia.

ACEM Directory (Association consulting engineers Malaysia). (2012). Retrieved from http://www.acem.com.my/index.php?option=com_content&task=view&id=58&Itemid=1

AHKS report. (2011). Malaysian-German Chamber of Commerce, Market Watch, the Construction Sector. Retrieved from http://malaysia.ahk.de/fileadmin/ahk_malaysia/Bilder/Others/Market_Watch_Malaysia_Construction_Industry_2011.pdf

Atsusaka, N. (2003). *Growing the Green Building Industry in Lane County* - a report for the lane county sustainable business and job project, report prepared from the program for Watershed and Community Health, Institute for a Sustainable Environment, University of Oregon, USA.

Bandy, R., Danckaert, C., Fetscher, G., Holmes, B., Gale, M., Mirsky, M., … Stewart, S. (2007). *Leed in upstate New York: an exploration of barriers, resources and strategies*. USGBC New York Upstate chapter and environment finance center, EPA region 2, Maxwell capstone project.

Bernard, H. R. (2000). *Social Research Methods: Qualitative and Quantitative Approaches.* London: Sage Publishing Ltd.

Bon, R., & Crosthwaite, D. (2000). *The Future of International Construction.* London: Thomas Telford Limited.

Bossel, H. (1999). *Indicators for Sustainable Developments - Theory, Method, Application.* International Institute for Sustainable Development, pp 2-8.

Bourdeau, L. (1999). Sustainable development and the future of construction: a comparison of vision from various countries. *Journal of Building Research and Information, 27*(6), 354-366. http://dx.doi.org/10.1080/096132199369183

Capital E Analysis. (2011). Green building, redirecting our campus toward environmental sustainability. Retrieved from http://www.colbysawyer.edu/assets/pdf/GreenBuildingPoster.pdf

Chan, E. W. W., Qian, K. Q., & Lam, I. P. (2009). The market for green building in developed Asian cities: the perspectives of building designers. *Energy Policy, 37*, 3061-3070. http://dx.doi.org/10.1016/j.enpol.2009.03.057

Choi, C. (2009). Removing Market Barriers to Green Development: Principles and Action Projects to Promote Widespread Adoption of Green Development Practices. *JOSRE, 1*(1), 107-138.

Chua, S. C., & Oh, T. H. (2011). Green progress and prospect in Malaysia. *Renewable and Sustainable Energy Reviews, 15*, 2850-2861. http://dx.doi.org/10.1016/j.rser.2011.03.008

CIDB. (2006). Strategic Recommendations for Improving Environmental Practices in Construction Industry. Construction Industry Development Board (CIDB) Malaysia. Kuala Lumpur.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334. http://dx.doi.org/10.1007/BF02310555

Davis. A. (2001). Barriers to Building Green. Retrieved from http://www.greenbiz.com/blog/2001/11/01/barriers-building-green

Dennis, K. (2006). The Compatibility of Economic Theory and Proactive Energy Efficiency Policy. *The Electricity Journal, 19*(7), 58-73. http://dx.doi.org/10.1016/j.tej.2006.07.006

Economic Planning Unit. (2010). The Tenth Malaysia Plan 2011-2015. Malaysia.

EPA. (2008). Report on the Environment. Retrieved from http://www.epa.gov/greenbuilding/pubs/whybuild.htm

Fisher, A. C., & Rothkop, M. H. (1989). Market failure and energy policy: a rationale for selective conservation. *Energy Policy, 17*(4), 397-406. http://dx.doi.org/10.1016/0301-4215(89)90010-4

Fisk, W., & Rosenfeld, A. (1998). *Potential nationwide improvements in productivity and health from better indoor environments.* Berkeley: Lawrence Berkeley National Laboratory.

Golove, W. H., & Eto, J. H. (1996). *Market barriers to energy Efficiency: a critical reappraisal of the rationale for public policies to promote energy efficiency, report done by Energy & Environment Division*. Lawrence Berkeley National Laboratory, University of California, USA.

Government of Malaysia, (2006). *Malaysian 9th Plan*, Chapter 21: 437.

Hill, M. (2003). Dictionary of Scientific and Technical Terms. USA, McGraw-Hill Publication.

Horvath, A. (1999). *Construction for Sustainable Development – A Research and Educational Agenda.* Submitted for the Berkeley-Stanford Construction Engineering and Management Workshop, August 26-28, 1999.

Kats, G., & Capital, E. (2003). *The cost and financial benefits of green buildings: a report to California's sustainable building task force, developed for the Sustainable Building Task Force*. California, USA.

Malaysian Construction Outlook. (2008). Presentation by Business Development Division, Construction Industry Development Board (CIDB), August 2008.

Master Plan(Construction Industry of Malaysia), 2006-2015 (CIMP). (2006).

McGraw-Hill Construction. (2008). Key Trends in the European and U.S. Construction Marketplace: Smart Market Report.

Miyatake, Y. (1996). Technology development and sustainable construction. *Journal of Management in Engineering, 12*(4), 23-27. http://dx.doi.org/10.1061/(ASCE)0742-597X(1996)12:4(23)

Ofori, G. (2006). *Attaining sustainability through construction procurement in Singapore*. CIB W092–Procurement Systems Conference 2006, Salford, UK.

Parkin, S. (2000). Context and drivers for operationalizing sustainable development. Proceedings of The Institution of Civil Engineers: *Civil Engineering, 138*, 9-15. http://dx.doi.org/10.1680/cien.2000.138.6.9

Qian, Q. K., & Chan, E. H. W. (2007). *Government measures for promoting Building Energy Efficiency (BEE): a comparative study between China and some developed countries*. The CRIOCM2007 International Symposium on Advancement of Construction Management and Real Estate, Sydney, Australia.

Raftery, J., Anson, M., Chiang, Y. H., & Sharma, S. (2004). *Regional overview*. In: Chiang, Y.H., et al. (Eds.). The Construction Sector in Asian Economics. London: Spon Press.

Samari, M., Ghodrati, N., & Shafiei, M. V. M. (2012). Implementation of Sustainable Development in Construction Industry in Penang Island, International Conference on Civil Engineering AICCE'12.

Samari, M. (2012). Sustainable Development in Iran: a Case Study of Implementation of Sustainable Factors in Housing Development in Iran. *IPEDR, 37*. Singapore: ACSIT Press.

Shafii, F., & Othman, M. Z. (2005). *Sustainable Building and Construction in South-East Asia*, Proceedings of The Conference on Sustainable Building South-East Asia, Malaysia, 11-13.

Sood, M. S., Chua, H. K. M., & Peng, Y. L. (2011). Sustainable Development in the Building Sector: Green Building Framework in Malaysia. ST-8: Best Practices & SD in Construction.

Sutherland, R. J. (1991). Market barriers to energy-efficient investments. *Energy Journal, 12*(3), 15-34. http://dx.doi.org/10.5547/ISSN0195-6574-EJ-Vol12-No3-3

Toronto green development standard report. (2006). Retrieved from http://www.toronto.ca/planning/environment/greendevelopment.htm

Varone, F., & Aebischer, B. (2000). Energy efficiency: the challenges of policy design. *Energy Policy, 29*, 615-629. http://dx.doi.org/10.1016/S0301-4215(00)00156-7

World Commission on Environment and Development. (1987). *Our Common Future*. Oxford: Oxford University Press.

Yung, E., & Chan, E. H. W. (2002). Evaluating Environmental Management Policies: International Trend Development of Construction Management and Real Estate; 100-111.

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Accounting for Dispersion and Correlation in Estimating Safety Performance Functions. An Overview Starting from a Case Study

Orazio Giuffrè[1], Anna Granà[1], Tullio Giuffrè[2] & Roberta Marino[1]

[1] Department of Civil, Environmental, Aerospace and Materials Engineering, Università degli Studi di Palermo, Palermo, Italy

[2] Faculty of Engineering and Architecture, Kore University, Enna, Italy

Correspondence: Anna Granà, Department of Civil, Environmental, Aerospace and Materials Engineering, Università degli Studi di Palermo, Viale delle Scienze al Parco d'Orleans, Palermo 90128, Italy. Tel: 39-91-2389-9718. E-mail: anna.grana@unipa.it

Received: November 20, 2012 Accepted: December 18, 2012 Online Published: January 15, 2013

doi:10.5539/mas.v7n2p11 URL: http://dx.doi.org/10.5539/mas.v7n2p11

Abstract

In statistical analysis of crash count data, as well as in estimating Safety Performance Functions (SPFs), the failure of Poisson equidispersion hypothesis and the temporal correlation in annual crash counts must be considered to improve the reliability of estimation of the parameters. After a short discussion on the statistical tools accounting for dispersion and correlation, the paper presents the methodological path followed in estimating a SPF for urban four-leg, signalized intersections. Since the case study exhibited signs of underdispersion, a Conway-Maxwell-Poisson Generalized Linear Model (GLM) was fitted to the data; then a quasi-Poisson model in the framework of Generalized Estimating Equations (GEEs) was performed in order to account for correlation.

Results confirm that dispersion and correlation are phenomena that cannot be eluded in the estimation of SPFs under penalty of loss of efficiency in estimating model parameters. Generalized Estimating Equations overcome this problem allowing to incorporate together dispersion and temporal correlation when a quasi-Poisson distribution is used for modeling crash data. Moreover, whereas GEE regression is handy (many statistical software packages have already implemented GEE functions), the interest of COM-Poisson regression, because of difficulties in interpreting the model parameters and in arranging COM-Poisson codes, is still limited to the research field.

Keywords: safety performance function, signalized intersections, COM-Poisson model, road safety

1. Introduction

It is well-known that Safety Performance Functions (SPFs) allow to estimate the expected number of crashes on entities (road sections or intersections). Differently from models to evaluate the potential accident rate (Mauro & Cattani, 2004), through a mathematical equation SPFs express the average crash frequency of a site as a function of traffic flow and other site geometric and/or functional characteristics. In statistical analysis of crash count data some problems must be addressed to improve the reliability of the estimations. Data and methodological issues associated with crash-frequencies are widely discussed by Lord and Mannering (2010) and by Turner et al. (2011); two of these issues will be focused in the following. First, the data structure can invalidate equidispersion hypothesis on which the Poisson model is founded; moreover, the same crash-data generating process makes inefficient estimations based on the traditional Poisson model. In order to relax the Poisson assumption of equidispersion, linear exponential family models incorporating a dispersion parameter, as well as quasi-likelihood methods, represent a potential solution to this question. Second, with many years of data, it is necessary to account for the year-to-year variations in crash counts because of the influence of factors that can change every year. This creates a temporal correlation that affects the reliability of the SPF estimates obtained through traditional estimation procedures of the model.

Starting from these considerations, the main purpose of the paper is to show how estimates efficiency can be improved taking into account either dispersion and temporal correlation in annual crash counts. In case of overdispersion, the Generalized Estimating Equations (GEEs) method with Negative Binomial distribution (NB1

or NB2) is an effective tool to address the mentioned problems and to increase efficiency of estimates (Lord, 2006); these models take into account for overdispersion by means of a parameter called overdispersion parameter α (with $\alpha>0$); moreover, it is possible to account for correlation, too. Underdispersion, instead, is a phenomenon which has been less convenient to model directly than overdispersion, mainly because it is less commonly observed. Oh et al. (2006) report that crash data show characteristics of underdispersion especially in cases where the sample is small and the sample mean is very low; underdispersion can be also caused by the data generating process that is independent from the size of the sample or its mean. To further improve estimation of the parameters in case of underdispersion, the Conway-Maxwell-Poisson (COM-Poisson) distribution can be introduced (Lord et al., 2010; Giuffrè et al., 2011). Unfortunately, currently COM-Poisson distribution cannot be used in GEE context; so, the analyst that decides to estimate a SPF with GLM method and COM-Poisson distribution is likely to overlook correlation phenomenon running the risk of impairing estimates efficiency.

Focusing on the latter question, the paper presents a case study that exhibited signs of underdispersion. The convenience of handling underdispersion through a GLM COM-Poisson model rather than a GEE quasi-Poisson model is evaluated comparing the efficiency of the obtained estimates. It has to be noted that the small sample size used in the study could have affected the estimation of model parameters (coefficients and dispersion parameter). Therefore, though results can help to highlight the potential of COM-Poisson model and of GEE quasi-Poisson model in handling under-dispersed data, further researches should be carried out using different dataset (namely larger sample size) to confirm them.

2. Method

The methodological approach applied for estimating a Safety Performance Functions at 4-leg signalized intersections is described in the following sections. The intersections here analyzed are characterized by being inserted in urban area, factor that may directly affect the expected number of crashes. Model performance measures used to verify the suitability of the predictive model are also introduced. Before introducing the case study, the main issues related to the dispersion and the temporal correlation of the data to be considered in the treatment of crash data are extensively accounted for and discussed.

2.1 Accounting for Dispersion

Poisson and Negative Binomial (NB) distributions in the context of Generalized Linear Models (GLM) have been widely used for some time to analyze crash count data in the estimation of Safety Performance Functions (SPFs). The basic regression model for count data is in fact the Poisson model: $y_{tj} \mid x_{tj} \sim$ Poisson with $E(y_{tj}|\boldsymbol{x}_{tj}) = \mu_{tj} = \exp(\boldsymbol{x}_{tj}'\boldsymbol{\beta})$, where y_{tj} denotes the crash count data at year t and at site j; $\boldsymbol{x}_{tj} = [x_{1tj}, \ldots, x_{ktj}]$ is a k dimensional vector of covariates at year t and at site j; $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_k]$ is the vector of parameters to be estimated.

In statistical analysis of crash count data, the Poisson model is often inadequate because of its implicit restriction on the distribution of observed crash counts; specifically in the Poisson model the variance of the random variable is constrained to equal the mean; this property is usually called equidispersion. Different factors can invalidate equidispersion hypothesis in the data: consequently, data often exhibit overdispersion (i.e. the variance is larger than the mean) and, occasionally they exhibit underdispersion with the mean exceeding the variance. The overdispersion can be caused by various factors, such as data clustering, unaccounted temporal correlation, model misspecification, but it has been shown to be mainly attributed to unobserved heterogeneity: counts are viewed as being generated by a Poisson process, but the analyst ignores that the rate parameter (μ) is itself a random variable and so he (or she) does not specify any distribution for it. In order to relax the Poisson assumption of equidispersion, quasi-likelihood methods represent a potential solution. By this way few assumptions about the distribution for the dependent variable are required; only the relationship between the outcome mean and the covariates, and between the mean and variance must be specified (McCullagh & Nelder, 1989): the variance v_{tj} of the observation y_{tj} is expressed as a known function, g, of the expectation μ_{tj}, i.e. $v_{tj} = \phi\, g(\mu_{tj})$ where ϕ is the scale parameter. Then a quasi-Poisson distribution can be used to model crash data: the mean is the same of the Poisson mean and $g(\mu_{tj}) = \mu_{tj}$; the variance is now a function of the mean: $v_{tj} = \phi\mu_{tj} = (1 + \alpha)\mu_{tj}$, where α is the dispersion parameter. In the case of under-dispersion $\alpha<0$ (and thus $0<\phi<1$); in the opposite case ($\alpha> 0$ and $\phi> 1$) the data are overdispersed. Although different Poisson-based distributions have been developed to accommodate over-dispersion, the most common distribution for crash count data remains the Poisson-gamma or Negative Binomial (NB) distribution. Properties of the traditional NB model were illustrated by Cameron and Trivedi (1998). The NB model arises mathematically by assuming that unobserved heterogeneity across sites is gamma distributed, while crashes within sites are Poisson distributed. According to Hauer (1997), the NB distribution offers a simple way to handle overdispersion, especially when the final equation has a closed form; in addition, the mathematics to manage the relationship between the mean and the

variance structures is relatively simple. There are two well-known nonnested forms of the negative binomial model, denoted as NB1 and NB2 (Greene, 2007). In overdispersed mixture models the Poisson mean is a random variable; when it is gamma distributed the model form will be $y_{tj}|x_{tj} \sim$ NB with $E(y_{tj}|x_{tj}) = \mu_{tj}$, $var(y_{tj}|x_{tj}) = \mu_{tj}+\alpha\mu_{tj}^{p}$; for p equals 1 we obtain the quasi-Poisson distribution above discussed. The choice of the specific form of NB depends on data set; NB2 can only accommodate overdispersion and it cannot be used in the case of underdispersion. Application of mixture models for crash data are referred by Park and Lord (2009). In NB regression the overdispersion parameter α is commonly assumed the same for all entities (intersection and/or road segment). Modeling crashes on road sections, Hauer (2001) showed that this assumption may have undesirable consequences when road sections in the data base differ in length; if model parameters are estimated by maximum likelihood, the relative influence of long road sections is much diminished whereas very short road sections exert an unduly large influence. The assumption of an overdispersion parameter constant within the sites can also lead to an inconsistency when the safety of a road section by the Empirical Bayes method is estimated (Hauer, 2001). In modeling crash-flow relationships for urban intersections, Miaou and Lord (2003) also challenged the assumption of fixed dispersion parameter. Because of the complexity and interaction of traffic flow in and around intersections, they supposed that the unmodeled heterogeneity of the mean of crash counts would be spatially structured. This means that the variance of NB models is not a simple function of mean, but contains a dispersion function that depends on site-specific characteristics such as major and minor road traffic flows. As well as overdispersion, underdispersion can violate some basic count data modeling assumptions; Winkelmann et al. (1995) proposed a correction for an underdispersed event count probability distribution. Several researchers recently have proposed new and innovative methods for analyzing under-dispersed crash data. Moreover, the Conway-Maxwell-Poisson (COM-Poisson) distribution has been re-introduced by statisticians to model count data characterized by either over- or under-dispersion (Shmueli et al., 2005; Guikema & Coffelt, 2008; Lord et al., 2010; Zou et al., 2011). The COM-Poisson distribution was first introduced in 1962 by Conway and Maxwell; only in 2008 it was evaluated in the context of a GLM by Guikema and Coffelt (2008), Lord et al. (2008) and Sellers and Shmueli (2010). The COM-Poisson distribution is a two parameter generalization of the Poisson distribution that is flexible enough to describe a wide range of count data distributions (Sellers & Shmueli, 2010); since its revival, it has been further developed in several directions and applied in multiple fields (Sellers et al., 2011).

For a random variable Y_{tj} (e.g., a discrete count at year t and at site j) COM-Poisson probability distribution function is given by the equation:

$$P\left(\mathrm{Y}_{tj} = \mathrm{y}_{tj}\right) = \frac{\lambda_{\mathrm{tj}}^{\mathrm{y}_{tj}}}{(\mathrm{y}_{tj}!)^{\nu}\, \mathrm{Z}\ (\lambda_{tj}, \nu)} \tag{1}$$

where:

$$\mathrm{Z}\ (\lambda_{tj}, \nu) = \sum_{s=0}^{\infty} \frac{\lambda_{tj}^{s}}{(s!)^{\nu}} \tag{2}$$

λ_{tj} = a centering parameter, denoting the expected value under a Poisson distribution associated with the generic observation at year t and at site j (Sellers and Shmueli, 2010);

$\nu\ (\geq 0)$ = the dispersion parameter (where $\nu < 1$ for over-dispersion and $\nu > 1$ for under-dispersion).

This formulation allows for a non-linear decrease in ratios of successive probabilities in the form:

$$\frac{P\left(\mathrm{Y}_{tj} = \mathrm{y}_{tj}\text{-}1\right)}{P\left(\mathrm{Y}_{tj} = \mathrm{y}_{tj}\right)} = \frac{y_{tj}^{\nu}}{\lambda_{tj}} \tag{3}$$

Shmueli *et al.* (2005) refer that the serie $\lambda^{s}/(s!)^{\nu}$ converges for any $\lambda > 0$ and $\nu > 0$, since the ratio of two subsequent terms of the serie λ/s^{ν} tends to 0 as $s \to \infty$. Moments of COM-Poisson distribution can be expressed using the recursive formula (Shmueli et al., 2005):

$$E\left(Y_{tj}^{r+1}\right)=\begin{cases}\lambda_{tj}\ \mathrm{E}^{1-\nu}\left(Y_{tj}+1\right) & r=0\\ \lambda_{tj}\ \dfrac{\partial}{\partial\ \lambda_{tj}}\mathrm{E}\left(\mathrm{Y}_{\mathrm{tj}}^{\mathrm{r}}\right)+E\left(Y_{tj}\right)\cdot\mathrm{E}\left(\mathrm{Y}_{\mathrm{tj}}^{\mathrm{r}}\right) & r>0\end{cases} \tag{4}$$

Using an asymptotic approximation for $Z(\lambda_{tj}, \nu)$, $E(Y_{tj})$ can be closely approximated by:

$$E\left(Y_{tj}\right)=\lambda_{tj}\ \ \frac{\partial\ \log \mathrm{Z}\left(\lambda_{\mathrm{tj}},\nu\right)}{\partial\ \lambda_{\mathrm{tj}}}\approx\lambda_{tj}^{1/\nu}-\frac{(\nu-1)}{2\nu} \tag{5}$$

This approximation is especially good for $\nu \leq 1$ or $\lambda_{tj} > 10^{\nu}$. Once the COM-Poisson regression model has been estimated, fitted values can be computed by equation 5, setting:

$$\hat{\lambda}_{\ tj}=\exp\ (\mathrm{x}_{\mathrm{tj}}^{'}\hat{\beta}) \tag{6}$$

2.2 Accounting for Correlation

Count data often consist of observations over several time periods, these are usually referred to as longitudinal data or panel data. A data structure of this kind creates a specific problem in safety modeling because of the failure of the independence hypothesis for the variate response. With respect to the precision of the parameters estimation, this is a serious issue in safety modeling; that is why elusion of the correlation within responses can lead to misleading conclusions in model interpretation on the basis of incorrect estimates of the variances and of an inefficient or biased estimate of the regression coefficients (Diggle et al., 2002; Giuffrè et al., 2007). Literature refers on several applications by GEE models (see by way of example, Lord & Persaud, 2000; Cafiso & D'Agostino, 2012).

It is well known that a robust model estimation based on Generalized Estimating Equations (GEEs) can still supply consistent estimates of the regression parameters even if the correlation matrix is incorrectly specified (Fitzmaurice, 1995). Standard application of GEEs to safety analysis uses robust (or sandwich) estimates of regression coefficients under an independence hypothesis for the working correlation matrix. Nevertheless it has been demonstrated that the efficiency of estimators declines as the correlation increases, and the decline becomes appreciable when the correlation is greater than 0.4 (Fitzmaurice, 1995). Furthermore, efficiency losses - when independence is a false assumption - will seriously compromise the significance of estimates for within-subject correlations greater than 0.5. Errors are particularly large when the correlation is highly positive or highly negative. Other researchers think that the search for the right correlation matrix becomes important only when marginal models are estimated by using data with missing values (Lord & Persaud, 2000); nevertheless, they agree that standard errors of the coefficients usually are underestimated when temporal effects are not included in the modeling framework (Hardin & Hilbe, 2003). As above referred, quasi-likelihood methods allow to relax Poisson assumption of equidispersion. The GEEs method is an extension of the quasi-likelihood approach (Liang & Zeger, 1986); by this method, once a proper distribution has been selected for the data set (e.g., Poisson, Quasi-Poisson, Negative Binomial), it is possible to improve the efficiency of parameters estimation specifying a "working" correlation matrix, in order to explicitly take into account for the correlation within observations. Parameters estimates can be found by solving the following estimating equation:

$$\sum_{\mathrm{t=1}}^{\mathrm{m}}\sum_{\mathrm{j=1}}^{\mathrm{n}}\mathrm{D}_{\mathrm{tj}}^{'}\mathrm{V}_{\mathrm{tj}}^{-1}\left[\mathrm{Y}_{\mathrm{tj}}-\mu_{tj}\right]=0 \tag{7}$$

for the year t ($t=1, 2, \ldots, m$) and for the entity j ($j = 1, 2, \ldots, n$)

$$\boldsymbol{D}_{tj}=\frac{\partial\boldsymbol{\mu}_{\mathrm{tj}}}{\partial\boldsymbol{\beta}}=\begin{bmatrix}\dfrac{\partial\mu_{1\mathrm{j}}}{\partial\beta_1} & \dfrac{\partial\mu_{1j}}{\partial\beta_2} & \cdots & \dfrac{\partial\mu_{1\mathrm{j}}}{\partial\beta_k}\\ \vdots & \vdots & \vdots\vdots & \vdots\\ \dfrac{\partial\mu_{\mathrm{mj}}}{\partial\beta_1} & \dfrac{\partial\mu_{\mathrm{mj}}}{\partial\beta_2} & \cdots & \dfrac{\partial\mu_{\mathrm{mj}}}{\partial\beta_k}\end{bmatrix} \tag{8}$$

where:

$\boldsymbol{V}_{tj}$ = covariance matrix at year t and site j

$\boldsymbol{\beta}$ = vector of regression parameters $(\beta_1,..., \beta_k)'$

$\boldsymbol{\mu_{tj}}$ = $n \times 1$ vector $(\mu_{1j}, \mu_{2j}, ..., \mu_{mj})'$ of expected values for the j^{th} site.

With the usual notation we can put $\mu_{tj} = g^{-1}(\boldsymbol{X}_{tj}\ \beta)$, where g is referred to as the "link" function. The temporal correlation in responses is described by means of a $m \times m$ matrix $R(\gamma)$, where γ represents the type of correlation with $\gamma = (\gamma_1, \ldots, \gamma_{m-1})'$ and $\gamma_j = corr(Y_{t_r j}, Y_{t_s j})$ for $r, s = 1, ..., m-1$ and $r \neq s$. The covariance matrices can be expressed as follows:

$$\boldsymbol{V}_{tj} = \boldsymbol{A}_{tj}^{1/2} \boldsymbol{R}_{tj}(\gamma) \boldsymbol{A}_{tj}^{1/2} \tag{9}$$

$$\operatorname{cov}\left(\hat{\beta}\right) = \sigma^2 \left[\sum_{t=1}^{m} \sum_{j=1}^{n} D_{tj}^{'} V_{tj}^{-1} D_{tj} \right]^{-1} \tag{10}$$

where $\boldsymbol{A}_{tj}$ is a diagonal matrix containing the variances of the elements of $\boldsymbol{Y}_{tj}$, expressed in terms of $\boldsymbol{\beta}$. The simultaneous solution of above equations with the iterative weighted least squares method gives the GEEs estimates for $\boldsymbol{\beta}$ and for the correlation type $(\boldsymbol{\gamma})$. Since it is not possible to know the proper correlation type for repeated observations, Liang and Zeger (1986) proposed the use of a "working" correlation matrix by replacing in the upper basic equation $\boldsymbol{V}_{tj}$ with $\hat{\boldsymbol{V}}_{tj}$ based on the correlation matrix $\hat{\boldsymbol{R}}_{tj}$.

Then, from the above-mentioned structure of the GEE procedure it has to be highlighted that the specification of the form of response correlation represents the central issue in obtaining more efficient estimates. In fact, although GEE models are generally robust to misspecification of the correlation structure (Liang & Zeger, 1986), when the specified structure does not incorporate all the information on the correlation of measurements within the subjects, loss of efficiency in estimates can be expected (Ballinger, 2004). In order to get the true correlation structure, it is necessary to test on different hypotheses of within-subject correlation. It is usually possible to choose from different types of correlation structures (e.g., *independence, exchangeable, unstructured, autoregressive, m-order dependence*). Assuming the independence structure (i.e. assuming subjects are independent of each other), one has to sacrifice the advantage of using GEE (because it does not consider the within-subject correlation); nevertheless, the independence structure can still be useful in fitting a base model. The exchangeable structure supposes no logical ordering for within-entity observations; when an unstructured working correlation matrix is chosen, estimates of all possible correlations of within-entity responses are made and they are included in the estimates of the variance. The m-order dependence structure implies that the γ_s take different values at different time points. Finally, for data that are correlated within cluster over time, an autoregressive correlation structure can be appropriated; in this case correlation within subject is specified as an exponential function of the lag time period (Ballinger, 2004).

In general, decisions about correlation structure should be guided first by theory; there are specific correlation structures that are appropriate for time-dependent correlation structures (e.g. autoregressive) and some that are not (e.g. exchangeable). For cases in which analyst may be undecided between few structures, Pan (2001) proposed a test that extends the Akaike's Information Criterion to allow comparison of covariance matrices under GEEs models to the covariance matrix generated under the independence hypothesis (Quasi-likelihood under the Independence model Criterion, QIC). The correlation structure with the QIC score closest to zero is judged to be the most appropriate. Applications of QIC in choosing the best correlation structure for a marginal GEEs model, as well as some useful general guidelines, are given by Hardin & Hilbe (2003).

2.3 Model Performance Measures

Technical literature suggests different goodness-of-fit methods to evaluate predictive performance of models and to find the model that best explains the data among all estimated models. The methods used in this paper include the following (where the subscript "*i*" denotes the generic observation at year *t* and at site *j*):

2.3.1 Mean Prediction Bias (MPB)

MPB gives a measure of the magnitude and direction of the average model bias (Oh et al., 2003). If the MPB is positive then the model over-predicts crashes and if the MPB is negative then the model under-predicts crashes. It is computed using the following equation:

$$MPB = \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - y_i) \tag{11}$$

where N is the sample size, $\hat{y}_i$ and y_i are the predicted and observed crashes at site *i* respectively.

2.3.2 Mean Absolute Deviance (MAD)

MAD gives a measure of the average mis-prediction of the model (Oh et al., 2003). The model that provides MAD closer to zero is considered to be the best among all the available models. It is computed using the following equation:

$$MAD = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \tag{12}$$

2.3.3 Mean Squared Predictive Error (MSPE)

MSPE is typically used to assess the error associated with a validation or external data set (Oh *et al.,* 2003). The model that provides MSPE closer to zero is considered to be the best among all the available models. It can be computed using the following equation:

$$MSPE = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2 \tag{13}$$

2.3.4 Akaike Information Criterion (AIC)

The AIC (Akaike, 1974) is a measure of the goodness-of-fit of an estimated statistical model and is defined as:

$$AIC = -2\ log\ L + 2\ p \tag{14}$$

where:

L = the maximized value of the likelihood function for the estimated model;

p = the number of parameters in the statistical model.

The *AIC* methodology is used to find the model that best explains the data with a minimum of free parameters, penalizing models with a large number of parameters. The model with the lowest *AIC* is considered to be the best model among all available models. The following expressions of log-likelihood were used in order to compute *AIC*:

Poisson and quasi-Poisson distributions: $log\ L = y_i\ log\ \mu_i - log\ (y_i!)$

COM-Poisson distribution: $\log L = \sum_{i=1}^{n} y_i \log \lambda_i - \nu\sum_{i=1}^{n} \log\left(y_i!\right) - \sum_{i=1}^{n}\log Z(\lambda_i, \nu)$

Quasilikelihood under the Independence model Criterion (QIC)

As above referred, since the GEE method is a quasi-likelihood based method, an extension of the Akaike's information criterion is needed to compare covariance matrices under GEE models to the covariance matrix generated under the independence hypothesis. So *AIC* statistic is replaced by the *QIC* statistic, defined as (Pan, 2001):

$$QIC = -2\ Q + 2\ p \tag{15}$$

where Q is the quasi-likelihood function ($Q = L/\phi$) and p is the number of parameters in the statistical model. The model with the lowest *QIC* is considered to be the best model among all available models.

The marginal R^2-test

The marginal *R^2-test* supplies a measure of improvement in fit between the estimated model and the intercepted-only one; it compares predicted values from the model (after it is estimated) against the actual values (observations) and against the squared deviations of the observations from mean values for the response variable:

$$R_m^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i - \overline{y}_i\right)^2} \tag{16}$$

R^2 statistics can be interpreted as the amount of variance in the response variable that is explained by the fitted model (Hardin & Hilbe, 2003).

3. Estimating SPF for Urban Four-Leg Signalized Intersections - A Case Study

3.1 Data Description

This section provides an overview of the characteristics of the data set used in this study.

The data collected at nineteen urban, four-leg, signalized intersections in the Municipality of Palermo City road network included: crash occurred from 2000 to 2007, crash-related maneuvers, infrastructure characteristics and traffic volumes; crash data were directly acquired from reports available at the Municipal Police Force; only fatal and injury crashes were considered. Note that, since eight years of crash data were considered as distinct observations, there were $8 \times 19 = 152$ observations. Table 1 shows yearly and 8-year crash statistics (minimum, median, mean, maximum), as well as total crashes for the entire dataset. Table 2 reports mean values over 8-years (2000-2007) of crashes and Annual Average Daily Traffic (AADT) for each site.

Table 1. Annual Crash Statistics, All Collision Types, 2000 - 2007

Year	Minimum	Median	Mean	Maximum	Total
2000	0	3	3.74	13	71
2001	0	3	3.63	7	69
2002	0	4	4.21	10	80
2003	0	3	3.58	8	80
2004	0	3	3.47	9	66
2005	1	3	3.42	10	65
2006	0	3	3.63	9	69
2007	0	3	3.68	13	70
2000-07	0	3	3.67	13	558

Description: Total crashes, yearly and 8-year crash statistics for the entire datase.

Table 2. Mean values of AADTmajor, AADTminor at intersections and crashes

Intersection	$(AADT_{major})_m$ [10^3 veh/d]	$AADT_{minor}$ [10^3 veh/d]	$(Crashes)_m$ [crash/year]
1	41.257	8.010	2.25
2	16.772	9.852	1.75
3	14.237	10.864	2.13
4	20.777	19.748	2.88
5	31.688	15.798	3.00
6	25.499	13.194	3.50
7	26.499	15.866	5.00
8	21.910	21.164	3.38
9	27.085	24.153	4.50
10	28.945	27.403	9.75
11	28.092	10.588	4.88
12	30.653	7.000	5.50
13	31.825	17.615	7.38
14	9.056	6.713	0.50
15	17.876	12.975	1.38
16	22.218	13.991	3.00
17	20.592	9.986	2.50
18	18.629	16.101	4.75
19	20.677	10.187	1.75

Description: 2000-2007 crashes and Annual Average Daily Traffic (AADT) by site.

3.2 Model Selection and Covariates

It is well known that the development of a Safety Performance Function (SPF) involves: i) which explanatory variables should be used; ii) how variables should enter into the model, i.e. the best model form. The results of

these tasks can be summarized as follows:

- all the covariates explored and the significant ones (at the 15% confidence level) are listed in Table 3;
- different model forms were investigated considering the combinations of all the significant variables listed in Table 3. The exploratory analysis suggested to insert into the model only the covariates listed on the right column in Table 3; it also revealed that the functional model form could be described using the power function for the variables ($F_1 + F_2$) and RW_2, and the exponential function for the variable PW1. Note that the variable F_1 was not introduced in the model because it was accounted in variable ($F_1 + F_2$); R_1 was not selected because its introduction in the model together with the other significant variables, either in the power or in the exponential form, produced no appreciable benefits on the model performance. Then the final selected model had the form:

$$y_{tj} = \beta_0 \left(F_1 + F_2\right)_{tj}^{\beta_1} RW_{2j}^{\beta_2} e^{\beta_3 PW_{1j}} \tag{17}$$

where:

y_{tj} = expected number of crashes for the year t and the intersection j;

$(F_1 + F_2)_{tj}$ = sum of Annual Average Daily Traffic on major- and minor-road for the year t and the intersection j;

RW_{2j} = minor-road roadway width at the intersection j;

PW_{1j} = major-road permitted ways at the intersection j (PW_1 = 0 for one way only, PW_1= 1 for two ways or more);

β_0, β_1, β_2, β_3 = parameters to be estimated.

Table 3. Variables explored and selected

Variables	Abbreviation	Significant variables	Selected
Annual Average Daily Traffic on major-road	F_1	✓	
Annual Average Daily Traffic on minor-road	$F_1 + F_2$	✓	✓
Major-road roadway width	RW_1		
Minor-road roadway width	RW_2	✓	✓
Major-road number of lanes	NL_1		
Minor-road number of lanes	NL_2		
Major-road red light time	R_1	✓	
Minor-road red light time	R_2		
Major-road permitted ways	PW_1	✓	✓
Minor-road permitted ways	PW_2		

Description: Variables represent all the possible covariates; in the third column only variables explored as significant are reported. The fourth column includes only the variables selected as coviariates of the model.

3.3 Accounting for Dispersion in GLM Regression

Generalized Linear Model was performed to estimate model coefficients assuming a Poisson error distribution and a quasi-Poisson distribution.

After model estimation with quasi-Poisson distribution it was observed that data clearly exhibited underdispersion ($\alpha < 0$). Then, to further improve parameters estimates, the COM-Poisson distribution was used.

For estimating Poisson and quasi-Poisson regression coefficients and standard errors GenStat software was used; since GenStat does not provide standard error for dispersion parameter, it was estimated iterating the following auxiliary regression (Cameron & Trivedi, 1998):

$$\frac{\left(y_{tj}-\hat{y}_{tj}\right)^2-y_{tj}}{\hat{y}_{tj}}=\alpha+\varepsilon_{tj} \tag{18}$$

where $\hat{y}_{tj}$ is the fitted value estimated by the Poisson model. Once the dispersion parameter was estimated its value was used to obtain new estimates of model parameters; this iteration was performed until all the values (i.e., dispersion parameter, coefficients) converged.

COM-Poisson model was estimated using *R* software, after implementation of codes arranged by Sellers and Shmueli (2010), available at www9.georgetown.edu/faculty/kfs7/research.

Table 4 shows coefficients estimates, standard errors and goodness-of-fit statistics for the Poisson model (model 1), the quasi-Poisson model (model 2), the COM-Poisson model (model 3). COM-Poisson coefficients are for centering parameter λ (see equation 6) and not for mean as in the case of Poisson/quasi-Poisson model (Lord *et al.* 2010; Sellers & Shmueli 2010); this is why COM-Poisson coefficients cannot be directly compared with the Poisson/quasi-Poisson ones.

Table 4. Coefficients estimates and goodness-of-fit for the three models

	Model 1			Model 2			Model 3		
variables	est	s.e.	t	est	s.e.	t	est(*)	s.e.	t
Constant (β_0)	-6.94	0.77	-9.01	-6.94	0.60	-11.49	-12.41	1.72	-7.22
$F_1 + F_2$ (β_1)	1.69	0.14	12.15	1.69	0.11	15.49	3.18	0.40	7.75
RW_2 (β_2)	0.73	0.17	4.23	0.73	0.14	5.40	1.38	0.28	4.79
PW_1 (β_3)	0.22	0.09	2.38	0.22	0.07	3.04	0.39	0.13	2.96
ν	-			-			1.99	0.25	
α	-			-0.46	0.06		-		
MPB	0.00			0.00			0.01		
MAD	1.09			1.09			1.09		
MSPE	1.98			1.98			1.97		
AIC	543			543			519		

(*) *model parameters to be used for determining* $\hat{\lambda}_i$ according to equation 6.

Description: Coefficients estimates, standard errors and goodness-of-fit statistics for the three estimated models: the Poisson model (model 1), the quasi-Poisson model (model 2), the COM-Poisson model (model 3).

From results showed in Table 4, although there is no difference between parameters estimates (and GOF) for Poisson and quasi-Poisson models, it can be seen that the consideration of underdispersion in the data improves the estimates accuracy, as it is shown by the reductions in the standard errors values.

The shape parameter of the COM-Poisson distribution again shows under-dispersion ($\nu > 1$). This confirms that the Poisson distribution is not appropriate to interpret the data-set. *MPB*, *MAD* and *MSPE* values of the models have slight differences, so they do not add any significant information about the models prediction capacity; on the contrary, *AIC* values indicate that the COM-Poisson model has to be considered the best among all estimated models.

3.4 Accounting for Correlation through GEE Regression

Considering that data consisted of repeated measures over the years that could be correlated within an entity, it seemed appropriate to account for the correlation within responses. For this reason GEE regressions were fitted under different working correlation matrices, these are assuming that repeated observations were correlated in different ways.

Again GenStat software was used for this purpose. As unfortunately up to now software packages do not allow to

perform a GEE model using a COM-Poisson distribution, we used the quasi-Poisson distribution to consider simultaneously both the correlation and the under-dispersion in the data.

As mentioned earlier, three forms of correlation were explored starting from the simplest one (*independence structure*) for which observations are thought (unrealistically) to be uncorrelated. In contrast to the hypothesis of independence, an *unstructured* working correlation matrix was stated to allow the free estimates on the within-site correlation from the data. A correlation structure of a stationary *7-dependent process* was assumed, too. The GEE regression results under the three named working correlation matrices are summarized in Table 5, in which both R^2_m and the Pan statistic (QIC) are shown.

Table 5. Coefficients estimates and goodness-of-fit in GEEs

variables	independence			unstructured			7-dependence		
	est	s.e.	t	est	s.e.	t	est	s.e.	t
Constant (β_0)	-6.94	0.82	-8.46	-7.63	0.66	-11.56	-7.35	0.63	-11.67
$F_1 + F_2$ (β_1)	1.69	0.09	18.78	1.78	0.08	22.25	1.77	0.08	23.11
RW_2 (β_2)	0.73	0.24	3.04	0.87	0.18	4.83	0.79	0.19	4.21
PW_1 (β_3)	0.22	0.10	2.20	0.31	0.08	3.88	0.29	0.08	3.42
$\alpha^{(*)}$	-0.47			-0,45			-0.48		
MPB	0.00			1.19			0.27		
MAD	1.09			1.14			1.15		
MSPE	1.98			2.09			2.12		
QIC	1009			1008			1032		
R^2_m	0.68			0.66			0.65		

(*) the unstructured working correlation matrix allows a dispersion parameter varying over time in the observation period; the mean value is reported in the table.

Description: GEE regression results under the *independence*, *unstructured* and *7-dependence* working correlation matrices and model performance measures.

The results in Table 5 show that unstructured and 7-dependence working correlation matrices give parameters estimates more accurate than under independence hypothesis; that is particularly evident with regard to the traffic variable (F_1+F_2), which most affects the model (i.e. the one with the highest parameter value). Nevertheless, the hypothesis of a correlation matrix different from the independent one does not provide any significant improvement in QIC and do not allow the best correlation structure to be determined clearly.

Thus it was decided to thoroughly analyze the model adequacy from another point of view.

According to the purpose of the current research it was thought that the width of the confidence interval for the mean could represent a criterion for model selection, as well as for deciding about the correlation structure of response.

As estimates of GEE parameters (β_{GEE}) are asymptotically normal, it can be derived that an asymptotic (1-α)100% confidence interval for the mean ($\boldsymbol{x}'_{\beta GEE}$) is given by:

$$x'\beta_{GEE} \pm z_{1-\frac{\alpha}{2}}\sqrt{x'V_{GEE}x}$$

Table 6 shows $\sqrt{x'V_{GEE}x}$ values for models estimated under unstructured, dependence and independence hypothesis for the working correlation matrix.

It can be easily seen that relaxing the independence hypothesis, the confidence interval is considerably reduced both in the case of unstructured and dependence correlation matrix.

That confirms benefits in explicitly considering correlation in the data as allowed by a GEE procedure.

Table 6. $\sqrt{x'V_{GEE}x}$ sample values for different hypothesis of correlation matrix

correlation matrix	min	mean	max	median
unstructured	0.042	0.079	0.127	0.077
dependence	0.043	0.078	0.121	0.079
independence	0.054	0.093	0.156	0.091

Description: Above values are related to models estimated under unstructured, dependence and independence hypothesis for the three working correlation matrix.

4. Discussion and Conclusion

Results reported in the previous section highlight that the dispersion and the correlation are phenomena that cannot be eluded in the estimation of SPFs under penalty of loss of efficiency in estimating model parameters. Generalized Estimating Equations (GEEs) procedure overcomes this problem because it allows to incorporate together the dispersion in the data and the temporal correlation. Moreover, GEE regression using different working correlation matrices (i.e. assuming that repeated observations are correlated in different ways) allows to gain a better understanding of the proper correlation structure in the crash count data.

The paper presents an application of the GLM and the GEE procedures in developing a SPF; data of the case study pertain to a sample of nineteen urban, four-leg, signalized intersections in Palermo, Italy, for the years 2000-2007. Since data were found to exhibit clear signs of underdispersion together with a quasi-Poisson distribution a COM-Poisson was considered in the GLM context; then a GEE quasi-Poisson model was performed under three different working correlation matrices.

Through the case study it was shown that:

- quasi-Poisson and COM-Poisson GLM regression model allow to handle underdispersion and to obtain more accurate estimates for the model parameters than the traditional Poisson model;
- according to the Akaike Information Criterion, COM-Poisson regression further improves the predictive performance of the proposed model and it provides a better goodness-of-fit than the quasi-Poisson model, at least for the case study dataset;
- GEE quasi-Poisson model, especially under working correlation matrices different from the independent one, allows more accurate estimates of model parameters than the correspondent GLM model (that do not account for the temporal correlation in crashes).

It has to be noted that regression models based on COM-Poisson distribution, despite their benefits, have disadvantages in terms of model estimation due to:

- difficulties in interpreting the model parameters (λ, ν), in obtaining fitted values and in comparing coefficients from a COM-Poisson regression model to those from other models (in fact, the comparison is possible only in terms of fitted values);
- difficulties in arranging COM-Poisson codes, that is actually possible using few statistical softwares just in GLM context (for example *R* software or WinBUGS package); this makes the interest in using COM-Poisson regression limited to the research field;
- difficulties in accounting for temporal correlation in the data since up to now it cannot be used in GEE context.

It follows that the practical utility of COM-Poisson regression may be limited only to the cases of underdispersed data. On the contrary, in more frequent cases of overdispersion, Negative-Binomial GEE models allow to obtain correct estimates for model parameters accounting simultaneously both for correlation and for dispersion in the data. This is handy since many statistical software packages have already implemented GEE functions.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716-723. http://dx.doi.org/10.1109/TAC.1974.1100705

Ballinger, G. A. (2004). Using generalized Estimating equations for longitudinal data analysis. *Organizational Research Methods, 7*(2), 127-150. http://dx.doi.org/10.1177/1094428104263672

Cafiso, S. D., & D'Agostino, C. (2012). Safety performance function for motorways using generalized estimation. *Procedia-Social and Behavioral Sciences, 53*(3 October, 2012), 901-910. http://dx.doi.org/10.1016/j.sbspro.2012.09.939

Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*, *Econometric Society Monograph No. 30*. Cambridge, UK: Cambridge University Press.

Conway, R. W., & Maxwell, W. L. A. (1962). Queuing model with state dependent service rates. *Journal of Industrial Engineering, 12*, 132-136.

Diggle, P. J., Heagerty, P., Liang, K. Y., & Zeger, S. L. (2002). *Analysis of Longitudinal Data* (2nd ed.). New York, NY: Oxford University Press Inc.

Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data. *Biometric, 51*(1), 309-317.

Giuffrè, O., Granà, A., Giuffrè, T., & Marino, R. (2007). Improving reliability of road safety estimates based on high correlated accident counts. *Transportation Research Record, 2019*, 197-204. http://dx.doi.org/10.3141/2019-23

Giuffrè, O., Granà, A., Marino, R., & Corriere, F. (2011). Handling Underdispersion in Calibrating Safety Performance Function at Urban, Four-Leg, Signalized Intersections. *Journal of Transportation Safety & Security, 3*(3), 174-188. http://dx.doi.org/10.1080/19439962.2011.599014

Greene, W. (2007). Functional forms for the negative binomial model for count data. *Economics Letters, 99*(3), 585-590. http://dx.doi.org/10.1016/j.econlet.2007.10.015

Guikema, S. D., & Coffelt, J. P. (2008). A Flexible Count Data Regression Model for Risk Analysis. *Risk Analysis, 28*(1), 213-223. http://dx.doi.org/10.1111/j.1539-6924.2008.01014.x

Hardin, J. W., & Hilbe, J. M. (2003). *Generalized Estimating Equations*. London, UK: Chapman & Hall/CRC Press.

Hauer, E. (1997). *Observational before-after studies in road safety*. *Estimating the Effect of Highway and Traffic Engineering Measures on Road Safety*. Oxford, UK: Pergamon Press.

Hauer, E. (2001). Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis & Prevention, 33*(6), 799,808. http://dx.doi.org/10.1016/S0001-4575(00)00094-4

Liang, K. Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*(1), 13-22. http://dx.doi.org/10.1093/biomet/73.1.13

Lord, D. (2006). Modelling motor vehicle crashes using Poisson-Gamma models: Examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. *Accident Analysis & Prevention, 38*(4), 751-766. http://dx.doi.org/10.1016/j.aap.2006.02.001

Lord, D., Geedipally, S. R., & Guikema, S. D. (2010). Extension of the Application of Conway-Maxwell-Poisson Models: Analyzing Traffic Crash Data Exhibiting Under-Dispersion. *Risk Analysis, 30*(8), 1268-1276. http://dx.doi.org/10.1111/j.1539-6924.2010.01417.x

Lord, D., Guikema, S. D., & Geedipally, S. (2008). Application of the Conway-Maxwell- Poisson Generalized Linear Model for Analyzing Motor Vehicle Crashes. *Accident Analysis & Prevention, 40*(3), 1123-1134. http://dx.doi.org/10.1016/j.aap.2007.12.003

Lord, D., & Mannering, F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A: Policy and Practice, 44*(5), 291-305. http://dx.doi.org/10.1016/j.tra.2010.02.001

Lord, D., & Persaud, B. (2000). Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations (GEE) Procedure. *Transportation Research Record, 1717*, 102-108. http://dx.doi.org/10.3141/1717-13

Mauro, R., & Cattani, M. (2004). Model to evaluate potential accident rate at roundabouts. *Journal of Transportation Engineering, 130*(5), 602-609. http://dx.doi.org/10.1061/(ASCE)0733-947X(2004)130:5(602)

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed.). London, UK: Chapman & Hall/CRC.

Miaou, S. P., & Lord, D. (2003). Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. *Transportation Research Record, 1840*, 31-40. http://dx.doi.org/10.3141/1840-04

Oh, J., Lyon, C., Washington, S. P., Persaud, B. N., & Bared, J. (2003). Validation of the FHWA Crash Models for Rural Intersections: Lessons Learned. *Transportation Research Record, 1840*(1), 41-49. http://dx.doi.org/10.3141/1840-05

Oh, J., Washington, S. P., & Nam, D. (2006). Accident Prediction Model for Railway-Highway Interfaces, *Accident Analysis & Prevention, 38*(2), 346-56. http://dx.doi.org/10.1016/j.aap.2005.10.004

Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics, 57*, 120-125.

Park, B. J., & Lord, D. (2009). Application of finite mixture models for vehicle crash data analysis. *Accident Analysis & Prevention, 41*(4), 683-691. http://dx.doi.org/10.1016/j.aap.2009.03.007

Sellers, K. F., & Shmueli, G. (2010). A flexible regression model for count data*, The Annals of Applied Statistics, 4*(2), 943-961. http://dx.doi.org/10.1214/09-AOAS306

Sellers, K. F., Borle, S., & Shmueli, G. (2011). The COM-Poisson model for count data: a survey of methods and applications. *Applied Stochastic Models in Business and Industry*. http://dx.doi.org/10.1002/asmb.918

Shmueli, G., Minka, T. P., Kadane, J. B., Borle, S., & Boatwright, P. (2005). A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution, *Journal of the Royal Statistical Society: Part C, 54*(1), 127-142. http://dx.doi.org/10.1111/j.1467-9876.2005.00474.x

Turner, S., Persaud, B., Lyon, C., Bassani, M., & Sacchi, E. (2011). International crash experience comparisons using prediction models. *Road and Transport Research, 20*(4), 16-27.

Winkelmann, R., Signorino, C., & King, G. (1995). A Correction for an Underdispersed Event Count Probability Distribution, *Political Analysis, 5*, 215-228.

Zou, Y., Lord, D., & Geedipally, S. R. (2011). Over- and Under-Dispersed Crash Data: Comparing the Conway-Maxwell-Poisson and Double-Poisson Distributions. *91st TRB Annual Meeting*, January 22-26, 2012. Retrieved from http://ceprofs.civil.tamu.edu/dlord/

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Solar Chimney Model Parameters to Enhance Cooling PV Panel Performance

Mohammed Sh-eldin[1], K. Sopian[1], Fatah O. Alghoul[1], Abdelnasser Abouhnik[2] & Ae. Muftah M.[3]

[1] Solar Energy Research Institute, Faculty of Eng., Universiti Kebangsaan Malaysia, Selangor, Malaysia

[2] School of Science & Engineering, Manchester Metropolitan University, Manchester, UK

[3] Faculty of Industrial Science and Technology, University Malaysia Pahang, Kuantan, Malaysia

Correspondence: Mohammed Sh-eldin, Solar Energy Research Institute, Faculty of Eng., Universiti Kebangsaan Malaysia, Selangor, Malaysia. E-mail: xiaoyao6554@gmail.com

Received: November 11, 2012 Accepted: January 14, 2013 Online Published: January 21, 2013

doi:10.5539/mas.v7n2p24 URL: http://dx.doi.org/10.5539/mas.v7n2p24

Abstract

The concept of using the Solar Chimney plays an important role in a wide range of topics to improve cooling system efficiency such as drying process, and single and multi-story buildings ventilation against temperature rising. In this paper, study the effective solar cooling chimney parameter model to enhance the performance of photovoltaic (PV) cooling system. First, a brief description of theoretical performance predictions of the solar cooling chimney also discusses the effect of the ambient wind velocity on the photovoltaic panel. Second, analysis air velocities at different points in solar cooling chimney are predicted and the temperature drop also estimated to predicted air velocities in the duct. Finally, from simulation result it was found for chimney height range 0.3 m - 3 m and at 60 ^{o}C, the air velocity increase from 0.6 to 1.78 m/s and Pressure difference between inlet and outlet increase from 0.5 to 5.3 KPa, which improve the PV panel voltage 8%.

Keywords: air cooling, natural convection, solar cell cooling, solar cooling chimney

1. Introduction

In last three decades solar chimneys (SC) are applied in different fields such as ventilation, drying process, or production of electricity systems. Solar chimney is passive elements and one of the most promising a natural power generator using the stack effect to induce buoyancy-driven airflow. Schlaich (1995) indicated solar chimney was primarily used for power generation. It utilizes solar radiation to increase the air temperature inside the SC channel and the resulting generates the buoyant flow through the channel. Photovoltaic panels have negative temperature coefficient due to which its open circuit voltage decreases by certain V/°C of rise in the panel temperature and it convert's small amount of the incident solar radiations to electricity while some amount of solar radiation is converted to heat.

Many researchers introduce many theoretical and experimental studies and focused mainly on SC structure optimization such as minimize installation cost and maximize power output. Various methods of cooling a solar cell have been previously suggested in the literature (Royne et al., 2005; Anderson et al., 2008; Kermani et al., 2009; Moshfegh & Sandberg, 1998). Most of the methods that are used for cooling of PV panel involve an active medium which requires auxiliary power. Having an active cooling system for photovoltaic cell will make it costly and complicated have the ongoing maintenance cost. Air flow due to buoyancy and heat transfer in a vertical channel heated with simulated heat source has been investigated numerically and experimentally (Akbarzadeh et al., 2009). Moshfegh and Sandberg (1998) has shown that 30% of heat flux is transfer to the unheated wall of the duct from the PV panel by radiation and then this heat is transferred to the air in the duct, this helps to increase the air temperature and in turn the air velocity. Brinkworth (2000) has presented a routine procedure to estimate the air flow under an inclined roof top PV panel and have mention the importance of radiation heat transfer in improving the air flow in the duct or the air gap between the PV panel and roof. Some work using commercial PV panels and outdoor experiments has been first carried out by (Tonui & Tripanagnostopoulos, 2008) and they focused on investigating the thermal performance of PV air collector.

This paper described the possibility of integrating the chimney effect to cool the photovoltaic cell. A mathematical model was developed and analytical performances prediction, such as the output and the power

delivered by a solar chimney power plant, according to geometrical parameters, such as the height, length, cavity width and the solar radiation.

2. Description Solar Cooling Chimney (SCC)

From Figure 1, the solar radiation is incident on the transparent bottom surface of the solar chimney. These solar radiations are absorbed by the air that is under the solar chimney which eventually gets warmed up and its density decreases. The warm air will try to rise up sue to the buoyancy force and escape through the chimney. Tall chimney will provide a pressure difference between the bottom and the top which will further assist in the draft of the air. This velocity of air is utilized to drive the turbine inside the chimney that will generate the electricity.

This design consists of a dedicated absorber section in addition to the design suggested by Tonui and Tripanagnostopoulos (2008) such that it will help to enhance the natural draft of air which can be utilized for cooling of PV panels. The velocity of air rising up in the chimney is directly proportional to the energy absorbed by the air.

SCC is decided into two main parts: top part, and the middle part as shown in Figure 1.

Top part of the SCC consists of the vertical extension to the middle part. Top part will act like a chimney in this setup to enhance the natural draft created by the warm rising air.

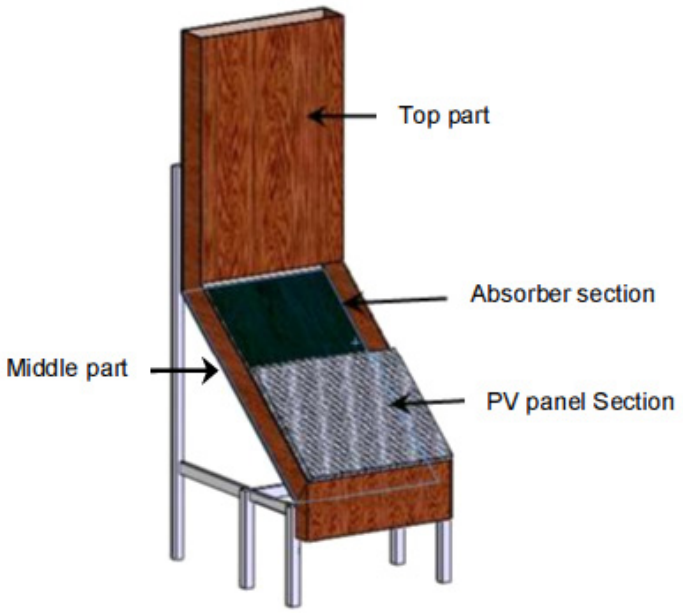


Figure 1. Solar cooling chimney

The middle part of Solar cooling chimney consists PV panel section and absorber section and designed such that its length is twice as long as the PV panel length. As shown in Figure 2 PV panel is mounted on the lower half of middle part while a transparent acrylic sheet is placed on the top half of the middle part. Middle part of solar cooling chimney is mounted at an inclination angle equal to local latitude to make sure that the PV panel and absorber section receives maximum solar radiation throughout the year.

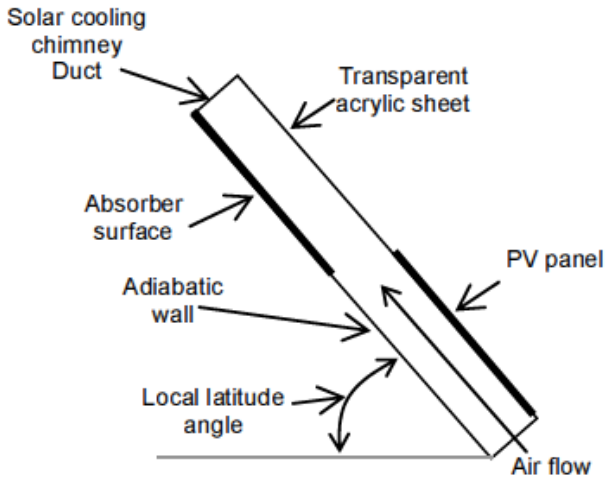


Figure 2. Middle section of solar cooling chimney

Most of the incident solar radiations will pass through the transparent acrylic sheet depending on its emissivity and reflectivity and fall on absorber surface. Absorber surface is painted black to maximize the amount of energy absorbed. As a result the temperature of the absorber surface will increase. High temperature absorber surface will transfer the heat to the air coming from the PV panel section.

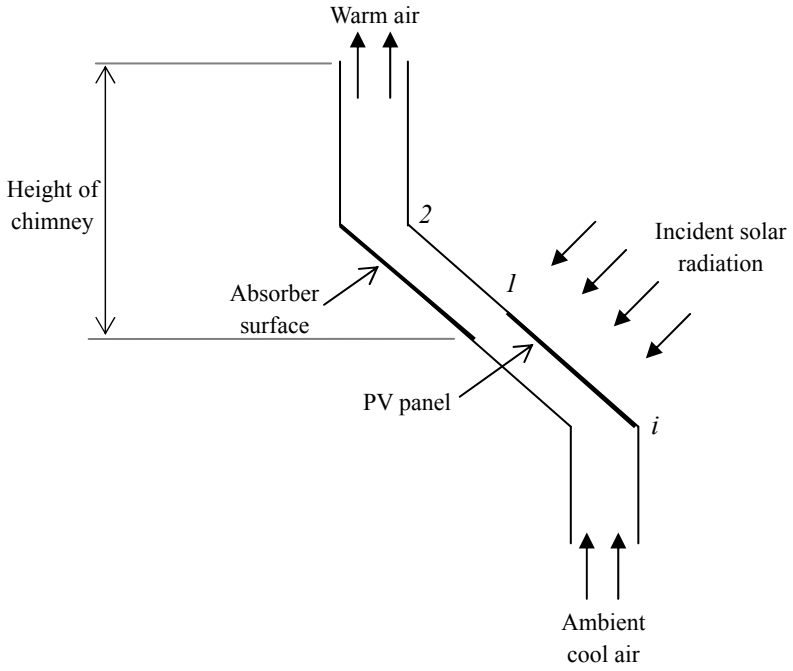


Figure 3. Solar cooling chimney dimension

Warm air with lesser density will rise upwards towards the top section. Height of the top part of SCC will assist the warm air to rise toward the top due to the static head difference. Air flow velocity will increase if we increase the height of chimney. The natural air draft in the duct induced due to the density difference between the air in the middle part and ambient air above the top part will create the suction to allow ambient air to enter from bottom of SCC as shown in Figure 3. Ambient cool air entering from the bottom of SCC will flow over the heated PV panel surface and reduce its temperature.

3. Model Theoretical Analysis

The following section presents the theoretical analysis of the solar cooling chimney as shown in Figure 4. The governing equations are derived using combination of conservation of energy and conservation of mass and momentum across the duct of the Solar Cooling Chimney to predict the values air velocities at the outlet. Energy balance equations for two separate sections in the solar cooling chimney are presented further. Figure 4 shows the PV panel section in which the PV panel is mounted on the inclined duct of the solar cooling chimney.

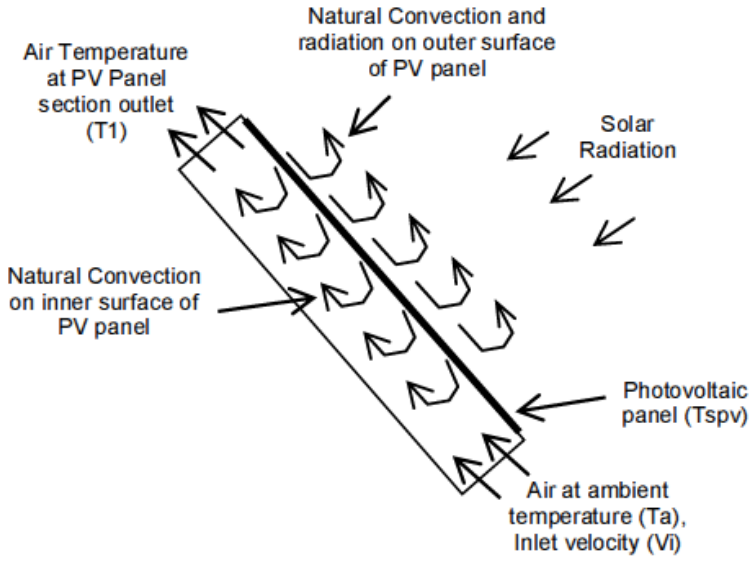


Figure 4. Heat transfer over the PV panel surface

PV panel section (point *i* to point 1)

Energy received by PV panel:

$$E_{in-pv} = I \times A_{pv} \times \left(1-\eta_{pv}\right) \times \tau_{pv} \tag{1}$$

Heat is transferred from outer and inner surface of the PV panel. Heat is transferred from outer of the PV panel via natural convection and radiation. Heat is transferred from inner surface via forced convection and radiation.

$$E_{loss-pv} = h_{nat} A_{pv} \left(T_{spv} - T_a\right) + h_{for} A_{pv} \left(T_{spv} - T_a\right) + \varepsilon\sigma A_{pv} \left(T_{spv}^4 - T_a^4\right) \tag{2}$$

Nusselt number equation for natural convection over inclined plate (Churchill & Chu, 1975):

$$Nu = \left(0.825 + \frac{0.387 Ra^{1/6}}{\left[1+(0.492/\mathrm{Pr})^{9/16}\right]^{8/27}} \right)^2 \tag{3}$$

Nusselt number equation for forced convection (Churchill and Ozoe, 1973):

$$Nu = \frac{0.3387\,\mathrm{Pr}^{1/3}\,\mathrm{Re}^{1/2}}{\left[1+\left(0.0468/\mathrm{Pr}\right)^{2/3}\right]^{1/4}} \tag{4}$$

The Reynolds number, Re is a function of the inlet velocity, V_i which is unknown.

The surface temperature T_{spv} is found by equating Equation (1) and Equation (2):

$$h_{nat}\left(T_{spv} - T_a\right) + h_{for}\left(T_{spv} - T_a\right) + \varepsilon\sigma\left(T_{spv}^4 - T_a^4\right) = I \times \left(1-\eta_{pv}\right) \times \tau_{pv} \tag{5}$$

Equation (5) is a non-linear equation.

The temperature of the air at outlet of the PV section T_1 can be determined using the log mean temperature difference between the ambient temperature and the PV surface temperature.

$$T_1 = \frac{T_{spv} - T_a}{\ln\left(\frac{T_{spv}}{T_a}\right)} \tag{6}$$

The density of moist inlet air is calculated by

$$\rho_i = \rho_{dry,i} + \rho_{v,i} \tag{7}$$

where density of dry air and water vapor are both estimated by ideal gas law.

Here, the partial pressure of water vapor can be determined by

$$P_{v,i} = \varphi_i P_{sat,i} \tag{8}$$

while the partial pressure of dry air be determined by

$$P_{dry,i} = P - P_{v,i} \tag{9}$$

where Φ_i is the relative humidity of air at inlet of PV panel, $P_{sat,i}$ is the partial pressure of vapor for saturated air at that temperature, and P is the absolute pressure.

Similar procedure is used to calculate density of air at outlet of PV panel, with

$$\frac{P_{v,i}}{P_{v,1}} = \frac{T_i}{T_1} \tag{10}$$

In the duct, pressure difference will be created by the warmer air at the inlet and cooler air at the outlet. The natural draft pressure caused by the difference in outside and inside air density is given by

$$\Delta P_s = (\rho_i - \rho_1) \times g \times H_c \tag{11}$$

where H_c is the height of chimney duct and g is the gravitational acceleration (9.81 m/s^2).

The total pressure loss in a duct can be calculated by

$$\Delta P_{lossPV} = \left(\frac{fL}{D_h} \times \frac{\rho_{avg} V_1^2}{2}\right) + \left(\frac{k \rho_{avg} V_1^2}{2}\right) \tag{12}$$

By equating Equation(11) and Equation(12), we can determined the velocity, V_1 of the air.

$$V_1 = \sqrt{\frac{(\rho_i - \rho_1) \times g \times H_c}{\frac{\rho_{avg}}{2}\left(\frac{fL}{D_h} + k\right)}} \tag{13}$$

Absorber section (point 1 to point 2)

Energy absorbed by absorber surface:

$$E_{in-abs} = I \times A_{abs} \times \tau_{trans} \times \tau_{abs} \tag{14}$$

Energy loss of the absorber:

$$E_{loss-abs} = h_{conv} A_{abs} \left(T_{sa} - T_1\right) + \varepsilon \sigma A_{abs} \left(T_{sa}^4 - T_1^4\right) \tag{15}$$

h_{conv} is calculated from Nusselt number for forced convection (Churchill & Ozoe, 1973).

$$Nu = \frac{0.3387\,\mathrm{Pr}^{1/3}\,\mathrm{Re}^{1/2}}{\left[1+\left(0.0468/\mathrm{Pr}\right)^{2/3}\right]^{1/4}} \tag{16}$$

Similar to the PV panel section, T_{sa} can be found by equating Equation(14) and Equation(15).

The log mean temperature of the absorber section is

$$T_2 = \frac{T_{sa} - T_1}{\ln\left(\frac{T_{sa}}{T_1}\right)} \tag{17}$$

Similar procedure is used to calculate the velocity of air, V_2 at this section.

$$V_2 = \sqrt{\frac{(\rho_1 - \rho_2)\times g \times H_c}{\frac{\rho_{avg}}{2}\left(\frac{fL}{D_h}+k\right)}} \tag{18}$$

Conservation of mass principle is used to calculate V_i. Mass balance across inlet (i) to outlet of absorber (2) is

$$\rho_2 V_2 A_{duct} = \rho_i V_i A_{duct} \tag{19}$$

$$V_i = \frac{\rho_2 V_2}{\rho_i} \tag{20}$$

The value of V_i from Equation(20) is now used in the calculation of Equation(4) in the PV section where heat transfer coefficient of forces convection can now be calculated. Iteration is continues until V_i is converged. However, since the h_c is rather small (comparative with h_n), so only one or two iteration is required (expected result).

4. Results and Discussion

In this section the theoretical solar cooling chimney model analysis as shown in Figures 5 to 7. In Figure 5, the total static pressure difference generated between the inlet and outlet of solar cooling chimney is plotted against height of chimney. It can be prominently seen that the static pressure difference across the inlet and outlet of the solar cooling chimney goes on increasing as the height of chimney increases. Temperature of outlet air has a similar effect on the static pressure of air across solar cooling chimney. This static pressure difference between the inlet and outlet of the solar cooling chimney does not behave like velocity in the Figure 7 since the major losses due to friction and minor losses due to the shape of the chimney are not includes in this static pressure.

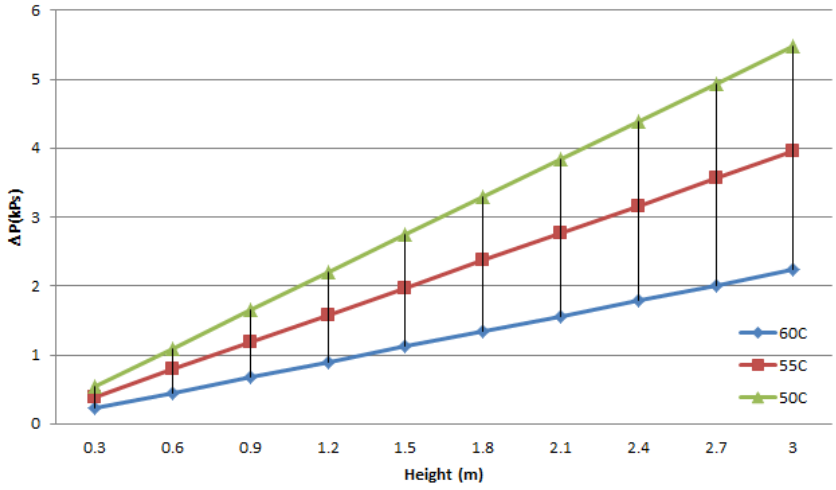


Figure 5. Static pressure difference between inlet and outlet of solar cooling chimney with respect to height

Figure 6 provides us information about the outlet air velocities that could be achieved in a chimney with its height starting from 0.1m and increasing till 1m tall. The air velocities for outlet air temperature of between 323

K (50 °C) to 333 K (60 °C) are presented. It is observed as expected that the air velocity due to natural draft will go on increasing as the solar cooling chimney height increases. Similar to the height of solar cooling chimney, temperature of outlet air will have a similar effect on the velocity if outlet air. As the temperature of outlet air will increase the velocity of air will rise. It is interesting to see that these curves have a logarithmic trend. It can be predicted that the velocity rise of air will be stagnated after certain increase of height of solar cooling chimney for different outlet air temperatures. The optimum solar cooling chimney height could be determined if this curve is extended for higher chimney heights.

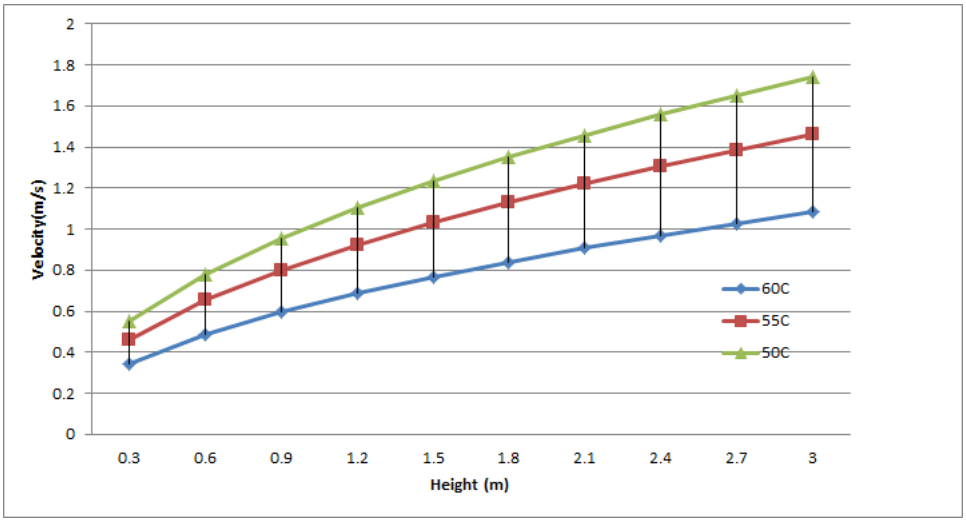


Figure 6. Air velocity in the solar cooling chimney with respect to the height of solar cooling chimney

Figure 7 illustrates the predicted surface temperature of photovoltaic panel attached with solar cooling chimney. These temperature values are determined using the velocity of air predicted in the earlier section. The natural draft in the duct will pull the ambient air from the bottom of the duct. Ambient air in this case is considered to be at 20 °C, which will flow over the back heated surface of the photovoltaic panel with the predicted velocity. This will cause the convection heat transfer from the photovoltaic panel. Nusselt number correlation for forced convection with either laminar or turbulent flow (Whitaker, 1972) is used to determine the convection heat transfer coefficient. Further the convection heat transfer coefficient along with Newton's law of cooling is used to predict the temperature of the photovoltaic panel. The temperature of the photovoltaic panel goes on reducing as the chimney height increases. This can be justified from the earlier graph in Figure 7, since the velocity of air increases it will increase the convection heat transfer coefficient reducing the temperature of the photovoltaic panel.

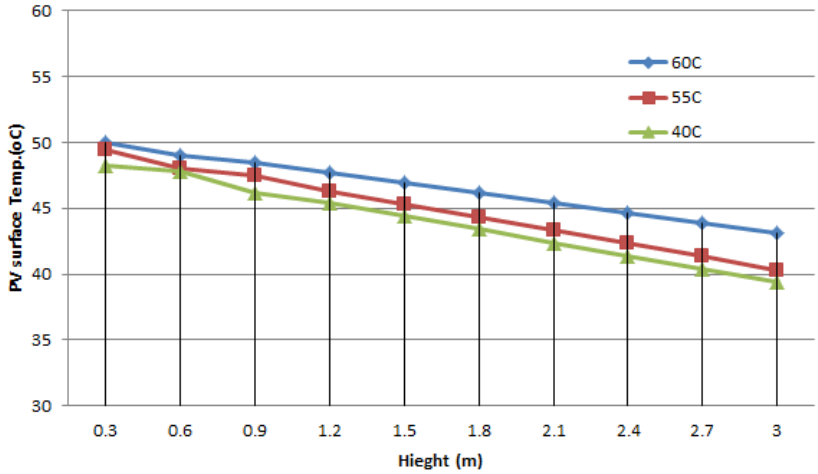


Figure 7. Predicted PV panel surface temperature with respect to change in height of solar cooling chimney

Figure 8 illustrates the improved voltage from the PV panel due to reduced surface temperature. As the solar cooling chimney height goes on increasing the output voltage from the PV panel will increase. In addition improving voltage will follow the outlet air velocity trend.

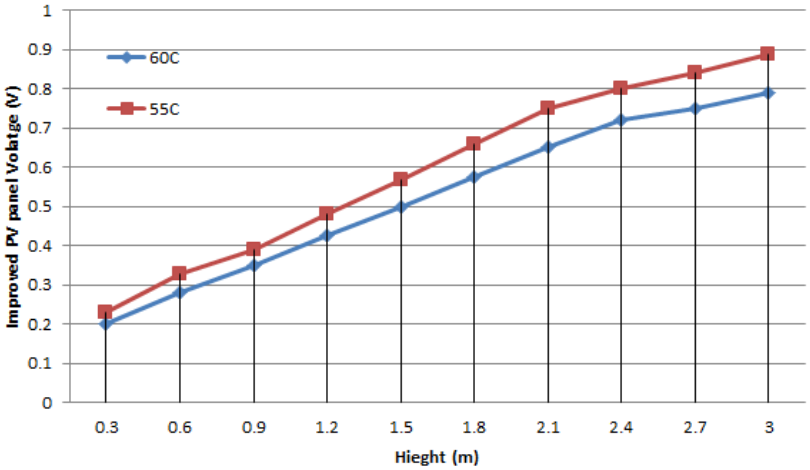


Figure 8. Effect of negative temperature coefficient on open circuit voltage of the PV panel

From above operation conditions, there is an optimum working temperature mentioned to achieve the said open circuit voltage. As the operating temperature of the PV panel goes on increasing the open circuit voltage will drop. This will reduce the efficiency of the PV panel as well.

5. Conclusion

The natural air draft that is achieved by buoyancy effects in a chimney can be used as a passive cooling medium for PV panels. Simple modification in the design of the system and including the additional absorber section can help to improve the induced natural draft of air and as a result helps improve the performance of the PV panel. Simple and preliminary analysis shows that by coupling the solar cooling chimney with a PV panel we can achieve considerable cooling of the PV panel and improve the efficiency of the PV panel. Taller chimneys can increase the velocity of air and further reduce the working temperature of PV panel to make it more efficient. Although taller chimneys will increase the overall cost of the system and reduce the simplicity in design. Taller chimney will also have a shading effect in case of PV panels are placed close to each other. Solar cooling chimney height should be optimized for the cost effectiveness and shading effect. Another way to improve the cooling / heat removal would be by increasing the surface area of the absorber section. Further investigation is required to accurately measure velocity of induced natural draft.

References

Akbarzadeh, A., Johnson, P., & Singh, R. (2009). Examining potential benefits of combining a chimney with a salinity gradient solar pond for production of power in salt affected areas. *Solar Energy, 83*, 1345-1359. http://dx.doi.org/10.1016/j.solener.2009.02.010

Anderson, W. G., Dussinger, P. M., Sarraf, D. B., & Tamanna, S. (2008). Heat pipe cooling of concentrating photovoltaic cells. *In:* Photovoltaic Specialists Conference, 2008. PVSC '08. 33rd IEEE. pp. 1-6. http://dx.doi.org/10.1109/PVSC.2008.4922577

Brinkworth, B. J. (2000). Estimation of flow and heat transfer for the design of PV cooling ducts. *SolarEnergy, 69*, 413-420. http://dx.doi.org/10.1016/S0038-092X(00)00082-7

Cengel, Y. A., Cimbala, J. M., & Turner, R. H. (2008). *Fundamentals of Thermal-Fluid Sciences*. McGraw-Hill Higher Education.

Churchil, S. W., & Ozoe, H. (1973). Correlations for laminar forced convection with uniform heating in flow over a plate and in developing and fully developed flow in a tube. *Journal of Heat Transfer-Transactions of the Asme, 95*, 78-84. http://dx.doi.org/10.1115/1.3450009

Churchill, S. W., & Chu, H. H. S. (1975). Correlating equations for laminar and turbulent free convection from a vertical plate. *International Journal of Heat and Mass Transfer, 18*, 1323-1329. http://dx.doi.org/10.1016/0017-9310(75)90243-4

Date, A. (2010). *Cooling of solar cells by chimney-induced natural draft of air*. 48th AuSES Conference SOLAR 2010. Canberra.

Haaf, W., Gmayr, K. F., & Schlaich, J. (1983). Solar chimneys Part I: principle and construction of the pilot plant in Manzanares. *International Journal of Solar Energy,* 3-20. http://dx.doi.org/10.1080/01425918308909911

Kermani, E., Dessiatoun, S., Shooshtari, A., & Ohadi, M. M. (2009). Experimental investigation of heat transfer performance of a manifold microchannel heat sink for cooling of concentrated solar cells. *In:* Electronic Components and Technology Conference, 2009. ECTC 2009. 59th. pp. 453-459.

King, D. L., Kratochvil, J. A., & Boyson, W. E. (1997). Temperature coefficients for PV modules and arrays: measurement methods, difficulties, and results. *In:* Photovoltaic Specialists Conference, 1997. Conference Record of the Twenty-Sixth IEEE. pp. 1183-1186.

Krauter, S., Araújo, R. G., Schroer, S., Hanitsch, R., Salhi, M. J., Triebel, C., & Lemoine, R. (1999). Combined photovoltaic and solar thermal systems for facade integration and building insulation. *Solar Energy, 67*, 239-248. http://dx.doi.org/10.1016/S0038-092X(00)00071-2

Mayhew, Y. R., & Rogers, G. F. C. (1995). *Thermodynamic and Transport Properties of Fluids*. Oxford: Blackwell Publishing Ltd.

Moshfegh, B., & Sandberg, M. (1996). Investigation of fluid flow and heat transfer in a vertical channel heated from one side by PV elements, part I - Numerical Study. *Renewable Energy, 8*, 248-253. http://dx.doi.org/10.1016/0960-1481(96)88856-2

Moshfegh, B., & Sandberg, M. (1998). Flow and heat transfer in the air gap behind photovoltaic panels. *Renewable and Sustainable Energy Reviews, 2*, 287-301. http://dx.doi.org/10.1016/S1364-0321(98)00005-7

Royne, A., Dey, C. J., & Mills, D. R. (2005). Cooling of photovoltaic cells under concentrated illumination: a critical review. *Solar Energy Materials and Solar Cells, 86*, 451-483. http://dx.doi.org/10.1016/j.solmat.2004.09.003

Sandberg, M., & Moshfegh, B. (1996). Investigation of fluid flow and heat transfer in a vertical channel heated from one side by PV elements, part II - Experimental study. *Renewable Energy, 8*, 254-258. http://dx.doi.org/10.1016/0960-1481(96)88857-4

Sandberg, M., & Moshfegh, B. (2002). Buoyancy-induced air flow in photovoltaic facades: Effect of geometry of the air gap and location of solar cell modules. *Building and Environment, 37*, 211-218. http://dx.doi.org/10.1016/S0360-1323(01)00025-7

Schlaich, J. (1995). *The Solar Chimney - electricity from sun*. Berlin: Edition Axel Menges.

Tonui, J. K., & Tripanagnostopoulos, Y. (2008). Performance improvement of PV/T solar collectors with natural air flow operation. *Solar Energy, 82*, 1-12. http://dx.doi.org/10.1016/j.solener.2007.06.004

Whitaker, S. (1972). Forced convection heat transfer correlations for flow in pipes, past flat plates, single cylinders, single spheres, and for flow in packed beds and tube bundles. *AIChE Journal, 18*, 361-371. http://dx.doi.org/10.1002/aic.690180219

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

An Analysis of the Environmental Vulnerability Index of a Small Island: Lipe Island, Kho Sarai Sub-District, Mueang District, Satun Province, Thailand

Nutsurang Pukkalanun[1], Wasin Inkapatanakul[1], Chucheep Piputsitee[2] & Kasem Chunkao[1]

[1] Faculty of Environment, Kasetsart University, Bangkok, Thailand

[2] Faculty of Economic, Kasetsart University, Bangkok, Thailand

Correspondence: Nutsurang Pukkalanun, Faculty of Environment, Kasetsart University, Bangkok 10900, Thailand. E-mail: to_gae@hotmail.com

This research was financed by the King's Royally Initiated Laem Phak Bia Environmental Research and Development Project

Received: December 16, 2012 Accepted: January 10, 2013 Online Published: January 21, 2013

doi:10.5539/mas.v7n2p33 URL: http://dx.doi.org/10.5539/mas.v7n2p33

Abstract

Thailand is located in South East Asia and is a popular tourist destination. It is rich in both natural resources and culture. There are 691 islands in Thailand, and more than 214 of these islands are used for tourism. Koh Lipe is very Small Island of approximately 2 square kilometers, located in Talutao National Park in the southern part of Thailand. This research aims to assess the sensitivity of the Island in terms of tourism development by using the Environmental Vulnerability Index, or EVI. The results showed that the EVI of Lipe Island is approximately 5.7, which represents a very high vulnerability score. Particularly, the REI, the level of risk to hazard, which measures influences on the environment within the island (e.g., loss of forestry, tourist accommodation, waste water and solid waste) was approximately 6.2, while the EDI, the natural resilience of the state based on its native characteristics, (e.g. water resources, protected area, marine protected area, and law enforcement), was approximately 5.7. This is also a very important indicator of the vulnerability of the Island. Thus, to reduce the overall vulnerability of the island, all indicators included in the REI and the EDI must become management priorities. Over time, this will increase the immunity of the island to of the impact of tourism development.

Keywords: environment vulnerability, small island, Thailand

1. Statement of the Problem

Thailand is located in South East Asia where highly potential of tourism destinations in both natural resources and culture. Moreover, tourism generates enormous income and creates economic activities, for instance, wage, employment and creates business enterprise such as hotel, transportation and tourist agency etc. However, there are some tourism activities still face the problems in term of management likes waste, garbage, water and air pollution which effect to natural degradation, especially, the Island where distance from the main land and difficult accessibility.

There are 691 islands in Thailand and more than 214 islands are used for tourism destinations (The King's Royally Initiated Laem Phak Bia Environmental Research and Development Project, 2009). The utilization of natural resources for tourism destinations without any evaluates the potential of island in order to generate tourism activities; this will produce the problem in economics, environmental and social aspects. Furthermore, an island is also in the ecology fragile system (The King's Royally Initiated Laem Phak Bia Environmental Research and Development Project, 2009), therefore, tourist activity might harmful for island structure and this will change function of the island which is hard to recovery.

Koh Lipe or Lipe Island is located in Satun province, Southern part of Thailand. In the last few years, Koh Lipe has been grown rapidly in tourism sector. Unfortunately, the consequence of dramatic developing in small Island generate both social and environment problems for instance local people loss of their land as well as huge of wastes and garbage from tourist sectors on the island. Therefore, this research aims to identify the sensitive area

particularly a small island by using environmental vulnerability index (EVI). To evaluate the island vulnerability in order to awareness that activity which takes place on the island might cause the environment degradation.

2. Materials and Method

2.1 Study Area

Lipe Island is Located in Talutao national park, Andaman sea, Southern part of Thailand. It is approximately 2 kilometer square and distance from main land, Satun province, about 70 kilometer. The characteristic of Lipe similar to J or upside down L. There are 187 households and the population is about 983 which are divided to men and men equally. Before Lipe Island become an attractive place of tourist destination, almost of men in the village are fishermen and women are housewife. In high season of tourism, some of fishermen work as a transfer boat from speech boat to each resort and this activity become local career.

Lipe Island is not connected beach to each other. There are 3 mains beach on the Island, Bandaya, sunrise and sun set beach. Because of monsoons from the Andaman sea, Lipe Island has only operate its tourism industry 6 months out of the year, from November to April.

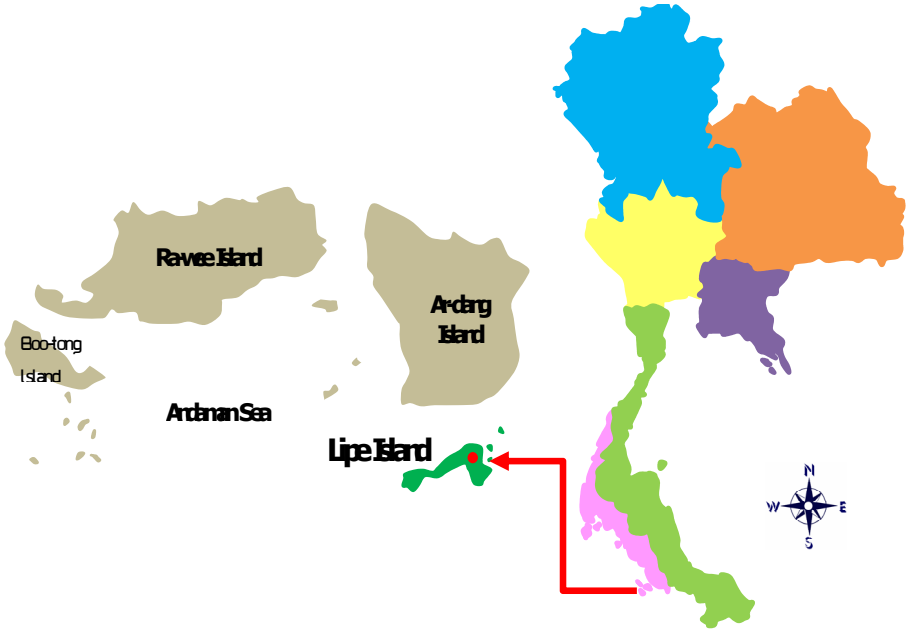


Figure 1. The location of Lipe Island

2.2 Method

In the past, the development of vulnerability indices focused on economic and social aspects which related to human dimensions. It has been clearly seen that environmental aspect was different from vulnerability of human systems because environment system is more complicated in ecological system (Kaly, 2000; Gowrie, 2003; Jens et al., 2006). However, in 1999, the Environmental Vulnerability Index (EVI) for Small Island developing States (SIDS) was developed by South Pacific Applied Geoscience Commission (SOPAC) which funded by New Zealand government. This was an initiative attempt to create environmental index and 54 indicators were built which three aspects of environmental vulnerability index were constructed. In this research, 18 indicators were used to measure the vulnerability of Lipe Islan as follow:

1) The Intrinsic Resilience sub-Index (IRI) that measures of natural resilience

$$IRI = \frac{Sc_1Wt_1 + Sc_2Wt_2 + ... + Sc_4Wt_4}{Wt_1 + Wt_2 + ...Wt_4} \quad (1)$$

2) The Environmental Degradation Sub-Index (EDI) which evaluates the degree of damage

$$EDI = \frac{Sc_5Wt_5 + Sc_6Wt_6 + ... + Sc_9Wt_9}{Wt_5 + Wt_6 + ...Wt_9} \quad (2)$$

3) The risk Exposure sub-index (REI) that illustrated the level of the risk natural hazard and human activities which the formula is as follow:

$$REI = \frac{Sc_{10}Wt_{10} + Sc_{11}Wt_{11} + ... + Sc_{18}Wt_{18}}{Wt_{10} + Wt_{11} + ...Wt_{18}} \quad (3)$$

4) The Environmental Vulnerability Index (EVI) which identifies

$$EVI = \frac{Sc_1Wt_1 + Sc_2Wt_2 + ... + Sc_{18}Wt_{18}}{Wt_1 + Wt_2 + ...Wt_{18}} \quad (4)$$

Where

Sc = Scoring

Wt = Weighting

Scoring is the standard universal number of each index which represents the level of Environmental vulnerability from 1-7 (less to more) (see Kaly, 1999)

Weighting is identified by the scholar who works in LipeIsland more than 10 year which rank from 1-3 (less to more) to illustrate of the important of each index in Lipe island.

From the calculating of environmental vulnerability index, the result of calculating can be identified the ranking of vulnerability of small island and the efficiency of environmental management which showed the table as follow:

Table 1. The result of calculating and evaluating of IRI, EDI, REI and EVI

Index	Ranking		
	0-2.00	2.01-4.00	4.01-7.00
IRI	Low natural and manmade resilience	Medium natural and manmade resilience	High natural and manmade resilience
EDI	High resistance and low degradation	Medium resistance and medium degradation	Low resistance and high degradation
REI	Low of risk exposure	Medium of risk exposure	High of risk exposure
EVI	Low Environmental Vulnerability	Medium Environmental Vulnerability	High Environmental Vulnerability

Source: Coastal Habitats and Resources Management: CHARM 2007

3. Results and Discussion

Koh Lipe is very SmallIsland and approximately 2 square kilometer. It is located in Talutoa national park in the south of Thailand. The rapid development of Koh lipe particularly tourism sector generates environment problem such as solid waste, water pollution and environmental degradation. This research identify Koh Lipe island is a sensitive area which is fragile ecosystem. 18 indicators of environmental vulnerability index were represented the vulnerability of Koh Lipe which divided into 3 categories, the Intrinsic Resilience sub-Index (IRI), the Risk Exposure sub-index (REI), and the Environmental Degradation Sub-Index (EDI). The Intrinsic Resilience sub-Index (IRI) consists of land area, Low land, Habitats diversity and Relief and most of indicators are related the physical of the island which means the natural resilience of the state based on its native characteristics. While the Environmental Degradation Sub-Index (EDI) is fisheries, water resource, protected area, marine protected area and law enforcement and almost of indicators are about the level of risk to hazard which influence to the environment within the island. Likewise, Risk Exposure sub-Index (REI) is dry period, wet period, tsunami, population density, loss of forestry, tourism accommodation, source of tourism, waste water, solid waste which incorporates measures of the level of risk to hazard. The result of calculating the vulnerability of Koh Lipe has been shown in the Table 2.

Table 2. The result of calculating the REI, IRI, EDI and EVI of Koh Lipe

No.	Index	Weighting (W)	Factors	Detail	Raw score	Scoring (S)	WxS
1.	IRI	3	Land area	Total land area (sq km)	1.902	7	21
2.	IRI	2	Relief	Altitude range (highest point subtract the lowest point in the Island)	110	7	14
3.	IRI	2	Lowlands	Percent of land area less than 10 meters above sea level	80	7	14
4.	IRI	3	Habitats diversity	The variety of coral leafs , mangroves, and seagrass	1	1	3
		10		**IRI = 5.2**			**52**
5.	EDI	3	Fisheries	The number of fishing boats less than 10 meters	≤500	1	3
6.	EDI	3	Water resource	Mean percentage of water usage per year (surface water and underground water)		7	21
7.	EDI	2	Protected area	Percentage of protected area on the Island	≤20	7	14
8.	EDI	3	Marine protected area	Percentage of marine protected area surrounding the Island	≤20	7	21
9.	EDI	3	Law enforcement	The enforcement or management of environmental issues	non	7	21
		14		**EDI = 5.7**			**80**
10.	REI	1	Dry period	Number of months over the last five years during which rainfall is greatest than 20 percent lower than the 30 year average.	10	2	2
11.	REI	2	Wet period	Number of months over the last five years during which rainfall was greatest than 20 percent higher than the 30 year average.	50	7	14
12.	REI	1	Tsunamis	Percentage of risky of tsunamis occurs on the island	≤20	1	1
13.	REI	1	Population density	Total human population density (number per square kilometer land area)	20	1	1
14.	REI	3	Loss of forestry	Percentage of land area change by removal of forest over the last 5 years	>60	7	21
15.	REI	3	Tourist accommodation	Number of rooms (for tourists)	>1,100	7	21
16.	REI	3	Sources of tourism	Potential of source of tourism evaluated by Tourism Authority of Thailand	>50	7	21
17.	REI	3	Waste water	Ratio of waste water from industries and communities	>7,000	7	21
18.	REI	3	Solid waste	The average amount of solid waste per year	>120	7	21
		20		**REI = 6.2**			**123**
				EVI = 5.7			

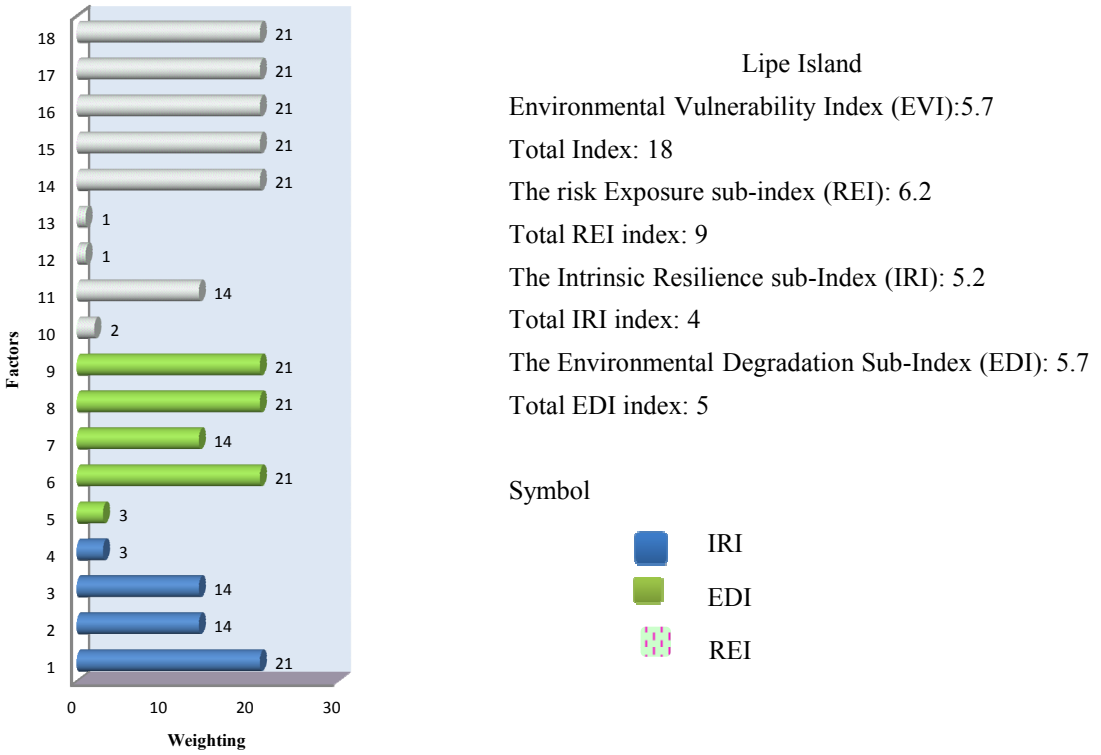


Figure 2. The scoring of IRI, EDI, REI and EVI of Lipe Island

From the Figure 2, it has been clearly seen that Lipe Island was extremely vulnerable area by calculating of EVI at 5.7. The result has shown that the Intrinsic Resilience Index (IRI) which consist of 4 factors, land area, relief, low land and habitats diversity, was 5.2. It was identified that this intrinsic Resilience of the island was highly vulnerability. Almost of factors are physical aspects therefore, to protect the island should consider the human activity which is not harmful for environmental degradation. Today Lipe is attractive tourist destination and welcome tourist all around the world, to generate tourism activity should appropriate to the function of the island. Moreover, the Environmental Degradation Index (EDI) has illustrated the ability of natural resistance from natural disaster in the future. The factors were fisheries, water resource, protected area, marine protected and law enforcement. The result of EDI was about 5.7 and almost of factors related to activities was taking place on the island. Thus, to enhance the ability of natural resistance was to manage the activity in the island such as increase the protected area in both marine and on the island. Furthermore, small island of Lipe need restrict law enforcement to against illegal activities like water treatment and fishery. Moreover, the highest score of vulnerability of Lipe Island was 6.2, which was about Risk Exposure Index: REI. It contained of dry period, wet period, tsunami, and population density, loss of forest, tourism accommodation, and source of tourism, waste water, and solid waste. The factors emphasize on the impact of both natural condition and human activities. Some of factors of the risk exposure index were related to human activities like tourism destination. The number of tourist accommodation, waste water and solid waste were harmful for ecological system of Small Island. It would be generated the following problem and still exist in the island for instance, overuse of natural resources, deforestation, and water pollution etc. Almost of the scoring to identify the vulnerability was about 7. To adjust the score to other ranking means factors need to be well management and the responsibility of all stakeholders who get the benefit from island should concern the activities taking place on the island.

4. Conclusion

It can be clearly seen that Lipe Island is vulnerable area. From the result of EVI has shown that 3 Indices and 18 factors represent the vulnerability of the island and about 13 factors reach the score about 7. Some of factors cannot adjust the scoring because it contains of physical aspects. However, some of factors need to be well management like water resource, law enforcement, protected area, loss of forestation, tourist accommodation, waste water and solid waste. To generate immune and reduce the vulnerability for the island and reach the sustainable development, all of factors must restrict implementation. However, the limitation of EVI evaluation

is not include social dimension, thus for the next research should focus on the impact of island development into the local or community who live on the Island.

References

Coastal Habitats and Resources Management. (2005). *Vulnerability Mapping and Quality Status of Phang Nga and Ban Don Bay in Southern Thailand Project.* Vilailuck University, Thailand

Coastal Habitats and Resources Management: CHARM. (2007). *Ban Don Bay and its offshore Island Management planning Project: Analysis and Diagnosis of the Coastal Production Systems.* Vilailuck University, Thailand

Jens, K., Grete, S., & Lars, E. (2006). Assessing environment vulnerability in EIA-The content and context of the Vulnerability concept in an alternative approach to standard EIA procedure. *Environmental Impact Assessment Review, 26*, 511-527. http://dx.doi.org/10.1016/j.eiar.2006.01.003

Kaly, U. L., Briguglio, L., McLeod, H., Schmall, S., Pratt, C., & Pal, R. (1999). *Environmental Vulnerability Index (EVI) to summarize national environmental vulnerability profile.* SOPAC Technical Report 275. Retrieved from http://www.sopac.int/data/virlib/TR/TR0275.pdf

Kaly, U. L., & Pratt, C. (2000). *Environmental Vulnerability Index: development and provisional indices for Fiji, Samoa, Tuvalu, and Vanuatu.* SOPAC Technical Report 306. Retrieved from http://www.unescap.org/STAT/envstat/stwes-evi.pdf

Marissa, N. G. (2003). Environmental Vulnerability Index for the Island of Tobago, West Indies. *Conservation Ecology, 7*(2), 11. Retrieved from http://www.consecol.org/vol7/iss2/art11

The King's Royally Initiated Laem Phak Bia Environmental Research and Development Project. (2009). The report of Environmental quality on Lipe Island. Kasetsart university, Bangkok.

The Southern Study Limited. (2009). *The final report of solid waste and waste water management in Lipe Island.* Resources and Environment office, Satun province. Thailand.

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

Peat Water Treatment Using Combination of Cationic Surfactant Modified Zeolite, Granular Activated Carbon, and Limestone

S. Syafalni[1], Ismail Abustan[1], Aderiza Brahmana[1], Siti Nor Farhana Zakaria[1] & Rohana Abdullah[1]

[1] School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal, Penang, Malaysia.

Correspondence: S. Syafalni, School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300, Penang, Malaysia. E-mail: cesyafalni@eng.usm.my

Received: December 3, 2012 Accepted: January 14, 2013 Online Published: January 22, 2013

doi:10.5539/mas.v7n2p39 URL: http://dx.doi.org/10.5539/mas.v7n2p39

Abstract

This research was conducted essentially to treat fresh peat water using a series of adsorbents. Initially, the characterization of peat water was determined and five parameters, including pH, colour, COD, turbidity, and iron ion exhibited values that exceeded the water standard limit. There were two factors influencing the adsorption capacity such as pH, and adsorbent dosages that were observed in the batch study. The results obtained indicated that the majority of the adsorbents were very efficient in removing colour, COD, turbidity at pH range 2-4 and Fe at pH range 6-8. The optimum dosage of cationic surfactant modified zeolite (CSMZ) was found around 2 g while granular activated carbon (GAC) was exhibited at 2.5 g. In column study, serial sequence of CSMZ, GAC, and limestone showed that the optimal reduction on the 48 hours treatment were found pH = 7.78, colour = 12 TCU, turbidity = 0.23 NTU, COD = 0 mg/L, and Fe= 0.11 mg/L. Freundlich isotherm model was obtained for the best description on the adsorption mechanisms of all adsorbents.

Keywords: cationic surfactant modified zeolite, granular activated carbon, limestone, peat water

1. Introduction

Water is essential and fundamental to all living forms and is spread over 70.9% of the earth's surface. However, only 3% of the earth's water is found as freshwater, of which 97% is in ice caps, glaciers and ground water (Bhatmagar & Minocha, 2006). In Malaysia, more than 90% of fresh water supply comes from rivers and streams. The demand for residential and industrial water supply has grown rapidly coupled with an increase in population and urban growth (WWF Malaysia, 2004). Water demand in affected populations such as rural areas also demands that attention is paid to providing more sustainable solutions rather than transporting bottled water (Loo et al., 2012). For this reason, it is essential to ensure availability of local sources of water supply and even develop new potential sources of water such as from peat swamp forest to overcome future water shortages.

River water surrounded by peat swamp forest is defined as peat water and is commonly available as freshwater since it has a low concentration of salinity. The previous study shows that peat swamp forest has high levels of acidity and organic material depending on its region and vegetation types (Huling et al., 2001). Under natural conditions, tropical peat lands serve as reservoirs of fresh water, moderate water levels, reduce storm-flow and maintain river flows, even in the dry season, and they buffer against saltwater intrusion (Wosten et al., 2008).

Due to the acidity and high concentration of organic material, selective treatment of peat water must be conducted prior to its use as water supply. Recently, many methods have been designed and have proven their effectiveness in treating raw water such as coagulation and flocculation (Franceschi et al., 2002; Liu et al., 2011; Syafalni et al., 2012a), absorption (Ćurković et al., 1997), filtration (Paune et al., 1998) and combining (Hidaka et al., 2003). Careful consideration of the most suitable method is important to ensure that the adsorption process is the most beneficial, economically feasible method as well as easy to operate for producing high quality of water in a particular location.

Many researchers have shown that activated carbon is an effective adsorbent for treating water with high concentrations of organic compounds (Eltekova et al., 2000; Syafalni et al., 2012b). Its usefulness derives mainly from its large micropore and mesopore volumes and the resulting high surface area (Fu & Wang, 2011). However, its high initial cost makes it less economically viable as an adsorbent. Low cost adsorbent such as zeolite nowadays has been explored for its ability in many fields especially in water treatment. Natural zeolite has negative surface

charge which gives advantages in absorbing unwanted positive ions in water such heavy metal. These ions and water molecules can move within the large cavities allowing ionic exchange and reversible rehydration (Jamil et al., 2010). The effectiveness of zeolite has been improvised by modified zeolite with surfactant in order to achieve higher performance in removing organic matter (Li & Bowman, 2001). Among tested cationic surfactants, hexa-decyl-tri-methyl ammonium (HDTMA) ions adsorbed onto adsorbent surfaces are particularly useful for altering the surface charge from negative to positive (Chao & Chen, 2012). Surfactant modified zeolite has been shown to be an effective adsorbent for multiple types of contaminants (Zhaohu et al., 1999).

Zeolite is modified to improve its capability of exchanging the anion by cationic surfactants, called CSMZ. CSMZ adsorbs all major classes of water contaminants (anions, cations, organics and pathogens), thus making it reliable for a variety of water treatment applications (Bowman, 2003). Nowadays, interest in the adsorption of anions and neutral molecules by surfactant modified zeolite has increased (Zhang et al., 2002). Modification of zeolite by surfactant is commonly done by cationic or amphoteric surfactants. By introducing surfactant to the zeolite, an organic layer is developed on the external surfaces and the charge is reversed to positive (Li et al., 1998). However, the present study used zeolite that had been modified using Uniquat (QAC-50) as cationic surfactant (CSMZ) and their performance towards the removal of color, COD, turbidity and iron ion from peat water were investigated.

2. Materials

Four adsorbents were used in these experiments which are natural zeolite, zeolite modified by cationic surfactant, activated carbon and limestone. All adsorbents were prepared with equivalent sizes of 1.18 mm - 2.00 mm. Hydrochloric acid (HCl) and sodium hydroxide (NaOH) were used for polishing zeolite during the preparation phase and for pH adjustment of the sample. Furthermore, potassium dichromate (K_2CrO_7), silver sulphate (Ag_2SO_4), sulphuric acid (H_2SO_4) and mercury (II) sulphate ($HgSO_4$) were used as digestion solution reagents and acid reagents for COD analysis. Lastly, Uniquat (QAC-50) was used as cationic surfactant to modify the zeolite.

2.1 Preparation of Surfactant Modified Zeolite

In these studies, 100 g of prewashed natural zeolite was contacted with 5.6 ml/l Uniquat (QAC-50) as cationic surfactant (CSMZ). The mixture was then stirred at room temperature for 4 hours at 300 rpm (Karadag et al., 2007). The zeolite then was filtered and washed with distilled water several times. After that, the absorbent was dried in an oven at a temperature of 105 °C for 15 hours.

2.2 Test Procedures

2.2.1 Batch Studies

Serial batch studies were conducted at room temperature (28 ± 1 °C) to investigate the influence of pH and dosage for removing colour, COD, turbidity and iron ion from peat water. Shaking speed of 200 rpm for 20 minutes were fixed and operated respectively. A working volume of 150ml peat water sample was set up in 250 ml conical flasks. Preceding the batch studies, initial concentration for those parameters was determined. The optimum pH and dosage of absorbent were determined. Subsequently, the percentage of removal was finally determined, plotted, and compared.

2.2.2 Batch Column Studies

Column studies were carried out using a plastic column with dimensions: 5.4 cm diameter and 48 cm length. Three adsorbents were filled inside the column at a specific depth with the supporting layers of marbles, cotton wool, and perforated net. Total volume of 2000 ml peat water was pumped in the up flow mode from the vessel into the column by using a Masterflex peristaltic pump at a minimum flow rate of (30, 60, 90) ml/min. In this study, however, column studies were performed un-continuously (batch) due to limitations of time. All parameters related to the column design are summarized in the following Table 1.

Table 1. Column studies parameters

Parameters	Unit	Value
Diameter, D	cm	5.4
Horizontal Surface Area, A	cm^2	22.9
Column volume, V	cm^3	1099.3
Flowrate, Q	ml/min	30, 60, 90
Surface Loading Rate, SLR= Q/A	cm/min	1.31, 2.62, 3.93

The serial sequence arrangements of adsorbents were conducted as shown in Figure 1 below. Effluent samples were collected at various time intervals, whilst maintaining room temperature, and analysed.

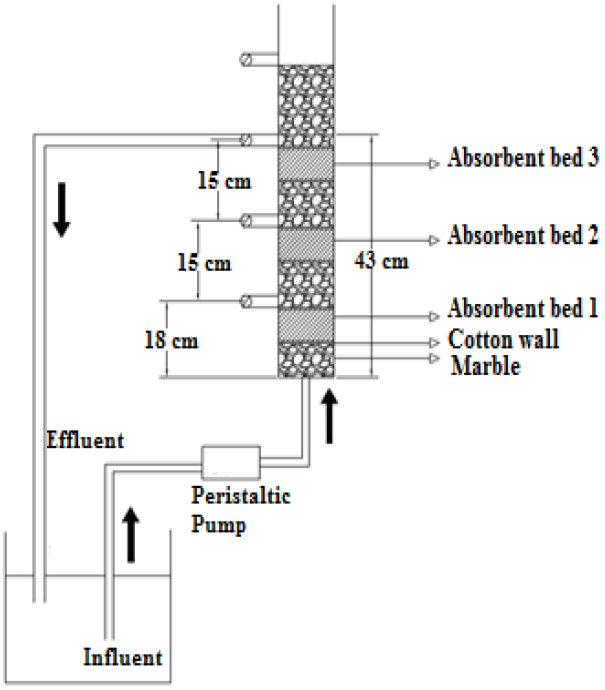


Figure 1. Schematic diagrams of lab-scale column studies

3. Results and Discussion

3.1 Peat Water Characterization

Surface water originating from the peat swamp forest was taken from Beriah peat swamp river along the Kerian River on several occasions as the main sample. The characterization of peat water was carried out at the sampling point (in-situ measurement) using a multi-parameter probe as well as in the environmental laboratory of civil engineering, USM. Fundamentally, the characterization procedures were based on the Standard Methods for the Examination of Water and Wastewater (APHA, 1992). Table 2 represents the peat water characteristics in average value and the comparison to the standard drinking water quality in Malaysia.

Table 2. The characteristics of peat water sample from beriah peat swamp forest

Parameters	Unit	Average Value
pH	-	4.67 - 4.98
Temperature	°C	27.8
TDS	mg/L	20.6
DO	mg/L	3.4
Conductivity	uS/cm	34.5
Salinity	Ppt	0.02
Color	TCU	224.7
Turbidity	NTU	20.8
COD	mg/L	33.3
Iron, (Fe)	mg/L	1.24
NH_3-N	mg/L	0.51

Thirteen parameters were successfully determined where the first six parameters, including pH, temperature, TDS, DO, conductivity, and salinity were measured at the sampling point, whilst the rest of the parameters, including colour, turbidity, COD, iron ion, Ammoniacal Nitrogen, NH_3-N, Ammonia (NH_3), and Ammonium ($NH4^+$) were examined from the sample brought to the environmental laboratory on the same day.

Acidic pH of the peat water was predicted due to the composition of the surrounding peat soil itself which had been formed by decaying material possessing humic substances (Rieley, 1992). Besides that, humic substances also lead to the high organic content as humic substances are comprised of numerous oxygen containing functional group and fractions (humic acid, fulvic acids and humin) with different molecular weights which mean yielding high concentration of turbidity and COD as well as coloured water (Torresday et al., 1996). Moreover, composition of peat soil may also have an impact on the iron ion concentration of peat water (Botero et al., 2010).

From the thirteen parameters, five parameters were indicated exceeding the standard limit. These parameters were pH, colour, turbidity, COD, and iron ion that showed values of 4.67 - 4.98, 224.7 TCU, 20.8 NTU, 33.3 mg/l, and 1.24 mg/l respectively while the standard limit of these parameters are 6.5 - 9.0, 15 TCU, 5 NTU, 10 mg/l, and 0.3 mg/l accordingly.

3.2 Effect of Initial pH on the Efficiency of Colour, COD, Turbidity, and Iron Ion (Fe) Removal

Influence of initial pH on the adsorption capacity for removing colour, COD, turbidity, and iron ion were investigated.

Figure 2(a) to Figure 2(d) below, displayed the percentage removal of colour, COD, turbidity, and iron ion against pH of adsorbents respectively.

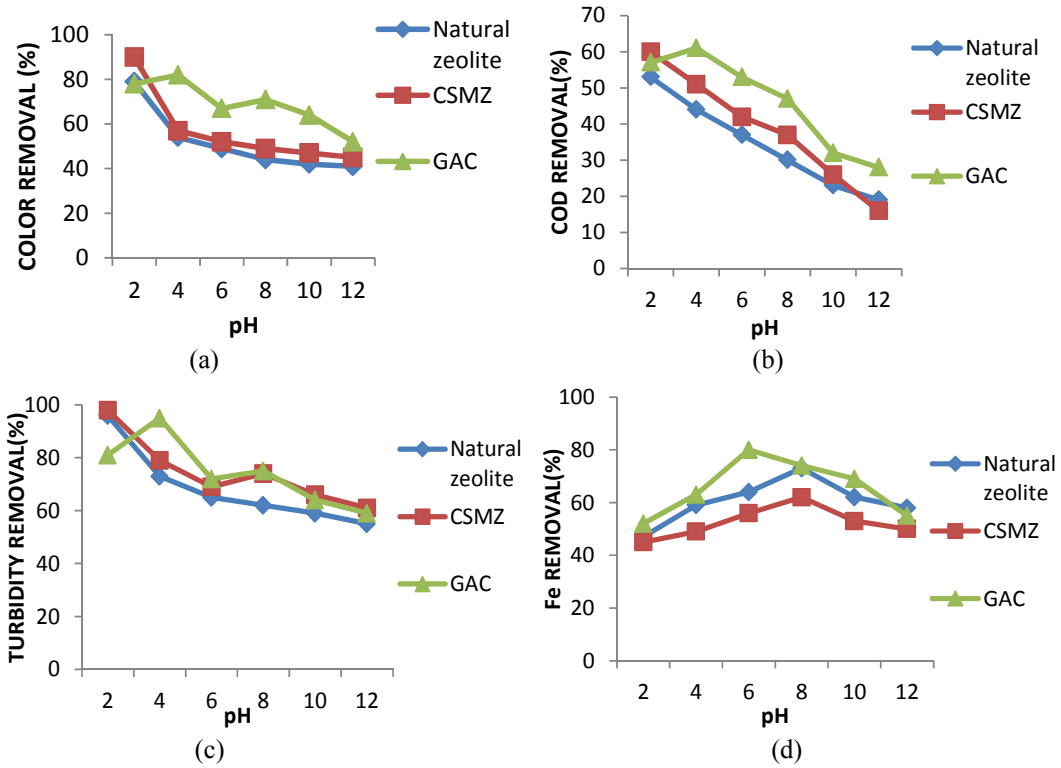


Figure 2. Percentage of color (a), COD (b), turbidity (c), and Fe (d) removal (%) against pH for NZ, CSMZ, and GAC

Figure 2(a) shows the maximum removal percentage of colour that was removed by natural zeolite, CSMZ, and granular activated carbon (GAC) which were 79%, 90%, 82% respectively. This adsorption is depended on the characteristic of adsorbents itself. For zeolite and CSMZ were related to the amount of cationic ions (Al^{3+}) increased, resulting in high reaction activity and GAC was related to the adsorption capacity. It was observed that the adsorption capacity was highly dependent on the pH of the solution, and indicated that the colour removal efficiencies decreased with the increase of solution pH.

The pH of the system exerts profound influence on the adsorptive uptake of adsorbate molecules presumably due to its influence on the surface properties of the adsorbent and ionization or dissociation of the adsorbate molecule. Figure 2(b) represents the percentage removal of natural zeolite and CSMZ where they reach optimum efficiency in removing organic compound (COD) at pH 2 with efficiency of 53% and 60% respectively. Meanwhile, the highest percentage removal of COD for GAC was achieved at pH 4 with efficiency obtained about 61%. Identical trends in colour removal were exhibited in percentage removal of COD for natural zeolite, CSMZ and GAC. In fact, this result also reveals that GAC has the highest percentage removal among natural zeolite and CSMZ yet optimum in difference pH solution. Neutralization mechanism occurs in low pH makes color removal, COD removal and Turbidity removals at pH 2 are higher for most of absorbents in this process.

In Figure 2(c), percentage turbidity removal against pH for each adsorbent revealed that optimal reduction of turbidity was obtained in an acidic environment with efficiency removal of 96%, 98%, 95% for natural zeolite, CSMZ, and GAC respectively. When the pH of the solution was adjusted above pH 6 to pH 12, the tendencies of all adsorption performances were gradually decreased. Moreover, it also showed that the lowest efficiency for the three adsorbents were identified at pH 12 with percentage values removal 55%, 61%, and 59% for natural zeolite, CSMZ, and GAC respectively.

Figure 2(d) demonstrates the removal efficiencies of iron ion as a function of the influent pH. The maximum removal of iron ion was observed at pH 8 for both natural zeolite and CSMZ whereas GAC had its optimum removal at pH 6. Natural zeolite and CSMZ only yielded 73% and 62% removal efficiency while GAC had more significant removal with removal efficiency of 80% to the iron ion concentration. Further, it is evident from the graph that gradual increment of removal efficiency for natural zeolite, CSMZ, and GAC occurred when the initial pH of the solution was increased to higher values. Somehow, at pH values greater than 6 the removal efficiency of GAC reduced slightly while for natural zeolite and CSMZ the reduction occurred from pH values above 8.

3.3 Effect of Adsorbent Dosage on the Efficiency of Colour, COD, Turbidity, and Iron Ion (Fe) Removal

The effect of adsorbent dosage was studied for all adsorbents employed on colour, COD, turbidity, and iron ion removal by varying the dosage of adsorbent and keeping all other experimental conditions constant. The pH was set to acidic conditions which were most favourable in obtaining the highest removal efficiency. In this study, to find optimal adsorbent dosage of natural zeolite and CSMZ, the appropriate experiments were carried out at adsorbent dosages in the range of 0.5 g to 5.0 g while for GAC, the adsorbent dosage was varied from 0.01 g to 4.0 g. The experimental results for all the adsorbents are represented by Figure 3(a) to Figure 4(d).

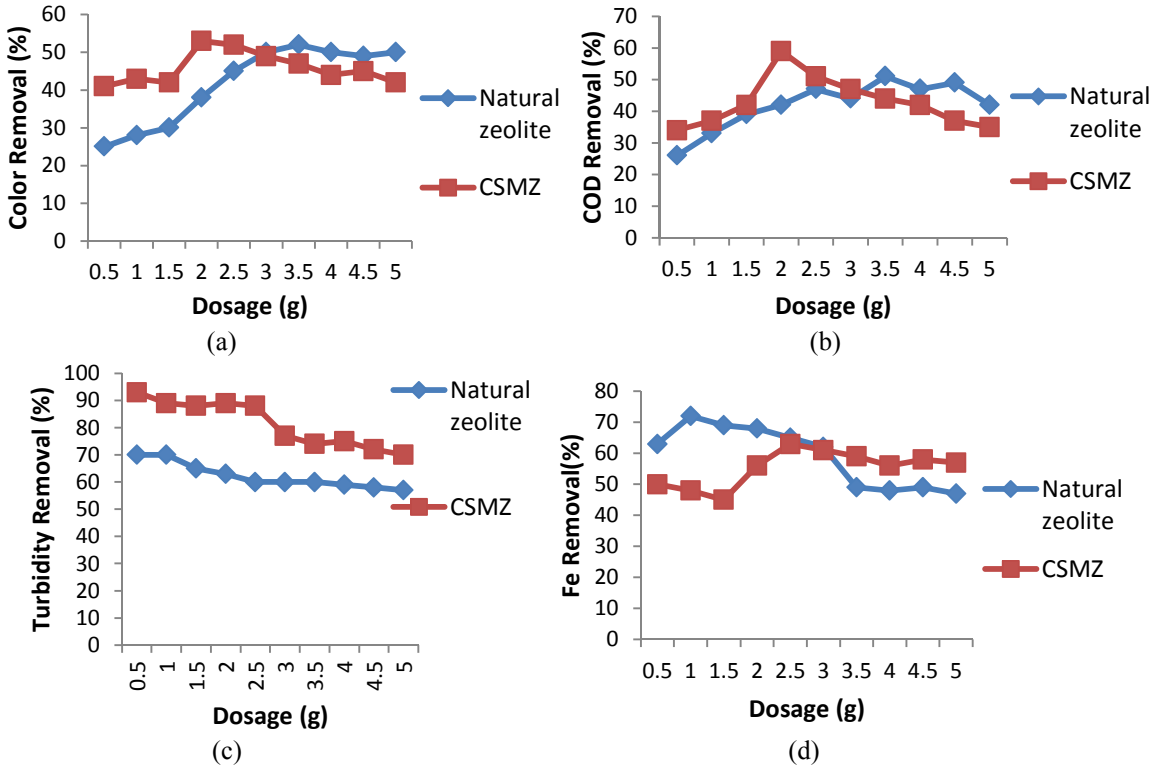


Figure 3. Percentage of color (a), COD (b), turbidity (c), and Fe (d) removal against pH for NZ, and CSMZ

Figure 3(a) displays the relationship between the amount of adsorbent mass (dosage) and adsorption efficiency for natural zeolite and CSMZ in terms of removing colour. The colour removal of peat water increased from about 25% to 52% with increasing adsorbent dosage of natural zeolite from 0.5 g to 3.5 g whereas for CSMZ, removal percentage increased from 41% to 53% with increasing adsorbent dosage from 0.5 g to 2.0 g. However, further increase in adsorbent dosage to 5.0 g only led to slight degradation of removal efficiency to 50% and 41% for natural zeolite and CSMZ respectively. This degradation with further increases in adsorbent dosage was due to the unsaturated adsorption active sites during the adsorption process since the adsorbates in the vessel were only shaken for 20 minutes (insufficient time). Besides, modification of zeolite by cationic surfactant had proven to have better colour removal as presented in the graph.

Percentage removal of COD against the adsorbent dosage is shown in Figure 3(b). It was observed that the highest percentage removal for both natural zeolite and CSMZ to remove COD were 51% and 59%, achieved at adsorbent dosage 3.5 g and 2.0 g respectively.

The variations in removal of turbidity of peat water at various system pH are shown in Figure 3(c). The removal rate of turbidity was highest at the adsorbent dosage of 0.5 g with 70% and 93% removal efficiency for respective natural zeolite and CSMZ. The removal rate showed a smooth downward trend with the increase in adsorbent dosage. Concurrently, the adsorption capacity gradually decreased with the increasing adsorbent dosage. The least efficient removal of turbidity was noted at dosage 5.0 g with percentage removal recorded for natural zeolite and CSMZ only 57% and 70% respectively.

Figure 3(d) demonstrates the percentage iron ion removal of natural zeolite and CSMZ with respect to their dosage. The result shows that there was a significant difference trend in iron ion adsorption efficiencies between natural zeolite and CSMZ. For natural zeolite, it was shown that the removal percentage of iron ion had increased until it reached 1.0g of dosage with 72% of removal efficiency. On the other hands, CSMZ was only able to remove about 63% of iron ion when its dosage was increased to 2.5 g. The lowest percentage removals were 47% and 57% recognized at the adsorbent dosage 5.0 g for respective natural zeolite and CSMZ.

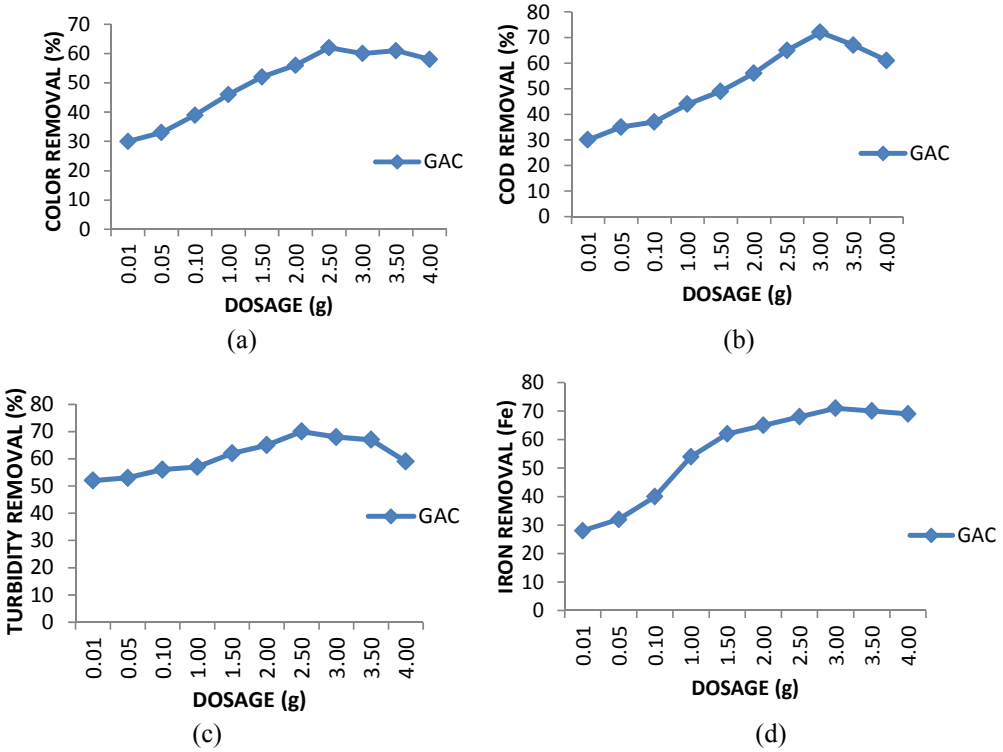


Figure 4. Percentage of color (a), COD (b), turbidity (c), and Fe (d) removal against dosage for GAC

The result illustrated in Figure 4(a) shows the maximum removal percentage of colour for GAC at 2.5 g dosage was 62%. Moderate increment in colour removal was identified along with the addition dosage of 2.5 g whilst abatement of removal efficiency began subsequently at adsorbent dosage of 3.0 g to 4.0 g.

The results from Figure 4(b) indicated that increasing the GAC dosage would increase the efficiency in removing COD respectively. The optimum dosage was recorded at 3.0 g with 72% of removal efficiency. Meanwhile, increasing the dosage above 3.0 g exhibited a slight decrease in removal efficiency with 67% to 61% for COD removal. A better result in removing COD was also shown by GAC compared to the natural zeolite and CSMZ.

The percentage of turbidity removed by GAC in different dosages is described in Figure 4(c). The highest removal was indicated at adsorbent dosage 2.5 g with removal efficiency of 70% while the minimum removal was 52% recorded at the adsorbent dosage 0.01 g. However, starting from adsorbent dosage of 3.0 to 4.0 g, removal efficiency began to decrease to 68%, 67%, and 69% respectively.

The result of percentage removal of iron ion by GAC in peat water is presented in Figure 4(d). It was found that the rate of removal was rapid in the initial dosage between 0.01 g to 3.0 g at which the removal efficiency increased from 28% to 71% accordingly. Subsequently, a few significant changes in the rate of removal were observed. Possibly, at the beginning, the solute molecules were absorbed by the exterior surface of adsorbent particles, so the adsorption rate was rapid. However, after the optimum dose was reached, the adsorption of the exterior surface becomes saturated and thereby the molecules will need to diffuse through the pores of the adsorbent into the interior surface of the particle (Ahmad & Hameed, 2009).

3.4 Batch Column Experiment

On the first running, the column was packed with natural zeolite (1st layer), limestone (2nd layer), and GAC (3rd layer) as shown in Figure 5(a). Removal efficiency for colour, COD, turbidity, and iron ion was recognized to be increased when the contact time was increased. At the time interval 1 hour to 6 hours, however, the increment was not so significant. The removal efficiency at 1 hour treatment was 39%, 21%, 54%, 36% while at 6 hours treatment was 77%, 65%, 73%, 60% recorded for respective colour, COD, turbidity, and iron ion. Poor removal efficiency at 1 hour treatment indicated that the required time to remove all parameters were insufficient. It is evident that if the adsorption process is allowed to run for 24 hours on the column, the removal efficiency shows notable removal. Percentage removals of colour, COD, turbidity, and iron ion at 24 hours were 83%, 72%, 76%, 65% respectively. Furthermore, the highest removal for respective colour, COD, turbidity, and iron ion were obtained at 48 hours treatment with 87%, 81%, 86%, and 79% of removal efficiency.

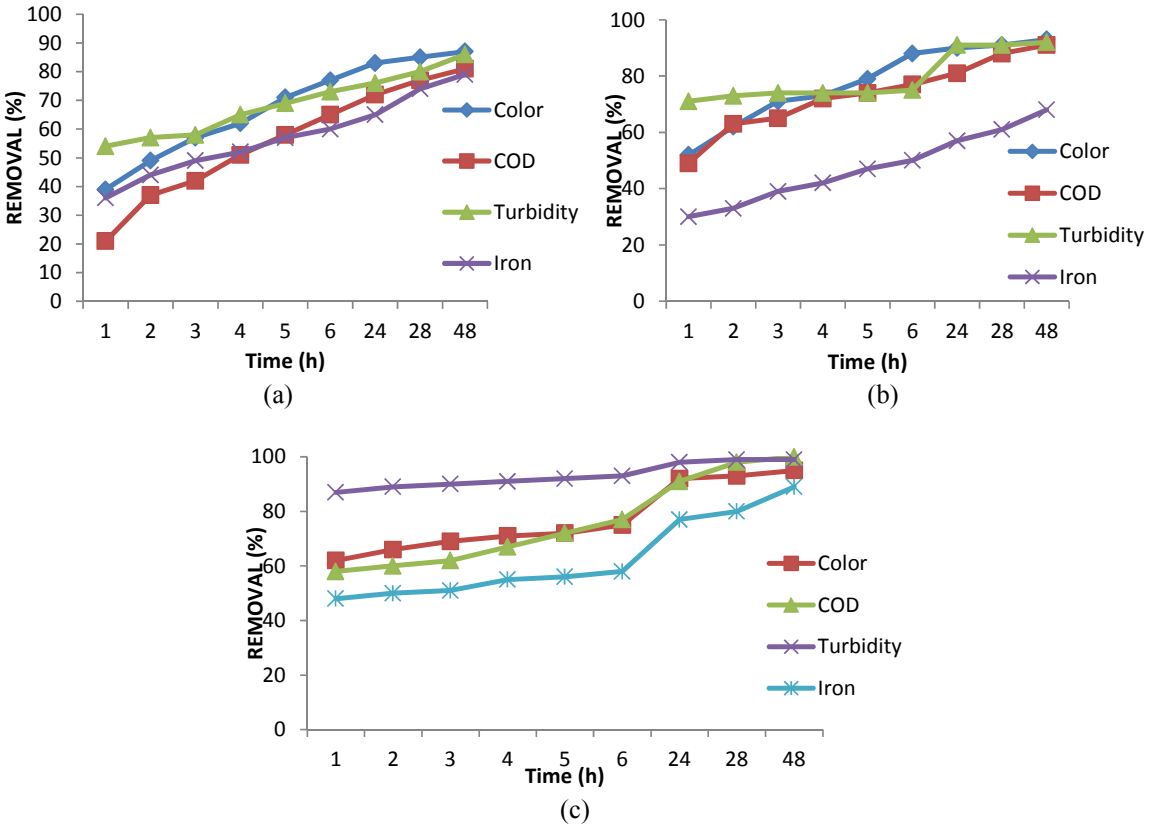


Figure 5. Percentage removal of color, COD, turbidity, and Fe for 1st run(a), 2nd run(b), and 3rd run (c) at flowrate 30 ml/min

On the second running, the column was packed with CSMZ (1st layer), limestone (2nd layer), and GAC (3rd layer) as presented in Figure 5(b). The removal percentages of colour, COD, turbidity, and iron ion were noticed after 1 hour to be 52%, 49%, 71%, and 30% respectively. The time of contact between adsorbate and adsorbent is proven to play an important role during the uptake of pollutants from peat water samples by adsorption process. In addition, the development of charge on the adsorbent surface was governed by contact time and hence the efficiency and feasibility of an adsorbent for its use in water pollution control can also be predicted by the time taken to attain its equilibrium (Sharma, 2003). Removal efficiency of 90% for colour, 81% for COD, 91% for turbidity, and 57% for iron ion were obtained at 24 hours of contact time.

On the third running, the column was packed with a difference sequence of CSMZ (1st layer), GAC (2nd layer), and limestone (3rd layer) demonstrated in Figure 5(c). It can be seen that the adsorption of these four parameters were slightly rapid at time interval 1 hour to 6 hours treatment. Further gradual increment with the prolongation of contact time form 24 hours to 48 hours has also occurred. Observation at 1 hour treatment recorded the removal efficiency of 62%, 58%, 87%, and 48% for respective colour, COD, turbidity, and iron ion. Whereby, 6 hours treatment had yielded higher removal percentage removal of 75%, 77%, 93%, and 58% respectively for colour, COD, turbidity, and iron ion. Further removal of colour, COD, turbidity, and iron ion was recorded when the treatment was run for 24 hours which exhibited 92%, 91%, 98%, 77% of removal efficiency respectively. Prolonged time to 48 hours indeed showed better removal of colour, COD, turbidity, iron ion with percentage removal of 95%, 100%, 99%, and 89% respectively. It can be seen that the arrangement of CSMZ, GAC, and limestone has the highest removal efficiency for all parameters at the flow rate influent of 30 ml/min.

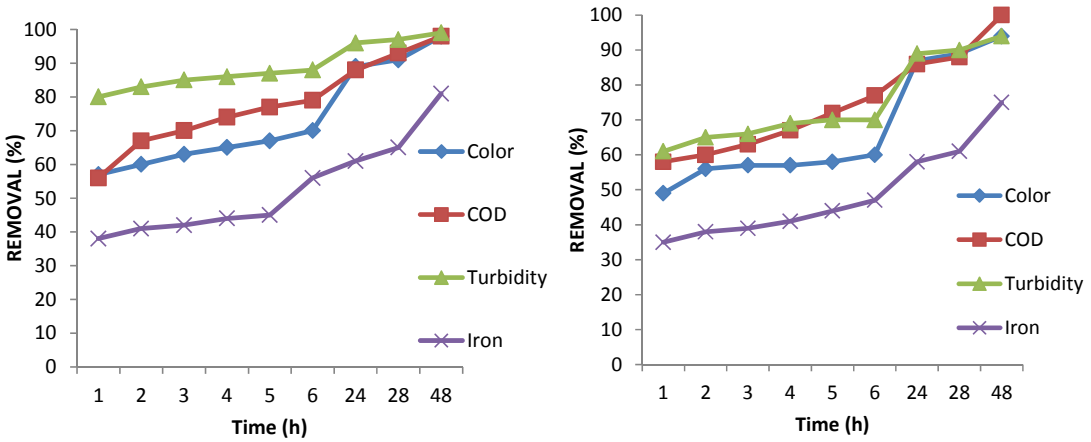


Figure 6. Percentage removal of color, COD, turbidity, and Fe against contact time for 2nd run(a) at flow rate 60 mL/min and at flowrate 90 mL/min (b)

The experimental adsorption behaviour was further seen for its adsorption capacity during 60 ml/min and 90 ml/min flow rate. In addition, the flow rate adjustment had also resulted in differences in surface loading rate in which the sample going through the surface area of adsorbent bed (horizontal surface area, A= 22.9 cm^2) for 30 ml/min equals to 1.31 cm/min while the flow rate of 60ml/min equals to 2.62 cm/min, and the flow rate of 90 ml/min equals to 3.93 cm/min. The percentage removal for both flow rate adjustments of CSMZ, GAC, and limestone arrangement were exhibited in Figure 6 (a) and Figure 6 (b). Based on these Figures, lower removal efficiencies were indicated at 1 hour time interval of 6 hours of contact time. The percentage removals for both 60 ml/min and 90 ml/min flow rate at 1 hour were 57%, 56%, 80%, 38% and 49%, 58%, 61%, 35% for colour, COD, turbidity, and iron ion respectively. Subsequently, when the contact time was at 6 hours, the removal percentage were 70%, 79%, 88%, 56%, and 60%, 77%, 70%, 47%. However, the maximum removal efficiency at 48 hours for both flow rates was not much different from the 30ml/min flow rate.

3.5 Adsorption Isotherm

In the present investigation, the experimental data were tested with respect to both Freundlich and Langmuir isotherms. Based on the linearized Freundlich isotherm models for natural zeolite, CSMZ, GAC in terms of adsorptive capacity to remove colour, COD, turbidity, and iron ion, the majority of them exhibited fits for all adsorbate with regression value (R^2) above 0.6, except for iron ion and turbidity for respective CSMZ, and GAC. On the other hand, the linearized Langmuir isotherm models for natural zeolite, CSMZ, GAC in terms of

adsorptive capacity to remove colour, COD, turbidity, and iron ion, had exhibited fits for all adsorbate with regression value (R^2) was at range of 0.242 to 0.912. The Langmuir isotherm model for all adsorption mechanisms were identified to have smaller R^2 values compared to the Freundlich isotherm model. Thereby, it can be concluded that the Freundlich isotherm model was more applicable in determining the adsorption mechanisms for this study.

3.6 Peat Water Quality Post Column Treatment

Peat water treatment in column with serial sequence of natural zeolite, CSMZ, and limestone had exhibited the highest removal with percentage removal at 48 hours at 95%, 100%, 99%, and 89% for colour, COD, turbidity, and iron ion respectively. Final readings at 48 hours treatment on pH, TDS, DO, conductivity, salinity, colour, turbidity, COD, and iron ion were 7.78, 74 mg/l, 4.03 mg/l, 137 uS/cm, 0.05 ppt, 12 TCU, 0.23 NTU, 0 mg/l, and 0.11 mg/l respectively (see Table 3). These findings, on the other hand, have indicated that peat water treatment had successfully produced water which satisfied the standard drinking water quality.

Table 3. The characteristics of results of peat water treatment from Beriah Peat Swamp Forest

Parameters	Unit	Average Value	Results after 48 hr	Standard Limit*)
pH	-	4.67 - 4.98	7.78	6.5 - 9.0
Temperature	°C	27.8	29	-
TDS	mg/L	20.6	74	1000
DO	mg/L	3.4	4.03	7
Conductivity	uS/cm	34.5	137	1000
Salinity	ppt	0.02	0.05	0.5
Color	TCU	224.7	12	15
Turbidity	NTU	20.8	0.23	5
COD	mg/L	33.3	0	10
Iron, (Fe)	mg/L	1.24	0.11	0.3
NH_3-N	mg/L	0.51	NA	1.5

Note: 1. *)Malaysian standard for drinking water quality;

2. NA = Not analyzed.

4. Conclusions

From the results presented in this paper, the following conclusions can be drawn:

1) The optimum removal of colour, COD, and turbidity for all adsorbents were observed to occur during acidic conditions at pH range 2 - 4 whereas for iron ion, the maximum removal was noted at pH range 6 - 8.

2) At pH 2, CSMZ yielded the highest removal for colour and turbidity with removal efficiency of 90% and 98% respectively. Meanwhile, GAC has the highest percentage removal of COD at pH 4 with removal efficiency obtained about 61% while at pH 6, GAC exhibited the best removal of iron ion with percentage removal around 80%.

3) CSMZ revealed stronger adsorptive capacity for colour, COD, and turbidity compared to natural zeolite.

4) The optimal removal was achieved for the serial sequence of CSMZ (1st layer), GAC (2nd layer), and Limestone (3rd layer) with the adsorbent media at 30 ml/min of flow rate.

5) Freundlich isotherm was more reliable to describe the adsorption mechanisms of colour, COD, turbidity, and iron ion for natural zeolite, CSMZ, and GAC.

Acknowledgement

The authors wish to acknowledge the financial support from the School of Civil Engineering, Engineering Campus, Universiti Sains Malaysia and Universiti Sains Malaysia (Short Term Grant No. 304/PAWAM/60312015).

References

Ahmad, A. A., & Hameed, B. H. (2009). Reduction of COD and colour of dyeing effluent from a cotton textile mill by adsorption onto bamboo-based activated carbon. *Journal of Hazardous Materials, 172*, 1538-1543. http://dx.doi.org/10.1016/j.jhazmat.2009.08.025

American Public Health Association (APHA), AWWA, WPCF. (1992). *Standard Methods for Examination of Water and Wastewater* (16th ed.). Washington.

Bhatmagar, A., & Minocha, A. K. (2006). Conventional and non-conventional adsorbents for removal of pollutant from water - A review. In *Indian Journal of Chemical Technology, 13*, 203-217

Botero, W. G., Oliveira, L. C., Rocha, J. C., Rosa, H. R., & Santos, A. D. (2010). Peat humic substances enriched with nutrients for agricultural applications: competition between nutrients and non-essential meals present in tropical soils. *Journal of Hazardous Materials, 177*, 307-311. http://dx.doi.org/10.1016/j.jhazmat.2009.12.033

Bowman, R. S. (2003). Applications of surfactant-modified zeolites to environmental remediation. *Microporous Mesoporous Materials, 61*, 43-56. http://dx.doi.org/10.1016/S1387-1811(03)00354-8

Chao, H. P., & Chen, S. H. (2012). Adsorption characteristics of both cationic and oxyanionic metal ions on hexadecyltrimethylammonium bromide-modified NaY zeolite. *Chemical Engineering Journal, 193-194*, 283-289. http://dx.doi.org/10.1016/j.cej.2012.04.059

Ćurković, L., Cerjan-Stefanović, Š., & Filipan, T. (1997). Metal ion exchange by natural and modified zeolites, *Water Research, 31*(6), 1379-1382. http://dx.doi.org/10.1016/S0043-1354(96)00411-3

Eltekova, N. A., Berek, D., Novak, I., & Belliardo, F. (2000). Adsorption of Organic Compounds on Porous Carbon Sorbents. *Carbon, 38*, 373-377. http://dx.doi.org/10.1016/S0008-6223(99)00113-X

Franceschi, M., Girou, A., Carro-Diaz, A. M., Maurette, M. T., & Puech-Coste, E. (2002), Optimisation of the coagulation–flocculation process of raw water by optimal design method. *Water Research, 36*(14), 3561-72.

Fu, F., & Wang, Q. (2011). Removal of heavy metal ions from wastewaters: A review. *Journal of Environmental Management, 92*, 407-418. http://dx.doi.org/10.1016/j.jenvman.2010.11.011

Hidaka, T., Hiroshi, T., & Kishimoto, N. (2003). Advanced treatment of sewage by pre-coagulation and biological filtration process. *Water Research, 37*(17), 4259-4269. http://dx.doi.org/10.1016/S0043-1354(03)00353-1

Huling, S. G., Robert, G. A., Raymond, A. S., & Matthew, R. M. (2001). Influence of Peat on Fenton Oxidation. *Water Research, 35*(7), 1687-1694. http://dx.doi.org/10.1016/S0043-1354(00)00443-7

Jamil, T. S., Ibrahim, H. S., Abd El-Maksoud, I. H., & El-Wakeel, S. T. (2010). Application of zeolite prepared from Egyptian kaolin for removal of heavy metals: I. Optimum conditions. *Desalination, 258*, 34-40. http://dx.doi.org/10.1016/j.desal.2010.03.052

Karadag, D., Akgul, E., Tok, S., Erturk, F., Arif Kaya, M., & Turan, M., (2007). Basic and reactive dye removal using natural and modified zeolite. *Journal of Chemical Engineering Data, 52*, 2436-2441. http://dx.doi.org/10.1021/je7003726

Liu, T., Chen, Zh. L., Yu, W. Z., Shen, J. M., & Gregory, J. (2011). Effect of two-stage coagulant addition on coagulation-ultrafiltration process for treatment of humic-rich water. *Water Research, 45*(14), 4260-4268. http://dx.doi.org/10.1016/j.watres.2011.05.037

Li, Zh. H., & Bowman, R. S. (2001). Regeneration of surfactant-modified zeolite after saturation with chromate and perchloroethylene. *Water Research, 35*(1), 322-326. http://dx.doi.org/10.1016/S0043-1354(00)00258-X

Li, Z. H., Jones, H. K., Robert, S., Bowman, & Helferich, H. (1999). Enhanced Reduction of Chromate and PCE by Pelletized Surfactant Modified Zeolite/Zerovalent Iron. *Environmental Science and Technology, 33*, 4326-4330. http://dx.doi.org/10.1021/es990334s

Li, Z., Roy, S. J., Zou, Y., & Bowman, R. S. (1998). Long Term Chemical and Biological Stability of Surfactant Modified Zeolite. *Environmental Science Technology, 32*, 2628-2632. http://dx.doi.org/10.1021/es970841e

Loo, S. L., Fane, A. G., Krantz, W. B., & Lim, T. T. (2012). Emergency water supply: A review of potential technologies and selection criteria. *Water Research, 46*(10), 3125-51. http://dx.doi.org/10.1016/j.watres.2012.03.030

Paune, F., Caixach, J., Espadaler, I., Om, J., & Riveraet, J. (1998). Assessment on the removal of organic chemicals from raw and drinking water at a Llobregat river water works plant using GAC. *Water Research, 32*(11), 3313-3324. http://dx.doi.org/10.1016/S0043-1354(98)00108-0

Rieley, J. O. (1992). The ecology of tropical peatswamp forest ± a South-east Asian perspective. In Tropical Peat, Proceedings of International Symposium on Tropical Peatland, Kuching, Sarawak, Malaysia, 6±10 May 1991 (B.Y. Aminuddin, ed.) pp. 244±54. Kuching, Malaysia: Malaysia Agricultural Research Development Institute & Department of Agriculture, Sarawak, Malaysia

Syafalni, S., Abustan, I., Dahlan, I., & Wah, C. K. (2012b). Treatment of Dye wastewater Using Granular Activated Carbon and Zeolite Filter. *Modern Applied Science, 6*(2), 37-51. http://dx.doi.org/10.5539/mas.v6n2p37

Syafalni, S., Abustan, I., Zakaria, S. N. F., & Zawawi, M. H. (2012a). Raw water treatment using bentonite-chitosan as a coagulant. *Water Science & Technology: Water Supply, 1*2(4), 480-488. http://dx.doi.org/10.2166/ws.2012.016

Torresdey, J. L., Tang, L., & Salvador, J. M. (1996). Copper adsorption by esterified and unesterified fractions of sphagnum peat moss and its different humic substances. *Journal of Hazardous Materials, 48*, 191-206. http://dx.doi.org/10.1016/0304-3894(95)00156-5

World Wildlife Fund (WWF) Malaysia. (2004). The importance of rivers.

Wosten, J. H. M., Clymans, E., Page, S. E., Rieley, J. O., & Limin, S. H. (2008). Peat- Water interrelationships in a Tropical Peatland Ecosystem in Southeast Asia. *Catena, 73*, 212-224. http://dx.doi.org/10.1016/j.catena.2007.07.010

Zhang, P., Tao, X., Li, Z., & Bowman, R. S. (2002). Enhanced Perchloroethylene Reduction in Column Systems Using Surfactant Modified Zeolite/zero-valent Iron Pellets. *Environmental Science and Technology, 36*, 3597-3603. http://dx.doi.org/10.1021/es015816u

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

The New Smart Eraser Design

Jinzan Liu[1], Zhong Zeng[1] & Lang Xu[1]

[1] University of Shanghai for Science and Technology, Shanghai, China

Correspondence: Jinzan Liu, College of Mechanical Engineering, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China. E-mail: jinzan_Liu@163.com

Received: December 5, 2012 Accepted: January 21, 2013 Online Published: January 29, 2013

doi:10.5539/mas.v7n2p50 URL: http://dx.doi.org/10.5539/mas.v7n2p50

Abstract

The paper puts forward a kind of mechanism design scheme, the mechanism can automatically detect the blackboard chalk stains, and erase the font, keep the blackboard clean. In this paper, the hardware system and software have been connected with each other, and tell us how to get installed. It introduces the principle of the mechanism, the method of image processing, and calculates the relationship of the displacement, velocity, acceleration and structure. This wipe mechanism adopts crank slider mechanism which has a good wipe effect.

Keywords: blackboard eraser institution, intelligent, wipe, image processing, crank slider mechanism

1. Introduction

The traditional blackboard chalk dust is a common problem in the traditional blackboard-eraser-chalk architecture (Ren, 2002). Lots of improved blackboard eraser structure came into being, but these improvements changed the original wipe, or not fundamentally solve the problem, or too costly and difficult to spread (Hu & Wan, 2008). For this reason, the design is based on the traditional blackboard-eraser-chalk called intelligent wipe chalk system, its connotation of "smart" includes moving, positioning, wipe (Wang, Song, & Wang, 2009). The design is able to achieve automated clean the blackboard and collect dust, which has a good prospect.

Bodies base on traditional blackboard and eraser size. The mechanism consists of three motors A, B, and C, three guide rails a, b, c, three sliders 1, 2, 3 and a slider-crank mechanism. Slider 1 and slider 4 are connected by rails c, and rails c is installed on them, can be moved in parallel with the slider, power-driven provided by three motors A, B, C. Motor A drives the left and right movement of beams c, and motor B drives the vertical movement of the slider 3, C is the power source of the slider-crank mechanism which driven rotary motion slider to rub c moving up and down along the rail, together with the installation in 2 below the blackboard (Chart 1).

The process is carried out as follows: the power of beam c comes from the motor A, the rail transfers the power. Motor A forward and reverse rotation drives the beam C forward or back movement. Motor B forward and reverse rotation drives the slider 3 upward and downward movement. The method is able to reach the special position, and then eraser cleans the blackboard. Motor C is installed in block 3 internal provides power for Slider-crank mechanism, pushes the slider 2 up-and-down along the rail c. Slider 2 is a combination of the movement of a, b, c, performs wipe. When the wipe is finished, A, B motor work and C motor stop, blackboard eraser returns to the original position, waiting for the next wipe.

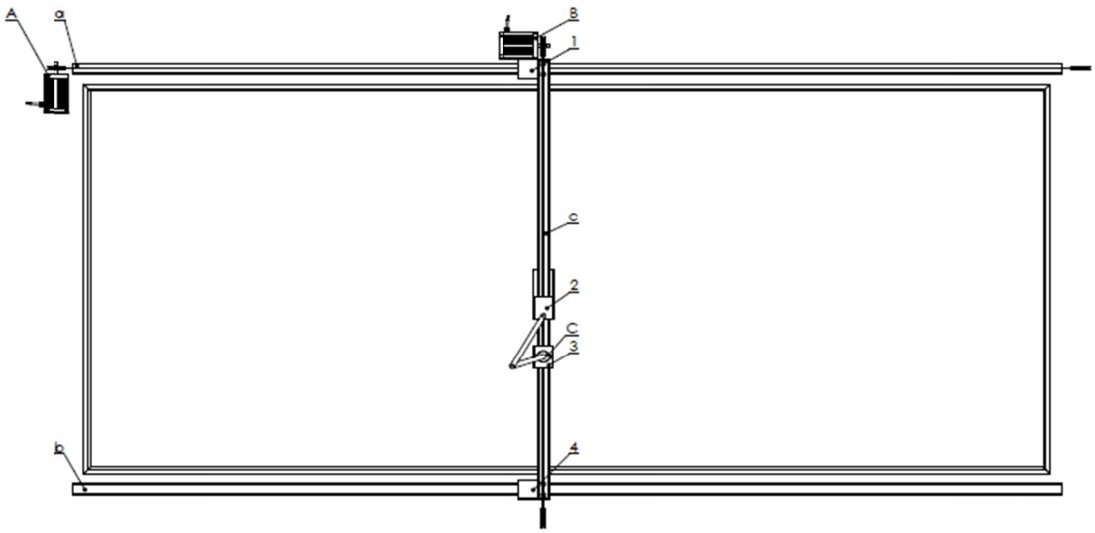


Chart 1. Mechanical structure

The blackboard surface is as the X-O-Y plane.

When the eraser begins to wipe, beam is at the left-most of the blackboard, and eraser on the top. Beam and eraser back to the original position after wiping completed. Before the eraser to wipe, CCD camera takes pictures of the entire blackboard. In the program, the stains are contained in the rectangle. By computer processing, the program determine the coordinates of the stains, PC can calculate the upper left corner of the rectangle's coordinate, this is the coordinate which the eraser should be reached. Then the motors release pulse signal, motors rotate a certain number of turns, eraser arrives at the designated location (Chart 2).

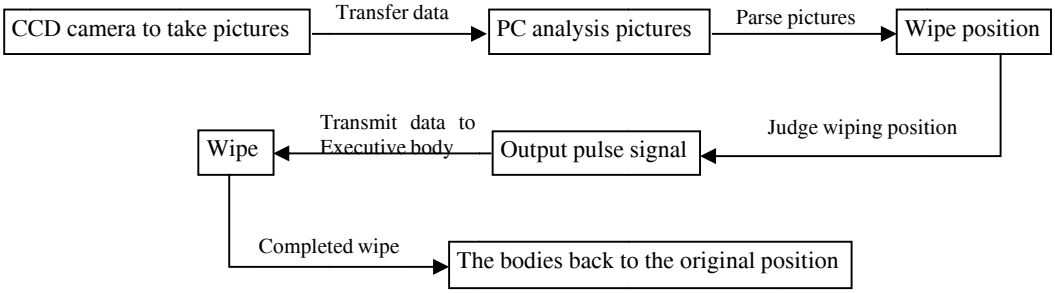


Chart 2. Flow chart

2. Visual Acquisition and Control

The automatic blackboard wipe system is divided into two parts of the information processing unit and motion control unit, whose hardware selection is run around the two parts (Zhang, 2008). The hardware of information processing unit involves to lighting source, CCD camera, lens, scene collect card & connection line parts and so on, whose main task is to use the image in the CCD camera collecting from the blackboard, then for analysis and processing of the acquired images to obtain the info. The blackboard is different from the background color of the blackboard surface. Finally output this information to provide the coordinates for the eraser wiping execution reference (Qian, 2005). The motion control unit includes motion control card, servo drives, servo motors and sensors, whose main task is to control the speed of the slider by the servo motor, make the eraser finish the wiping action successful (Yang & Chong, 2007).

The hardware of information processing (Deng, 2002; Zhou & Wu, 2008) unit involved mainly CCD cameras, lenses, frame grabbers and the connection line accessories. Motion control unit related to the hardware mainly have motion control card, servo drives, servo motors and sensors. The agency will choose the PC system as the hardware development platform, the visual system will choose the PCI interface frame grabber and CCD image sensors, motion control system will choose the PCI interface motion control card and servo motion controller.

The control system software design requires software platform with easy-to-use, modular, standardize and repeatable development features. Choosing a graphical programming language LabView make it convenient for maintenance and improvement.

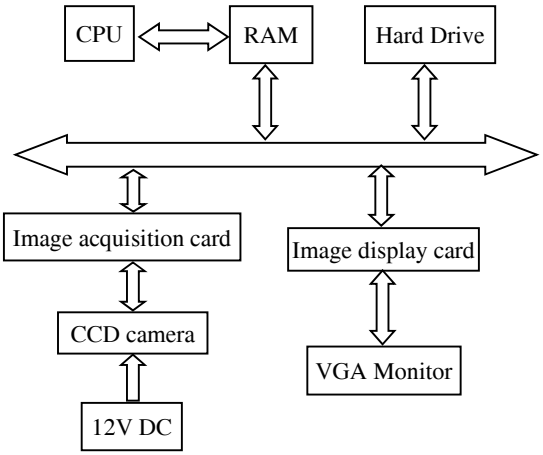


Figure 1. The acquisition system

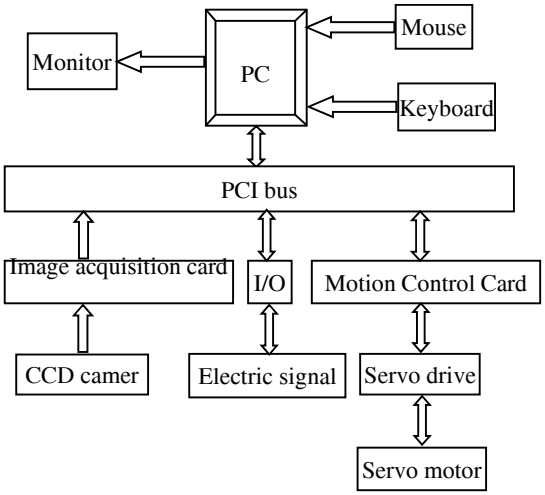


Figure 2. Control System Design

3. Image Processing

Image acquisition is a very important step in the visual collection. Because the unevenness of the CCD sensitivity and quantizing noise in photovoltaic conversion process is existent, and the reason of light. Actually, the quality of the image becomes poor, we cannot get a clear image. In order to facilitate the subsequent image analysis and understanding, it is necessary for image processing.

(1) Image Pretreatment

Image pretreatment contains image smoothing (Grosky & Tamburino, 1990; Sbnchez & Martinez, 2000; Sun & Fu, 2005) and image sharpening (Cai, Peng, & He, 2003).

The purpose of image smoothing is to suppress the image noise and improve image quality. In the practical application, the image noise source is come from blackboard surface's spots and the thermal noise of the flow of electron. When the noise reaches a certain intensity, will make subsequent image processing work cannot be carried out, so the image processing smoothing is required. In order to get a better picture of the edge information, the next is image sharpening. The purpose of sharpening processing of the image is mainly used to enhance the outline of the edge in the image, details and gradation to form a complete boundary of the object, to achieve the separation of the object from the image, or indicating that the surface region of the same object is detected. The Gradient and Laplace operator due to the isotropic nature, and therefore is a commonly used

method for sharpening. Gradient method is direct, simple, fast operation, but where there is a big change of image grayscale, there is a gradient's image output, but homogeneous regions in grayscale-scale gradient value is zero, so this method only display the edge contour where there is a big change of image grayscale.

Laplace is also a common edge enhancement processing operator, to enhance boundary contour. But it should be noted that, while using the Laplace operator sharpen the image，the image noise will be came together. Then enhance the image edge will enhance the image noise, the Laplace image sharpening first image smoothing. Before Laplace image sharpening, the first is image smoothing. In summary, Laplace as the subject sharpening method is suitable for selection under the premise of the smoothed image, the method can enhance the image edge, anti-interference, high positioning accuracy border and the edge of the continuous is good.

(2) Image Segmentation

Image segmentation is based on some of the characteristics of the image similarity criteria, grouping the image pixel, the image plane is divided into a series of special area, it will reduce the data which should be processed by the stage of the image analysis, recognition, etc. while retaining the information about the image structure characteristics. The paper uses threshold method to achieve image segmentation algorithm. The threshold is the method which we put each pixel to compare with the threshold value T, and we determined its category according to whether the grayscale value exceeds the threshold value. In this paper, the peak-to-valley method is determined segmentation threshold. The principle is: If the target area and the background area have some differences in grayscale image, then the histogram of image grayscale is bimodal shape, that is to say, the target and background pixel according to its probability in the image-grayscale histogram will form the respective crest and a trough between the two peaks occurs, so choose the pixel grayscale scale value of the trough at the image segmentation can be used as the threshold value T (chart 3). In this paper, chalk color and background color have a lot of contrast, so we can get a better bimodal pattern.

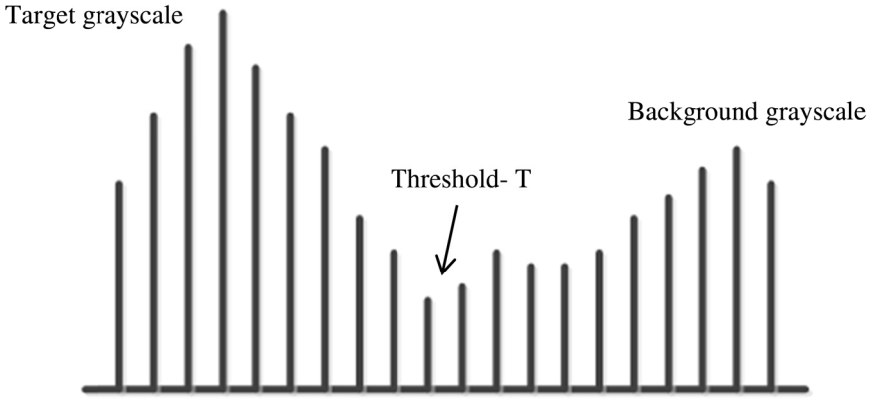


Chart 3. Image Grayscale

4. Slider-crank Mechanism Motion Analysis (Pu & Ji, 2001)

First, establishing a Cartesian coordinate and marking the vector and the azimuth of the rods. Known: the eraser width of D, A motor drives beam c, speed is V_x, the motor C drives slider-crank perform wiping action.

$$V_{cy} = L_1 \cdot \cos\theta \tag{1}$$

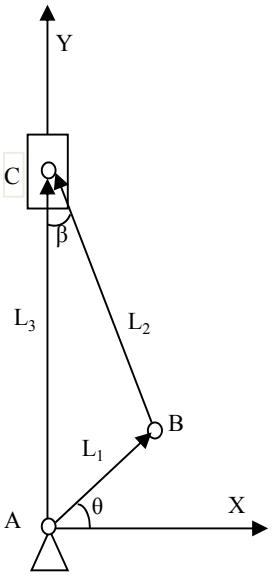
$$L_3 = L_1 \cdot \sin\theta + L_2 \cdot \cos\beta \tag{2}$$

$$l_2 \cdot \sin\beta = l_1 \cdot \cos\theta \tag{3}$$

(1) The first derivative with respect to time:

$$\alpha_{cy} = -L_1 \cdot \omega \sin \omega t \tag{4}$$

Blackboard eraser in the midline wipes to repeat D/2 traces, when the eraser is horizontal movement D/2, $\boldsymbol{L}_1$ rotational is half cycle:



$$\frac{D}{2v_t} = \frac{T}{2} \tag{5}$$

$$\omega = \frac{2\pi}{T} \tag{6}$$

$$L_3 = L_1 \cdot \sin\theta + \cos\theta\sqrt{L_1^2 - L_2^2} \tag{7}$$

We can get the conclusion v_c, α_c, L_c :

$$V_c = V_x + L_1 \cdot \cos\omega t$$

$$v_c = \sqrt{v_x^2 + v_{cy}^2}$$

$$= \sqrt{v_x^2 + \left(L_1 \cdot \cos\omega t\right)^2}$$

$$\alpha_{cy} = \alpha_{cy}$$

$$= -L_1 \cdot \omega \sin\omega t$$

$$\mathbf{S}_c = \mathbf{L}_1 \sin\omega t + \mathbf{V}_x t + \cos\omega t\sqrt{\mathbf{L}_1^2 - \mathbf{L}_2^2}$$

Written in the form of a matrix:

$$\begin{pmatrix} 0 & L_1 & V_x \\ -L_1\omega & 0 & 0 \\ L_1 & \sqrt{L_1^2 - L_2^2} & V_x t \end{pmatrix} \begin{pmatrix} \sin\omega t \\ \cos\omega t \\ 1 \end{pmatrix} = \begin{pmatrix} V_x \\ \alpha_c \\ S_c \end{pmatrix}$$

Already know D=60 mm, $\boldsymbol{L}_1$=100 mm, $\boldsymbol{L}_2$=180 mm, $\boldsymbol{V}_x$=6 mm/s, we can draw the graphics:

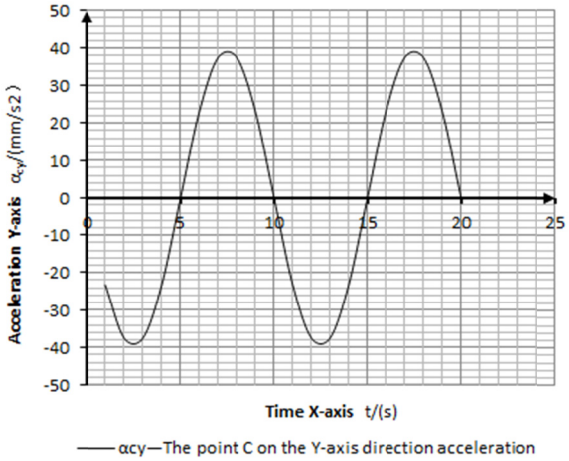


Figure 3. Acceleration Y-axis

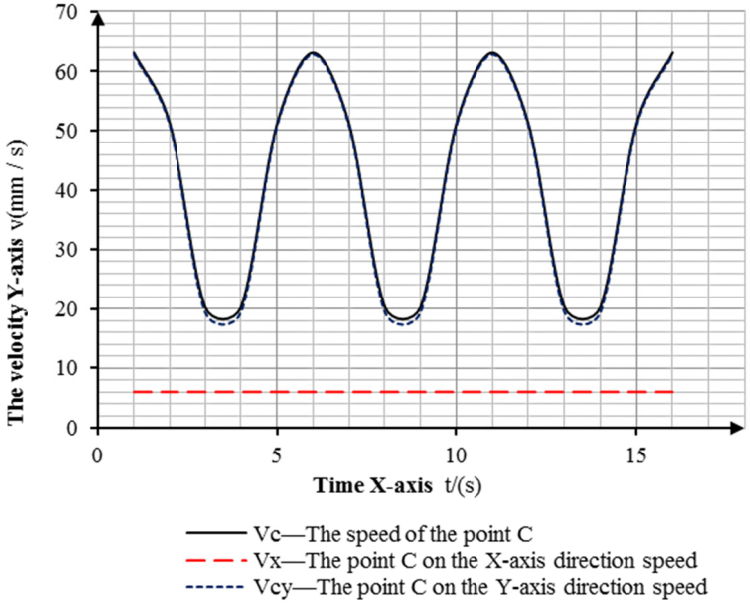


Figure 4. The velocity Y-axis

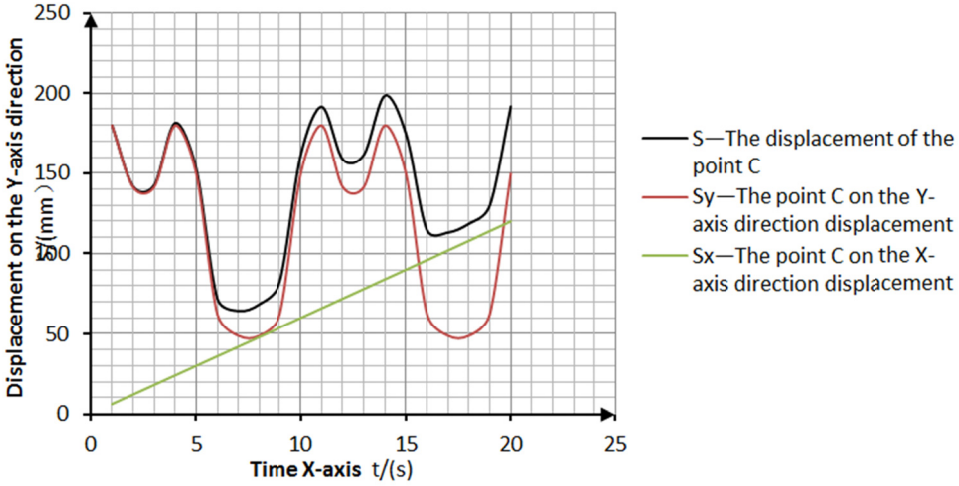


Figure 5. Displacement on the Y-axis direction

According to the calculation and analysis, we get a conclusion that blackboard erase is able to run smooth, and complete the expected movement.

5. Summary

Compared with manually wipe, smart wipe has a good effect and runs smooth with good reaction speed. The rate of rotation of the motor can be set in accordance with the requirements of the wiping speed to suit the requirements of different occasions.

1) The smart eraser has a simple structure, easy to operate, easy to obtain raw materials, manufacturing equipment simple process.
2) Control functions, and less susceptible to interference, high reliability, ease of use, can make products with high performance and low cost.
3) The product is suitable for large, medium and small institutions, the promotion of a certain significance.

References

Cai, Y. J., Peng, T., & He, J. (2003). The Research of Packaging-Checking Based on Application of Machine Vision. *Packaging Engineering, 24*(2), 45-48.

Deng, X. Z. (2002). *Electric Drive Control*. Wuhan: Hua zhong University of Science and Technology Press.

Grosky, W. I., & Tamburino, L. A. (1990). A Unified Approach to the Linear Camera Calibration Problem. IEEE Transactions on *Pattern Analysis and Machine Intelligence, 12*(7), 663-671. http://dx.doi.org/10.1109/34.56209

Hu, Y. K., & Jing, R. C. (2006). *Theoretical Mechanics.* Beijing: Higher Education Press.

Hu, Z. Y., & Wan, H. (2006). A pressure cleaner of blackboard eraser. *Journal of Wuhan Institute of Shipbuilding Technology*, Wuhan.

Pu, L. G., & Ji, M. G. (2001). *Mechanical design* (7th ed.). pp. 167-170. Beijing: Higher Education Press.

Qian, H. (2005). *Key Technology of the Vision System of High-speed Parallel Robot*. Tianjin: Tianjin University.

Ren, Z. G. (2002). Chalk dust on the health hazards of teachers (in Chinese). *Chinese Journal of School Health, 2*, 189.

Sbnchez, A. J., & Martinez, J. M. (2000). Robot-arm pick and place behavior programming system using visual perception. Proceedings 15th International Conference on *Pattern Recognition, 4*, 507-510.

Sun, H. Y., & Fu, D. S. (2005). Multi-target tracking technology in the video surveillance system. *Computer Applications and Software, 22*(11), 22-25.

Wang, X. F., Song, X. T., & Wang, Z. (2009). Innovative Design of a New Mechanism to Clean the Blackboard. *Mechanical Engineer, 10*.

Yang, S. Z., & Chong, Y. K. (2007). *Mechanical engineering control*. Wuhan: Hua zhong University of Science and Technology Press.

Zhang, C. (2008). *R & D of Vision Control System for a High Speed Packaging Robot*. Tianjin: Tianjin University.

Zhou, H. Q., & Wu, Q. X. (2008). *Microcomputer Principle and Interface Technology* (4th ed.). Hefei: University of Science and Technology of China Press.

Modern Applied Science; Vol. 7, No. 2; 2013
ISSN 1913-1844 E-ISSN 1913-1852
Published by Canadian Center of Science and Education

A Review of Passive Wireless Sensors for Structural Health Monitoring

Arvind Deivasigamani[1], Ali Daliri[1], Chun H. Wang[1] & Sabu John[1]

[1] School of Aerospace, Mechanical & Manufacturing Engineering, RMIT University, Melbourne, Australia

Correspondence: Arvind Deivasigamani, School of Aerospace, Mechanical & Manufacturing Engineering, RMIT University, GPO Box 71, Bundoora, Melbourne, Victoria 3083, Australia. Tel: 61-425-564-825. E-mail: arvind.deivasigamani@student.rmit.edu.au

Received: July 9, 2012 Accepted: January 16, 2013 Online Published: January 29, 2013

doi:10.5539/mas.v7n2p57 URL: http://dx.doi.org/10.5539/mas.v7n2p57

Abstract

Wireless sensors for Structural Health Monitoring (SHM) is an emerging new technology that promises to overcome many disadvantages pertinent to conventional, wired sensors. The broad field of SHM has experienced significant growth over the past two decades, with several notable developments in the area of sensors such as piezoelectric sensors and optical fibre sensors. Although significant improvements have been made on damage monitoring techniques using these smart sensors, wiring remains a significant challenge to the practical implementation of these technologies. Wireless SHM has recently attracted the attention of researchers towards un-powered and more effective passive wireless sensors. This article presents a review of some of the underlying technologies in the field of wireless sensors for SHM - with a focus on the research progress towards the development of simple, powerless, yet effective and robust wireless damage detection sensors. This review examines the development of passive wireless sensors in two different categories: (1) use of oscillating circuits with the help of inductors, capacitors and resistors for damage detection; and (2) use of antennas, Radio Frequency Identification (RFID) tags and metamaterial resonators as strain sensors for wireless damage monitoring. An assessment of these electromagnetic techniques is presented and the key issues involved in their respective design configurations are discussed.

Keywords: structural health monitoring, wireless sensors, damage monitoring, oscillating circuits, metamaterial resonators, RFID

1. Introduction

Over the past few decades, the field of Structural Health Monitoring (SHM) has attracted considerable research. Several effective damage monitoring techniques like strain measurement, electro-mechanical impedance, scattering of guided waves, acoustic emissions, dynamic response and optical techniques have been developed. Several sensors like strain gauges, piezoelectric sensors and optical fibre sensors have been employed (Housner et al., 1997; Chang, 2002; Auweraer & Peeters, 2003; Chang, Flatau, & Liu, 2003; Wang & Rose, 2003; Wang & Chang, 2005).

The above mentioned sensors are extensively used for damage monitoring; however, they do present certain limitations. Most of the existing sensors require an input power supply. When the sensors detect any change in strain or stress, they need to transfer the information, for signal processing and analysis, to the data acquisition system which may be located at the base station far away from the structure being monitored. The necessary connection of sensors by wires for power and data transmission often renders the SHM system complex to implement and difficult to maintain. The whole structure sometimes needs to be redesigned to accommodate the connections among these sensor networks; therefore, increasing the cost of manufacture. The technical difficulties of designing sensor systems along with their connections become more pronounced when the structure under investigation contains moving parts, such as a helicopter rotor. Furthermore, wiring between sensors and base-stations increases the cost of replacing damaged or degraded sensors.

In order to tackle these challenges, researchers have started investigating options which could result in wireless SHM. By making the sensors wireless through the incorporation of energy coupling and communication functionalities, it is possible to integrate the data acquisition and signal processing system in the same sensor unit.

One example is the use of microwave antennas along with the application of Micro Electro-Mechanical System (MEMS) technology, which could be integrated with the sensor nodes to communicate with the base station.

Although the above mentioned research contributed substantially to the development of wireless SHM, significant gaps still remain. In many of aforementioned cases, the power supply to the sensors and the signal conditioning system require wires or cables. Sometimes, the electronic chips in MEMS sensors contain integrated solid state batteries to supply power (Lynch & Loh, 2006). Recharging or replacing these batteries remains a major issue, particularly when the sensors are embedded in composite structures.

Recently, researchers have begun investigating alternative techniques to enable the sensor system to become completely wireless and passive; where the power is supplied by external sources wirelessly and the sensor transmits the signals wirelessly back to the base station (Spencer, Ruiz-Sandoval, & Kurata, 2004). In this regard, researchers have demonstrated sensor systems which can detect damage and monitor the structure effectively with less power consumption or completely wireless. Therefore, to address the issue of power supply to the sensors and signal conditioning system, several developments have been reported where the electromagnetic resonators could receive power wirelessly through Radio Frequency (RF) signals and utilize them to measure the strain in the structure (Chaimanonart & Young, 2006). The measured strain can be wirelessly transmitted to the base station for data analysis. Alternatively, on-chip energy harvesting systems have been reported as a means to solve the power requirement issue (Anton & Sodano, 2007).

This article presents a review of pertinent work in the field of wireless SHM, with a focus on the electromagnetic techniques. This article groups the use of electromagnetics in development of passive wireless sensors into two sections: (1) use of oscillating circuits with the help of inductors, capacitors and resistors for damage detection; and (2) use of antennas, Radio Frequency Identification (RFID) tags and metamaterial resonators as strain sensors for wireless damage monitoring. This article aims to review the current state of the art and progress towards the development of effective and simple passive wireless damage monitoring techniques.

2. Use of Oscillating Circuits and Wireless Interrogation

Although the work mentioned in the previous section contributes significantly to the development of wireless SHM, they still have certain limitations. In most of the above cases, the power input to the sensors and the signal conditioning system are given using wires or cables. Sometimes, these chips are integrated with solid state batteries to supply power. It is however, important to ensure that the charge in the batteries is sufficient for the useful life of these systems. This becomes even more important when the sensors are embedded in the host materials. This provides the need for developing a sensor system that could be completely wireless in terms of power supply and data transmission. Therefore, researchers started investigating the application of oscillating circuits in strain monitoring. Their objective was to transform the strain in the structure into the change in the frequency of the oscillating circuits in order to read the values wirelessly. Several RC and LC tank circuits are used in this regard which could transform the strain in the structure to the change in inductance, capacitance or resistance which in turn changes the resonant frequency of the circuit. Several researchers have reported developments in the use of oscillating circuits for wireless SHM and this section reviews this body of work.

Watters, Jayaweera, Bahr, and Huestis (2001) suggested a very effective wireless sensor for structural health monitoring. In this technique, a sensor was integrated with a passive RFID chip. The chip receives the RF power wirelessly and converts it to DC power. The received power is utilized by the sensor and the signals are sent back to the receiver which makes the system completely wireless. This system requires the integration of the chip with the signal conditioning unit and the sensor. The diagram of the RFID/sensor is shown in Figure 1.

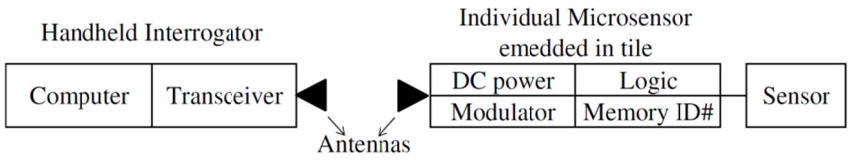


Figure 1. Principle of operation of an RFID/sensor hybrid (Watters et al., 2001)

Butler, Viglio, Vendi, and Walsh (2002) developed a strain sensor with an inductively coupled resonant circuit (LC tank circuit) for wireless damage detection. The governing equation for the resonant frequency of the circuit is,

$$f = 1/(2\pi\sqrt{LC}) \tag{1}$$

Where f is the natural frequency; L is the inductance; and C is the capacitance. The inductance of the solenoid can be calculated by,

$$L = k\mu N^2 A/l_0 \tag{2}$$

Where k is the form factor; μ is the permeability; N is the number of turns; A is the cross sectional area; and l_0 is the solenoid height.

Thus,

$$f = (1/2\pi)\sqrt{l_0/Ck\mu N^2 A} \tag{3}$$

By applying strain the cross sectional area of the solenoid changes and from Equation 3 its resonant frequency changes accordingly. A high frequency oscillator was used to measure the resonant frequency of the solenoid. The dip in the RF power is measured to find the frequency of the sensor and thus the applied strain. The experimental setup of this concept is illustrated in Figures 2 and 3.

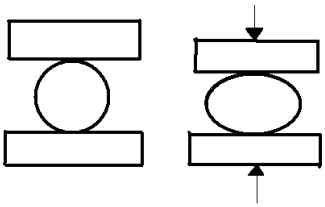


Figure 2. Schematic of guillotine compressing non-embedded sensor coil (Butler et al., 2002)

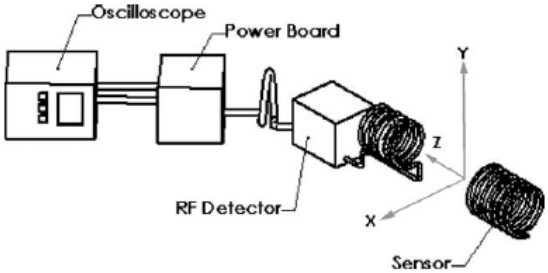


Figure 3. Illustration of experimental setup (Butler et al., 2002)

Chuang, Thomson, and Bridges (2005) developed an embeddable wireless strain sensor which works with the same principle as the previous case. The difference is that instead of a solenoid a coaxial resonant RF cavity was used as the sensor (Figure 4). The cavity length changes under the applied load thereby changing its resonant frequency. The shift in the resonant frequency with respect to the applied strain is shown in Figure 5. This sensor was shown to be linear up to 130 $\mu\varepsilon$. The shift in the frequency was about 2.42 kHz per $\mu\varepsilon$. The relationship between the applied strain and the resonant frequency of the cavity was derived using the following equations.

$$f_{str} = \frac{C}{2(l+\Delta l)} = \frac{C}{2l}\left(\frac{1}{1+\varepsilon}\right) \approx f_{unstr}(1-\varepsilon) \tag{4}$$

$$\Delta f_r = f_{str} - f_{unstr} \approx -f_{unstr}\varepsilon \tag{5}$$

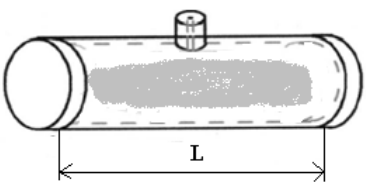


Figure 4. Electromagnetic coaxial cavity sensor. The dominant TEM001 resonant mode is shown (Chuang et al., 2005)

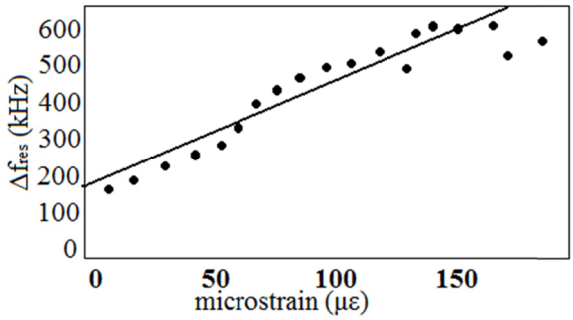


Figure 5. Change in Δ*fr* as a function of the strain applied to a concrete block (Chuang et al., 2005)

Umbrecht, Wendlandt, Juncker, Hierold, and Neuenschwander (2005) developed a wireless passive strain sensor for bio-medical applications. This sensor uses an incompressible liquid for strain measurement. As the load is applied, the liquid moves outward through a capillary tube proportional to the amount of load applied. This change in the liquid level is read wirelessly using an ultrasound imaging technique (Figure 6).

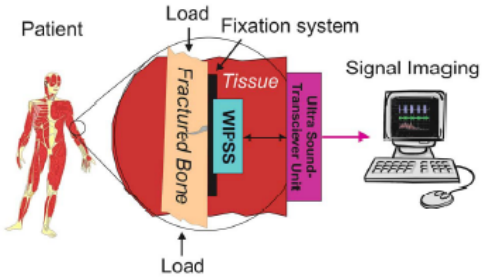


Figure 6. Typical application of a wireless implantable passive strain sensor system (Umbrecht et al., 2005)

Matsuzaki and Todoroki (2005) extended the use of wireless strain sensors to the automotive sector. A strain monitoring system was developed to measure and monitor the strain induced in an automotive tire by using the electrical capacitance variations within an oscillating circuit. The steel wires between the tread and the carcass were used as capacitors along with the resistors to form a RC oscillating circuit with a specific resonant frequency. When the tire is strained the distance between the wires is altered; hence, the resultant resonant frequency of the circuit shifts. These frequency signals were read wirelessly by inductive coupling using an external antenna. The schematic of the strain monitoring system using an oscillating circuit is shown in Figure 7.

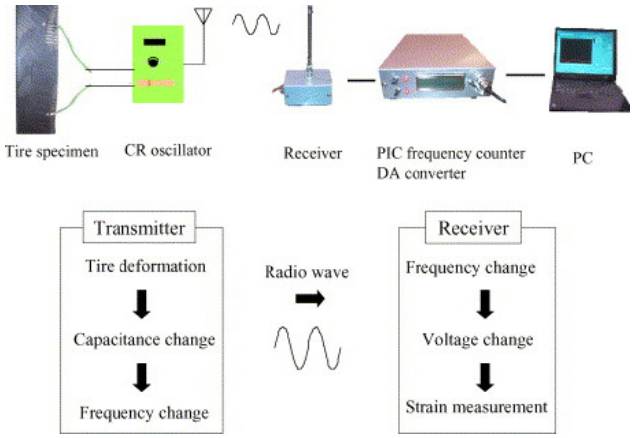


Figure 7. Schema of the strain monitoring system using an oscillating circuit (Matsuzaki & Todoroki, 2005)

Matsuzaki and Todoroki (2006) developed a wireless detection technique for internal delaminations/cracks in Carbon Fibre Reinforced Polymer (CFRP) laminates using frequency changes of an oscillating circuit. The CFRP laminate was itself used as the sensor because the electric current can flow through the conductive carbon

fibres. Two electrodes were attached to two sides of the laminate and were connected to an oscillating circuit. Presence of a crack/delamination in the laminate changes the resistance of the material and therefore shifts the frequency of the oscillating circuit. The resonant frequency of the circuit can be read wirelessly through an external antenna. The schematic of a CFRP laminate and the electrical current flow are shown in Figure 8.

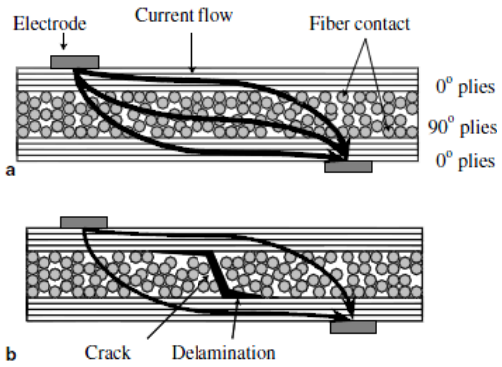


Figure 8. (a) Electrical network structure of the fibres in a CFRP laminate; (b) The electrical network is broken with a delamination (Matsuzaki & Todoroki, 2006)

Jia, Sun, Agosto, and Quinones (2006a) designed a passive wireless strain sensor for structural health monitoring. A plane spiral inductor as shown in Figure 9 was coupled with an interdigital capacitor to form a LC oscillating circuit which acts as a strain sensor as well as an antenna. The capacitance value is dependent on the distance between the fingers of the interdigital capacitor and hence it changes when there is a strain. The inductor receives the electromagnetic (EM) waves and sends the energy to the capacitor. This energy is then received back by the inductor and the signals are transmitted. This sensor was tested under load when attached to a cantilever beam showing acceptable linearity and sensitivity.

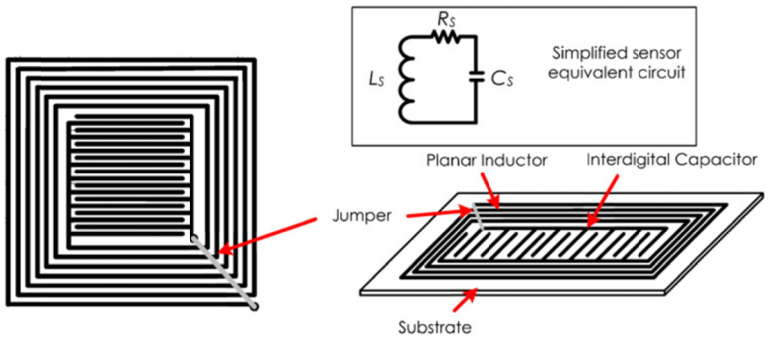


Figure 9. Sensor unit with a laminated sandwich structure (Jia et al., 2006a)

Jia and Sun (2006b) presented a novel wireless and powerless strain sensor with a multilayer thick film structure. The sensor consists of a planar inductor (L), a capacitive transducer (C), and a strain sensitive polarized polyvinylidene fluoride (PVDF) piezoelectric thick film. The resonant circuit was used to realize the wireless strain sensing through strain-to-frequency conversion and also by receiving radio frequency electromagnetic energy for powering the sensor. The results of the calibration on a strain constant cantilever beam showed marked linearity and the sensitivity of 0.0013 in the strain range of 0 to 0.018.

Tan, Pereles, Shao, J. Ong, and K. Ong (2008) developed a wireless passive strain sensor based on the harmonic response of magnetically soft materials. The sensor includes a deformable layer between a magnetically soft ferromagnetic alloy (sensing element) and a permanent magnet (biasing element) (Figure 10). The magnetic harmonic spectrum changes under the applied load which can be measured wirelessly when the sensor is under an AC magnetic field. The shifts in the magnetic harmonic spectrum of the sensor were linearly correlated with the mechanical strain. This sensor demonstrated good stability, linearity and repeatability. This passive wireless sensor is useful for long-term detection of mechanical loading from within an object such as inside a concrete structure or a human body.

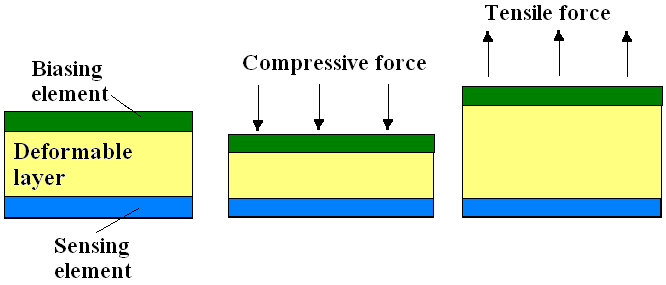


Figure 10. The biasing and sensing elements of the sensor are separated by a flexible layer to provide proper strain for a given compressive force (Tan et al., 2008)

In this section, several wireless strain sensing techniques utilising LC or RC oscillating circuits are discussed. The summary of the most important techniques are presented in Table 1. For these circuits to resonate at a specific frequency, power has to be supplied to the circuits. Some of these techniques employed external frequency oscillators and hand interrogators to provide power to the LC/RC circuits wirelessly. Thus, it is important to ensure an efficient energy coupling between the transmitter and the receiver. In SHM in order to determine and predict the crack propagation, it is essential to measure the strain in the structural member and to determine its spatial distribution. The work reviewed in this section does not refer primarily to the strain spatial distribution which is a major concern for researchers. Although these techniques are shown to be linear, it is important to ensure an acceptable reliability and repeatability for practical applications. These limitations provided the direction for further research, leading to the development of techniques discussed in the next section.

Table 1. Various techniques employed using oscillator circuits for wireless strain measurement

Authors	Technique
Butler et al., 2002	The strain changes the dimensions of the inductor and hence the inductance in the LC circuit, thus changing the resonant frequency.
Chuang et al., 2005	The strain changes the cavity length of the coaxial RF cavity thereby changing the resonant frequency.
Umbrecht et al., 2005	The strain moves the incompressible liquid through the capillary which is wirelessly read using ultrasound imaging technique.
Matsuzaki & Todoroki, 2005	The strain changes the capacitance of the RC oscillating circuit and hence changes the resonant frequency.
Jia et al., 2006a	The strain changes the capacitance of the interdigital capacitor coupled with the spiral inductor thereby changing the frequency of the LC circuit.
Tan et al., 2008	The strain deforms the flexible layer between the magnetically soft material and the permanent magnet hence changing the harmonic spectrum.

3. Use of Resonators and Antennas as Strain Sensors

The works discussed in the previous section are fairly simple and effective; however, researchers started to develop sensors which could directly convert strain into frequency shifts that could be read wirelessly. Antennas and electromagnetic resonators are passive devices which could be illuminated by incident electromagnetic waves and the backscattered signals could be received wirelessly using other antennas. Researchers have further tried to determine the direction of the strain induced in the structure. However, at present, it seems the work in this field is just starting. This section presents the recent techniques employed for wireless strain and damage monitoring which utilise resonators or antennas as strain sensors.

Das, Khorrami, and Nourbakhsh (1998) designed a novel sensor/actuator system which utilizes a patch antenna with a multilayer substrate (Figure 11). The multilayer consists of a dielectric layer and a piezoelectric layer. The piezoelectric layer is the sensing unit which converts the measured strain/vibration into voltage. The antenna receives wireless EM signal from the base station and generates a voltage which gets added up to the

piezoelectric voltage and this modulated signal is transmitted back to the base station. This antenna can also be used for actuating the piezoelectric layer by supplying the required voltage to the piezoelectric by receiving wireless electromagnetic power. Das et al. (1998) developed a dielectric-piezoelectric grating technique to distinguish sensing and actuating activities. Due to this grating technique, the sensing and actuating functions are activated separately using orthogonal polarization orientation techniques. It is also feasible to stack such microstrip patch antennas with dimensions to operate at different frequencies. This sensor integrates wireless power reception, sensing and data communication in one simple unit. However, this sensor could function well only when it is interrogated from a very close distance.

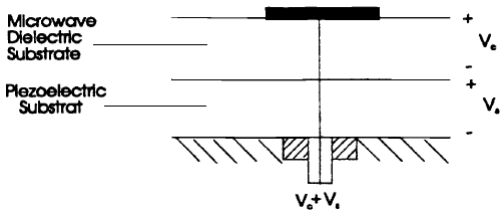


Figure 11. Microstrip antenna with Dielectric-Piezoelectric multilayer substrate (Das et al., 1998)

Loh, Lynch, and Kotov (2007) developed a wireless RFID based sensor by incorporating the field of nanotechnology. They utilized a layer by layer fabrication technique of Single Walled carbon Nano-Tube (SWNT) films. These films could act as a strain or a pH sensor because their capacitance or resistance changes accordingly. When these films are integrated with a coil antenna, they could be inductively coupled using a RFID reader and thus rendering the sensor completely wireless. Because these sensors act as a RLC oscillating circuit, the resonant frequency changes with the change in mechanical behaviour of the structure. Use of conducting carbon nanotube-gold nanocomposites as an inductor for wireless coupling was also investigated. However, the inductance was shown to be low, thereby limiting its wireless range to a very small distance. The size of the film sensor is 2.5 cm × 2.5 cm and is stated to be sensitive and linear. Although this technique might be useful, it is believed that the manufacturing of such film nanocomposites could be expensive.

Matsuzaki, Melnykowycz, and Todoroki (2009) developed a very innovative technique for wireless detection of damage in CFRP. The CFRP structure (e.g. the wing structure) can be modeled as a half-wavelength dipole antenna (Figures 12 and 13). The resonant frequency of the antenna is dependent on the length of the structure. When there is a crack perpendicular to the fibre direction, the dipole length decreases and hence the resonant frequency increases. Therefore, by measuring the frequency, the length of the dipole could be back calculated. With the length value, the crack location could be precisely identified. This method can only be used in structures with a specific geometry and it can only detect the crack when the crack reaches its critical length.

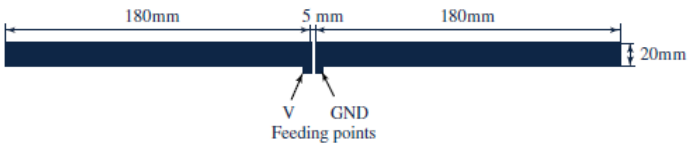


Figure 12. Simulation model of a rectangular dipole antenna (Matsuzaki et al., 2009)

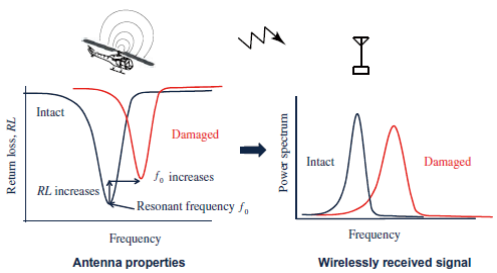


Figure 13. Schematic of the wireless crack detection mechanism (Matsuzaki et al., 2009)

Bhattacharyya, Floerkemeier, and Sarma (2009) investigated a RFID tag antenna sensor for displacement measurement (Figure 14). A simple RFID tag was kept at a very close distance to a metal surface which was attached to the structure. As the structure deforms, the metal surface comes closer to the RFID tag which affects the antenna's impedance and hence changes the backscattered power. It also affects the threshold power required to turn the RFID tag 'on'. This RFID tag can be queried wirelessly from a convenient location using an RFID tag reader/transmitter. By processing the backscattering from the RFID tag, the displacement of the structure could be evaluated. Although this sensor is very cheap and simple to design, there are certain challenges associated with this design. Obtaining the displacement data from backscattering becomes difficult if there are other metallic elements in the host structure. Due to the randomly moving metallic components, the sensor might give false positive results. Moreover, the sensor is sensitive to the displacements of the structure only in one direction. However, it is mentioned that this sensing technique could be optimized and utilized for an effective passive wireless displacement sensing system.

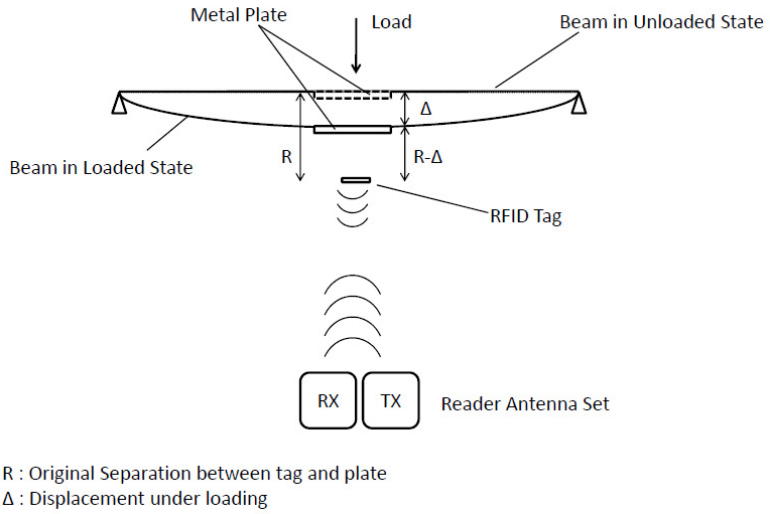


Figure 14. RFID sensor setup (Bhattacharyya et al., 2009)

Occhiuzzi, Paggi, and Marrocco (2011) proposed a meandered RFID tag sensor. This sensor can measure strain based on the change in the impedance and gain of the tag as a result of the deformation in the meandered line. This RFID tag requires an IC chip which increases the complexity of the structure. The shift in the power level as a result of the applied strain cannot be distinguished from the shift caused by other parameters that may influence the power transmitted (i.e. propagation path-loss, reflection, diffraction etc.). Caizzone and Marrocco (2012) further studied the application of this sensor in a RFID network grid to monitor the deformation of the structure. Their study shows that increasing the number of RFID tags does not further improve the sensitivity of the grid when the spacing between the sensors becomes lower than an optimum threshold.

Mandel, Schussler, and Jakoby (2011) proposed another concept for wireless passive strain measurement based on RFID tags principles. The proposed structure is composed of two layers of metal divided by a dielectric layer. The two metal layers are connected through the dielectric using an interconnecting via. This "mushroom structure" can be considered as a special case of a short-circuited microstrip patch antenna. One sensor structure comprised different elements, which were separated by gaps in the top metal layer and the substrate. The resonance frequency of each element is determined by the gap capacitance and the via inductance. Different fixed resonant frequencies can be used for identification purposes. The performance of this sensor was investigated using numerical simulations and experimental measurements. However, the linearity of sensor with applied strain is not discussed and further studies are required to quantify its performance.

Another recent study on RFID tags for passive wireless strain measurement was presented by Yi et al. (2011a). The sensor consists of a folded rectangular microstrip patch antenna with an IC chip. This passive wireless strain sensor operates based on a change in the impedance of the patch antenna as a result of the applied strain, which introduces a mismatch between the antenna and the IC package. When the EM power is sent wirelessly from a remote interrogator, the patch antenna receives the power and transfers it to the chip. This transfer from the tag to the chip is maxima when the interrogation frequency matches with the resonance frequency of the patch

antenna. The effect of changes in the impedance of the microstrip line between the rectangular patch and the IC was not considered in the model. The strain measurement was based on the change in the transmitted power which could also be affected by other factors. Due to the lack of sharpness in the transmitted power trace the exact resonant frequency of the RFID tag was difficult to ascertain. Hence, the resonant frequency of the tag was extracted using curve fitting techniques for the wireless strain measurements. Effect of antenna substrate thickness on the interrogation distance and strain transfer rate is further studied by Yi et al. (2011b).

Further investigations by Yi et al. (2012a) show that the shift in the resonant frequency of this sensor can be identified without the need for curve fitting for higher strain values (>4997 με). Also, this sensor can be used for monitoring crack growth when the crack propagates in one direction through the sensor. Yi et al. (2012b) implemented this sensor in an array to measure strain in different locations wirelessly. It is shown that, using the RFID tag protocols, different sensors in the array can be activated individually with minimum interference with other neighbouring sensors.

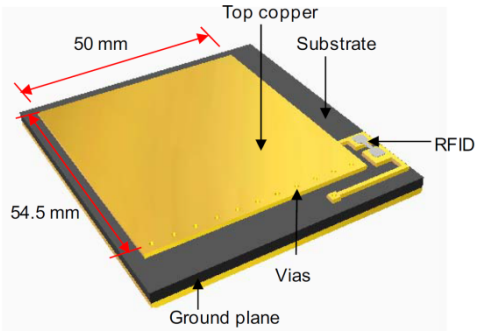


Figure 15. The RFID folded patch antenna strain sensor (Yi et al., 2012a)

Melik, Pergoz, Unal, Puttlitz, and Demir (2008) developed a passive on-chip RF-MEMS strain sensor for bio medical applications. As the material is stressed, the area of the sensor (spiral resonator) decreases thus its resonant frequency shifts. To make it completely wireless, two antennas of the same configuration were used for receiving and transmitting the signals. The micrograph of the fabricated sensor system is shown in Figure 16. This system is very small in size and has a very high quality factor.

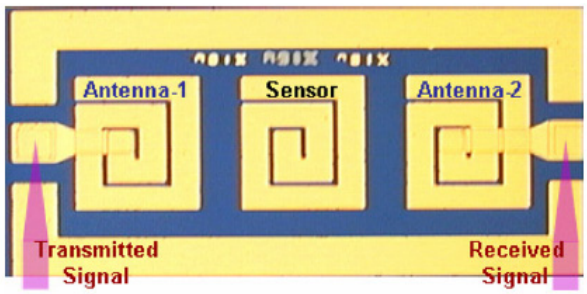


Figure 16. A plan-view micrograph of 270 μm × 270 μm on-chip sensor along with the on-chip antennas for communication (Melik et al., 2008)

Melik, Unal, Perkgoz, Puttlitz, and Demir (2009a) published another paper which demonstrated the use of a metamaterial-based wireless strain sensor for bio-medical applications. Metamaterials are artificial materials engineered to provide properties which may not be readily available in nature. Since metamaterials are fabricated for specific requirements, they could have extremely useful electromagnetic properties. Split Ring Resonator (SRR) is one of the geometrical configurations used in the fabrication of metamaterials. Melik et al. (2009a) used SRRs instead of the spiral case (Melik et al., 2008) because they have more gaps between the rings. These gaps reduce during compression and increase during tension both changing the capacitance and thus shifting the resonant frequency. This sensor was designed to be used for monitoring fracture healing and other biomedical applications. Figure 17 shows the shift in the resonant frequency of the SRR sensor under compressive strain.

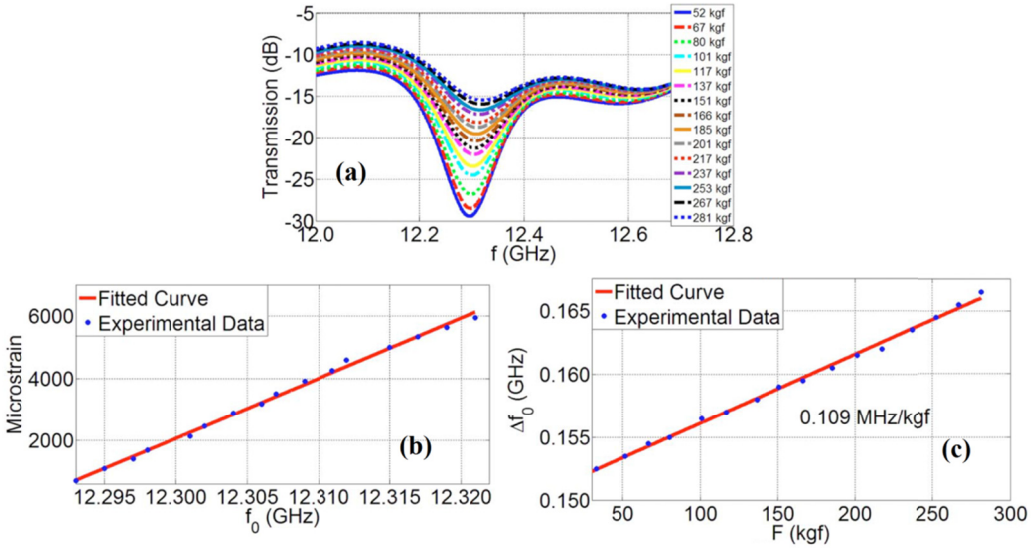


Figure 17. (a) Transmission spectra of the metamaterial strain sensor parameterized with respect to the external force, (b) the strain vs. the resonance frequency and (c) the resonance frequency shift vs. the applied force (Melik et al., 2009a)

Melik, Pergoz, Unal, Puttlitz, and Demir (2009b) reported that a flexible metamaterial sensor can be employed for effective strain measurement which could deliver greater sensitivity. These tape-based flexible SRR sensors also helped in reducing the non-linearity error. Figure 18 shows the step wise fabrication of this sensor. These tape-based flexible SRR sensors exhibit a significantly improved sensitivity level of 0.292 MHz/kgf with a substantially reduced nonlinearity error of 3% for externally applied mechanical loads up to 250 kgf. These resonators are very thin and hence could be bonded to the structure as a tape resulting in a very simple installation process.

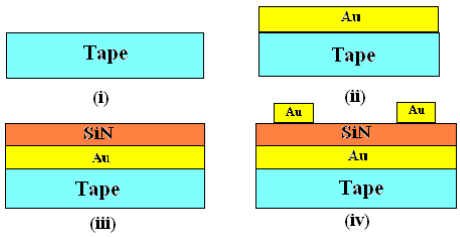


Figure 18. Fabrication procedure of the tape-based flexible sensor. 0.1 μm thick Au and Si3N4 deposited using lithography and lift off techniques (Melik et al., 2009b)

Melik, Unal, Perkgoz, Puttlitz, and Demir (2010) presented telemetric sensing of surface strains on different industrial materials using SRR based metamaterials. For wireless strain sensing, they utilized metamaterial array architectures for high sensitivity and low nonlinearity errors. Telemetric strain measurements were performed by observing operating frequency shift under mechanical deformation and these data were compared with commercially-available wired strain gauges. It was shown that hard material (cast polyamide) show low slope in the frequency shift vs. applied load curve (corresponding to high Young's modulus), whilst soft material (polyamide) exhibit high slope (corresponding to low Young's modulus). Figure 19 compares the change in the resistance of a conventional strain gauge compared to the change in the resonant frequency of the SRR array for the same applied load. This figure shows the low non-linearity error in the SRR strain sensor.

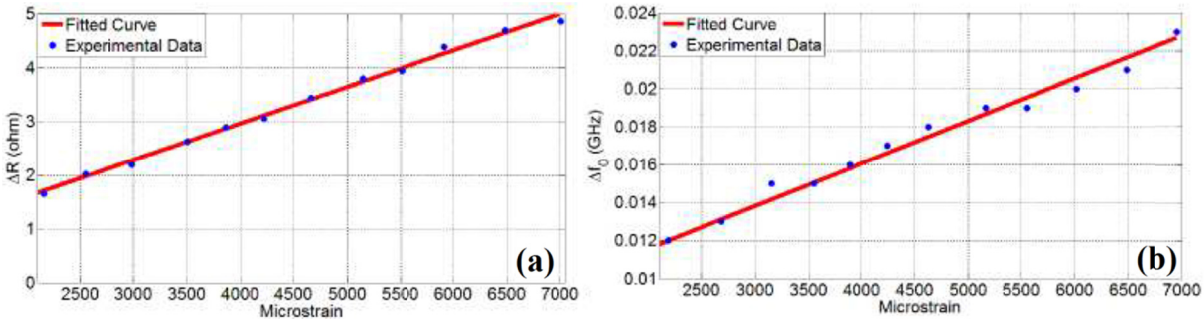


Figure 19. (a) Resistance change of the strain gauge and (b) Frequency change of the resonator with respect to the applied strain (Melik et al., 2010)

Other metamaterial resonator structures were also studied in the last few years, which follow the work by Melik and his colleagues. Ekmekci and Turhan-Sayan (2011) investigated the BC-SRR structure using numerical simulations. They suggested this sensor structure for applications such as pressure, temperature and humidity measurement. Li, Withayachumnankul, Chang, and Abbott (2011) extended this topic to terahertz frequencies. Again, the proposed resonator structures (I-shaped and crossed-I-shaped) were investigated using numerical simulations. Although these structures showed promising features, their performance has not yet been investigated using experimental measurements.

Naqui, Duran-Sindreu, and Martin (2011) introduced a resonator structure for position sensing applications. This structure consists of a SRR resonator at one side of a substrate material and a coplanar waveguide at the other side of the substrate. The principle of operation of this sensor is based on the symmetry of the resonator. When two rings of the resonator are rotated relative to each other, the resonance of the structure changes accordingly. This phenomenon was investigated theoretically and validated experimentally by fabrication and measurement of different SRR structures. However, in practice the sensor structure must be fabricated in a way to allow relative movement of two rings.

In a more recent study, Albishi, Boybay, and Ramahi (2012) proposed another metamaterial inspired resonator, Complementary-SRR (CSRR), as a sensor for crack detection. The structure of the proposed CSRR sensor is shown in Figure 20. The CSRR can be etched on a normal RF substrate and can be fed using a microstrip line behind the substrate. The CSRR has a stop band resonance which creates a strong electromagnetic field in the near field of the resonator. This resonant frequency shifts when the sensor is placed near a metallic structure with micro millimetre size surface crack compared to the same structure without any defects. Therefore, the resonant frequency of the CSRR can be used as an indicator for detection of crack in the structure. Experimental results from (Albishi et al., 2012) show the feasibility of crack detection using this sensor with high sensitivity. However, this sensor can only be used for near field measurements and required to be moved over the surface of the structure.

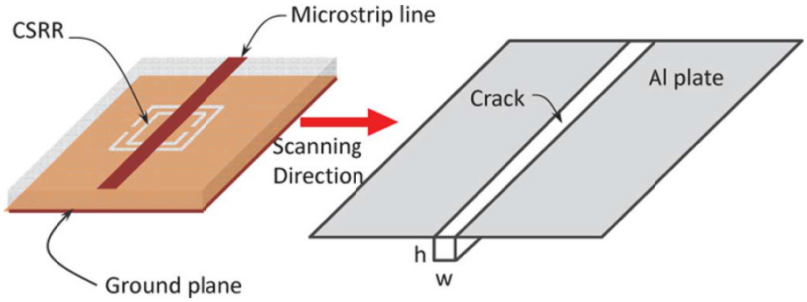


Figure 20. Schematic of the CSRR sensor and crack detection method (Albishi et al., 2012)

Tata, Huang, Carter, and Chiao (2009a) exploited a rectangular microstrip patch antenna for strain measurement. A rectangular patch antenna has the ability to resonate at two distinct frequencies, i.e. along its length and width. Thus it has the resonant frequencies f_{10} and f_{01} along its length and width, respectively (Figure 21). The resonant frequency is dependent on the electrical path length of the antenna. When the antenna is strained this length

changes and hence shifts the corresponding frequency (length or width depending on the direction of strain). The resonant frequency of the rectangular patch antenna can be derived as,

$$f_0 = C_1/(L_{e0} + C_2 h_0) \tag{6}$$

Where f_0 is the unstrained resonant frequency; L_{e0} is electrical length; h_0 is height of the patch; and C_1 and C_2 are antenna constants. After applying strain,

$$f_\varepsilon = \frac{C_1}{L_{e0}(1+\varepsilon_L) + C_2 h_0 (1 - \vartheta \varepsilon_L)} \tag{7}$$

Where f_ε is the strained resonant frequency; ε_L is the applied strain; and ϑ is the antenna substrate Poison's ratio.

Therefore, strain is derived as,

$$\varepsilon_L = C \frac{\Delta f}{f_\varepsilon} \tag{8}$$

The strain values can be calculated from the change in the resonant frequency. This sensor antenna was tested for a cantilever beam and the results were compared with analytical values. The resonant frequency of this antenna sensor was measured by connecting the antenna to a Vector Network Analyser (VNA) thorough a SMA connector and a coaxial cable.

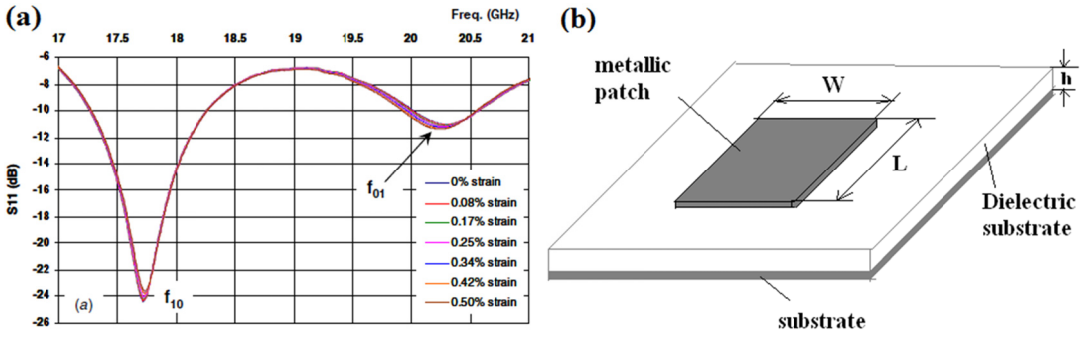


Figure 21. Rectangular microstrip patch antenna and its frequencies f_{10} and f_{01} (Tata et al., 2009a)

The technique required for wireless interrogation of microstrip patch antennas was introduced by Tata, Deshmukh, Chiao, Carter, and Huang (2009b) and Deshmukh, Mohammed, Tentzeris, Wu, and Huang (2009). In Tata et al. (2009b), the antenna sensor was excited by a dual polarized horn antenna which has polarizations along the length and width of the patch antenna to read both frequencies. Because the metallic host material also scatters the EM signals there are two scattering modes, namely the antenna and structural mode. A microwave switch was used which produced a phase shift of 180° to the antenna mode thereby normalizing and subtracting the structural mode for analysis. The experimental setup is illustrated in the Figure 22.

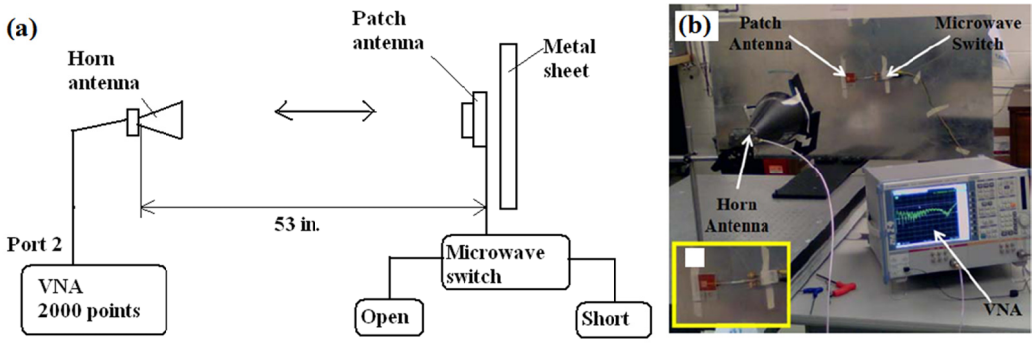


Figure 22. (a) Experimental setup diagram; (b) experimental setup (Deshmukh et al., 2009)

Deshmukh and Huang (2010) extended the previous work by utilizing a light activated microwave switch consisting of a pHEMT (pseudomorphic High Electron Mobility Transistor) and a photo cell to change the

terminating impedance of the antenna from open to short circuit. This change in circuit provides a 180° phase change and thereby helps in distinguishing the antenna backscattering mode from the structural backscattering mode. When the backscattering is measured using a horn antenna, a light beam is also sent to activate the switch in order to implement a phase change.

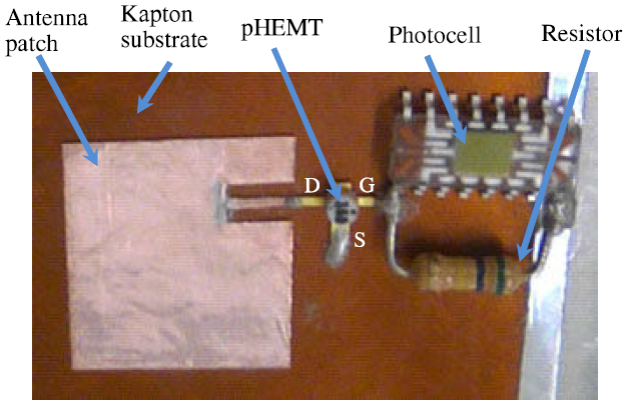


Figure 23. Passive antenna sensor with a light activated RF switch (Deshmukh & Huang, 2010)

Deshmukh and Huang (2010) studied the theoretical wireless interrogation range using the power budget model. The interrogation distance is,

$$R_{\max} = \frac{c}{4\pi f}\left[\frac{(D_{sd}D_{ds} + S_{11})P_t G_h^{\,2} G_s^{\,2}}{NF \times SNR}\right] \tag{9}$$

Where c is the speed of light; f is the frequency; D_{sd} and D_{ds} are insertion losses due to the switch; $|S_{11}|$is the scattering parameter; P_t is the transmitted power; NF is the noise factor; SNR is the signal to noise ratio; and G_h and G_s are gains of the horn and antenna sensor, respectively. This formula gives a good idea about the factors that affect the wireless interrogation range. The wireless range could be increased by increasing the transmitted power, antennas' gain, and reducing the SNR. It is mentioned that by increasing the interrogation power to 30 dBm and gain of the horn antenna to 20 dB, the interrogation distance could be up to 3.5 m.

Deshmukh et al. (2009), Erdmann, Deshmukh, and Huang (2010) and Mohammad and Huang (2010) tried to employ the rectangular patch antenna for sensing of fatigue crack growth. The patch antenna was placed on the crack in a lap joint structure and the shift in the resonant frequency with respect to crack growth was investigated (Figure 24). The effect of the resonant frequency due to the plate on top of the antenna sensor was also studied.

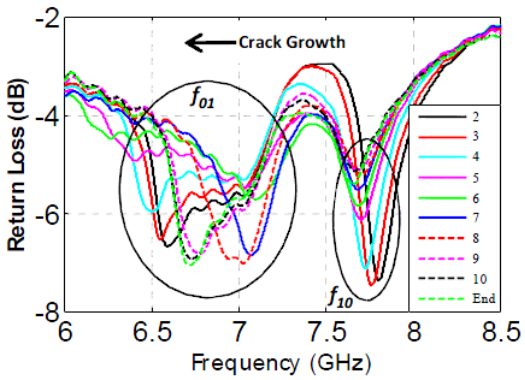


Figure 24. Shift of $|S_{11}|$ curves with crack growth (Erdmann et al., 2010)

The concept of crack detection using microstrip patch antennas was followed by Mohammad, Gowda, Zhai, and Huang (2012). Numerical simulations and experimental measurements were employed to detect the direction of the crack in the structure in addition of its presence. It was shown that a normalized frequency ratio (ratio of

frequencies of two modes of rectangular patch antenna) can represent a unique indication for crack orientation. However, cracks that are symmetric about the centre line of the patch antenna cannot be differentiated. Xu and Huang (2012) introduced multiplexing of dual-antenna arrays to cover a larger area of the structure for crack monitoring. A two-antenna array has two different frequencies and monitoring both antennas at the same time provides more information about the structure. However, the illuminating light must activate one antenna array each time to read the antenna information.

Daliri, Galehdar, John, Rowe, and Ghorbani (2010a) and Daliri, John, Galehdar, Rowe, and Ghorbani (2010b) developed a circular microstrip patch antenna (CMPA) strain sensor. It was shown that the resonant frequency of the CMPA is dependent on the radius of the patch antenna regardless of the material used as the substrate. Thus, as the material is strained, the resonant frequency shifts due to the change in the radius. The CMPA sensor was trialed on three different materials; namely aluminium, CFRP and glass fibre reinforced polymer (GFRP). The system was tested for different bending angles to determine the direction of strain. It was noted that depending on the strain along the magnetic or electrical plane of the antenna, the frequency shift was positive or negative. When the bending was in 45°, there was not much shift in the frequency. It was also interesting to note that in a GFRP plate, the frequency shift was not significant compared to aluminium and CFRP plates. The results showed that the antenna sensor is linear and effective in strain measurement. Computer simulations were carried out for 3-point bending of the antenna sensor on an aluminium plate and the shift in the frequency due to strain was determined. These values were compared with the experimental values and it was shown that these patch antennas could be employed as wireless strain sensors. The results showing the strain frequency shift relationship and the shift in the resonant frequency are shown in Figures 25, 26 and 27.

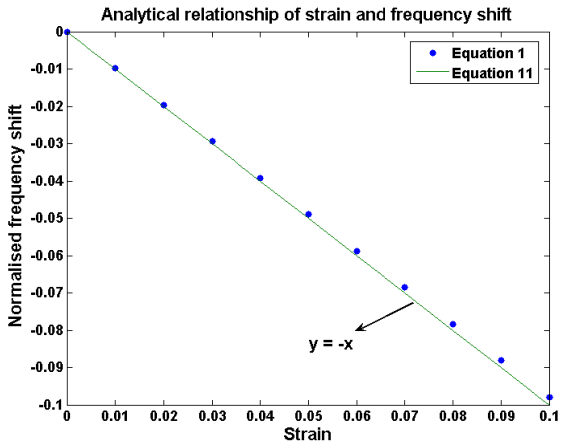


Figure 25. Linear (theoretical) relationship between strain and frequency shift (Daliri et al., 2010b)

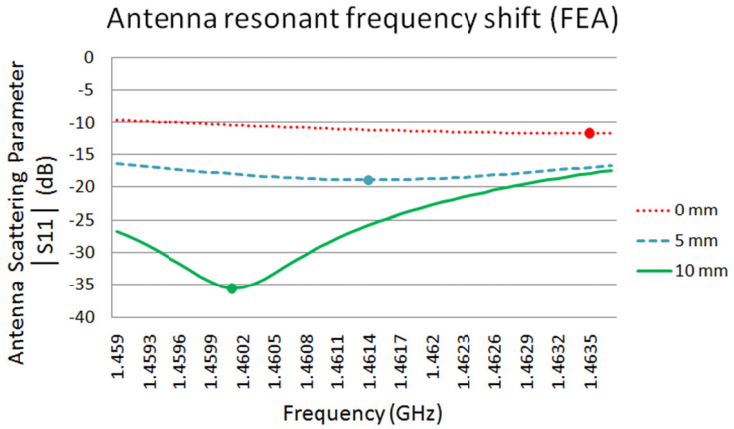


Figure 26. Simulation results of the antenna behaviour with 5 mm steps of bending (Daliri et al., 2010b)

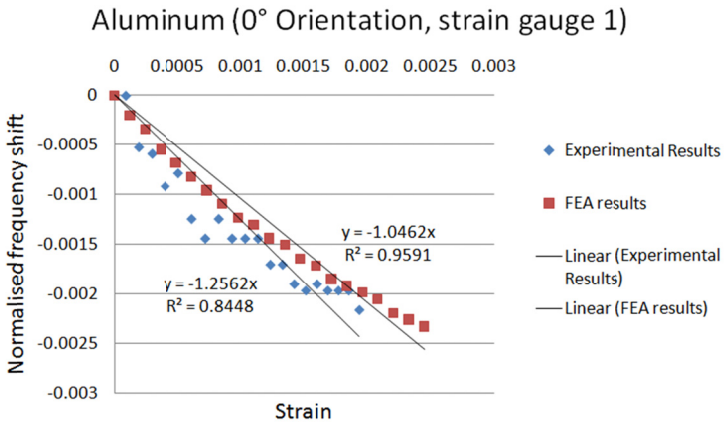


Figure 27. Comparison of measured and simulation results of strain and frequency shift for aluminium plate at 0° orientation (Daliri et al., 2010b)

Daliri, Galehdar, John, Rowe, and Ghorbani (2011a) studied the effect of slots in the CMPA on its directional capability and sensitivity. It was shown that introducing a slot in the CMPA introduces another resonant frequency which can help detect strain in two perpendicular directions. Also, this new sensor is more sensitive compared to a normal CMPA. To further tackle the directional problem of a normal CMPA, Daliri et al. (2011b) developed an omnidirectional CMPA strain sensor by introducing several slits on the edges of the CMPA. The resultant strain sensor is capable of measuring strain in all directions and it exhibits a five-fold reduction in the size of the antenna sensor and a three-fold increase in its sensitivity. A more comprehensive study on circular microstrip patch antennas is presented in (Daliri, Galehdar, Rowe, Ghorbani, & John, 2012a).

A slotted rectangular microstrip patch antenna (Salmani, Xie, Zheng, Zhang, & Zhang, 2011) and a rectangular microstrip patch antenna (Qian, Tang, Li, Zhao, & Zhang, 2012) have also been proposed to address sensitivity concerns in antenna strain sensors. Thai et al. (2011) introduced a novel idea to increase the sensitivity of microstrip patch antennas by adding an open loop to non-fed end of a rectangular microstrip patch antenna (Figure 28). The resonant frequency of the patch antenna depends on the impedance of the open loop, which can be altered by the loop length. However, the interference between the loop frequency and the antenna frequency may cause major problems. The manufacturing of this sensor requires special care.

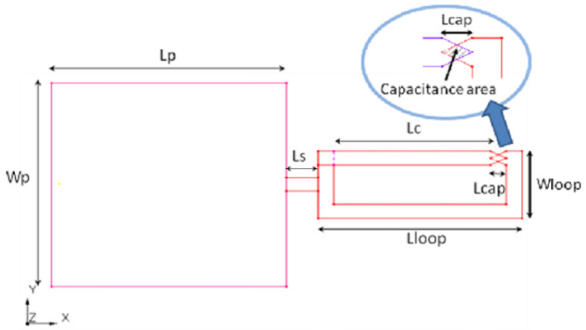


Figure 28. Details of the rectangular patch antenna design with an open loop (Thai et al., 2011)

Another example of using microstrip patch antennas for strain measurement is the application of multi-layer micro-fluidic stretchable radiofrequency electronics for wireless strain sensing (Cheng & Wu, 2011). A rectangular microstrip patch antenna fabricated using this technique was shown to be capable of measuring strain of up to 15% repeatedly based on the change in the strength of radio frequency (RF) signals emitted by the proposed antenna when connected to an RF transmitter. The RF transmitter is composed of miniaturized rigid active integrated circuits (ICs) and discrete passive components. This, with the addition of two AA batteries as a power source, increases the complexity and weight of the sensor unit, whilst decreasing its durability. Other concepts for strain measurement including printed dipole antennas (Jang & Kim, 2012) and dual-band microstrip patch antennas (Ahbe, Beer, Zwick, Wang, & Tentzeris, 2012) have also been proposed recently.

None of the above work on microstrip patch antenna strain sensors provided the answer for a passive wireless strain sensor which does not require additional integrated circuit units for wireless interrogation of the sensor. In the work by Deshmukh et al. (2009) and Tata et al. (2009a) the wireless interrogation of the rectangular patch antenna was achieved using an additional switch which requires wires or batteries. In the study of Deshmukh (2010) this switch was activated using a photo cell which further increases the complexity of the sensor unit and also requires a direct light source to activate the switch. In these works the resonant frequency of the patch antenna was read wirelessly; however, the strain was not measured. In other works on microstrip patch antennas a SMA connector was used to read the resonant frequency of the antenna sensor.

Daliri et al. (2012b) developed a method for wireless interrogation of CMPA strain sensors without the need for additional circuit elements. The CMPA was excited using a linearly polarised double ridged horn antenna to read its resonant frequency. This concept was studied using computational simulations and experimental measurements. The strain in aluminium and CFRP panels was measured wirelessly using this technique. However, the interrogation distance was limited to 5 cm. This technique also enables measuring strain in any desired direction because the linear horn antenna excites the CMPA in the direction of its polarization. By rotating the horn antenna the strain can be measured in the corresponding direction. Daliri et al. (2012c) further increased the interrogation distance of the CMPA sensor up to 20 cm by using a high quality factor CMPA. The high quality factor CMPA was developed using a substrate with low loss and high permittivity.

Regardless of the novelty and promising future of the discussed antenna sensor structures, there are certain key issues which have to be addressed to make these aforementioned works suitable for practical applications. Metamaterial sensors are very small and sensitive; however, they have not been employed to detect cracks till date and have practical limitations for monitoring strain in metallic structures. Moreover, there is not much information about the distance up to which these sensors could be wirelessly read. The work of Matsuzaki et al. (2009) involves the use of CFRP due to its good electrical conductivity. This technique cannot be extended to GFRP and other non-conducting structures.

Microstrip patch antenna sensors have been very useful in predicting the strain along with its direction. However, it is essential to design the antenna to be very sensitive to strain such that the shift in the frequency is more detectable. The wireless interrogation distance of these sensors needs to be improved significantly in order to make these sensors suitable for practical applications. Once these issues are addressed, these sensors could be installed in an array on the host structure for wireless SHM. Table 2 summarizes the important techniques that are provided in this section. Table 3 compares the various techniques covered in this article in general and in terms of important design parameters based on the authors' understanding of the sensor types. It is clearly evident from this table that each sensor type has its own advantages and limitations and no single technique has all the desirable properties.

Table 2. Various techniques investigated for use of resonators as strain sensors

Authors	Technique
Melik et al., 2009 a & b, 2010	Use of split rings, spirals and metamaterial based strain sensors whose resonant frequency changes due to the change in their dimensions.
Matsuzaki et al., 2009	Use of carbon fibres in the structure as a dipole antenna. The crack reduces the length of the dipole thereby shifting its frequency.
Tata et al., 2009 a & b; Deshmukh & Huang, 2010	Use of rectangular microstrip patch antennas whose resonant frequency changes with change in its dimension due to strain.
Daliri et al., 2010 a & b; 2011 a & b; 2012 a & b	Use of circular microstrip patch antennas whose resonant frequency changes with change in its dimension due to strain.
Yi et al., 2011 a & b; 2012 a & b	Use of rectangular microstrip patch antennas in combination with IC chips to form a RFID strain sensor whose resonant frequency changes with change in its dimension due to strain.

Table 3. Comparison of various types of sensors based on design parameters

Parameter	MEMS sensors	LC/RC circuit sensors	Antenna sensors
Power supply	Need battery for wireless monitoring.	LC circuits could receive power wirelessly. RC circuits need cables/wires.	Can receive power wirelessly from a transmitter antenna.
Active/Passive	Could be active and passive.	Could be active and passive	Passive.
Ability to detect strain direction	Could be designed to detect strain in any direction.	Generally unidirectional.	Could be designed to detect strain in any direction.
Wireless range	Can transmit signals over long distances.	Can transmit signals and receive power effectively only over few centimetres.	Has a small wireless range (< 1m).
Design complexity	Highly complex due to integration of various components.	Complex due to presence of inductor and capacitor integration.	Simple design due to the presence of only the antenna.
Size and weight	Large in size and heavy due to presence of battery and antenna.	Moderately large in size and less heavy compared to MEMS.	Smaller in size and much lighter compared to other sensors.

4. Conclusion

Structural health monitoring has been an important research area for the last few decades. The cumbersome use of wires to connect sensors with base stations has prompted researchers to investigate the feasibility of wireless SHM by incorporating various electromagnetic theories to overcome the limitations of the wired sensors.

Significant amount of research has been done in the field of MEMS systems and their application to SHM. Several different types of architectures with different specifications of microprocessors have been investigated. In order to address the power requirements of these sensing systems, few power harvesting techniques have also been investigated for effective functioning of these sensor systems. However, the design of such sensor units are relatively more complicated due to the incorporation of the sensing unit, signal conditioning unit, antennas and power harvesting devices. Thus, these sensors become expensive and the performance of these sensors has to be monitored regularly. In recent years, researchers have started using resonators and antennas as strain sensors where the resonant frequency of the resonator shifts due to the change in their dimensions during a structural deformation. These devices could be made small and simple which becomes elegant to install and read wirelessly. Moreover, the use of metamaterial based resonators could make these sensors very small thereby providing the feasibility for an array of sensors for damage monitoring.

In the use of patch antennas, along with the strain, the spatial orientation of the strain could also be wirelessly estimated which becomes crucial in terms of crack direction and growth prediction. However, the work on the use of resonators and antennas as sensors in SHM is presently at the embryonic stage. More effort towards these sensors could prove fruitful, especially work directed towards effective wireless SHM with good damage discrimination, sensitivity and linearity. It is envisaged that once these resonators or antennas can be designed to exhibit linear frequency shifts with strain, its utility can be extended to assess damage in structures, irrespective of the dielectric properties of the host structure. Given the comparative parameter-based assessment of the technologies discussed in this paper (in Table 3), it might be that the way ahead in wireless SHM should consist of a combination of the technologies discussed in this paper, since no single technique had all the desirable properties.

References

Ahbe, D., Beer, S., Zwick, T., Wang, Y., & Tentzeris, M. M. (2012). Dual-band antennas for frequency-doubler-based wireless strain sensing. *IEEE Antennas and Wireless Propagation Letters, 11*, 216-219. http://dx.doi.org/10.1109/LAWP.2012.2188014

Albishi, A. M., Boybay, M. S., & Ramahi, O. M. (2012). Complementary split-ring resonator for crack detection in metallic surfaces. *IEEE Microwave and Wireless Components Letters, 22*, 330-332. http://dx.doi.org/10.1109/LMWC.2012.2197384

Anton, S. R., & Sodano, H. A. (2007). A review of power harvesting using piezoelectric materials. *Smart Materials and Structures, 16*, R1-R21. http://dx.doi.org/10.1088/0964-1726/16/3/R01

Auweraer, H., & Peeters, B. (2003). International research projects on structural health monitoring: an overview. *Structural Health Monitoring, 2*(4), 341-358. http://dx.doi.org/10.1177/147592103039836

Bhattacharyya, R., Floerkemeier, C., & Sarma S. (2009). Towards tag antenna based sensing – an RFID displacement sensor. *IEEE International Conference on RFID*. http://dx.doi.org/10.1109/RFID.2009.4911195

Butler, J. C., Viglio, A. J., Vendi, F. W., & Walsh, S. M. (2002). Wireless passive resonant-circuit, inductively coupled, inductive strain sensor. *Sensors and Actuators A: Physical, 102*, 61-66. http://dx.doi.org/10.1016/S0924-4247(02)00342-4

Caizzone, S., & Marrocco, G. (2012). RFID-grids for deformation sensing. *IEEE International Conference on RFID*, 130-134. http://dx.doi.org/10.1109/RFID.2012.6193040

Chang, F. K. (2002). *Structural Health Monitoring-the Demands and Challenges*. Boca Raton: CRC Press.

Chang, P. C., Flatau, A., & Liu, S. C. (2003). Review paper: health monitoring of civil infrastructure. *Structural Health Monitoring, 2*(3), 257-267. http://dx.doi.org/10.1177/1475921703036169

Cheng, S., & Wu, Z. (2011). A microfluidic, reversibly stretchable, large-area wireless strain sensor. *Advanced Functional Materials, 21*(12), 2282-90. http://dx.doi.org/10.1002/adfm.201002508

Chuang, J., Thomson, D. J., & Bridges, G. E. (2005). Embeddable wireless strain sensor based on resonant RF cavities. *Review of Scientific Instruments, 76*, 094703-7. http://dx.doi.org/10.1063/1.2051808

Daliri, A., Galehdar, A., John, S., Rowe, W. S. T., & Ghorbani, K. (2010a). Circular microstrip patch antenna strain sensor for wireless structural health monitoring. *World Congress of Engineering 2010*, London, U.K., 1173-1178.

Daliri, A., Galehdar, A., John, S., Rowe, W. S. T., & Ghorbani, K. (2011a). Slotted circular microstrip patch antenna application in strain based structural health monitoring. *7th DSTO International Conference on Health & Usage Monitoring (HUMS 2011)*, Melbourne, Australia.

Daliri, A., Galehdar, A., John, S., Rowe, W. S. T., Ghorbani, K., & Wang, C. H. (2011b). Multidirectional circular microstrip patch antenna strain sensor. *ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems (SMASIS2011)*. http://dx.doi.org/10.1115/SMASIS2011-5065

Daliri, A., Galehdar, A., John, S., Wang, C. H., Rowe, W. S. T., & Ghorbani, K. (2012b). Wireless strain measurement using circular microstrip patch antennas. *Sensors and Actuators A: Physical, 184*, 86-92. http://dx.doi.org/10.1016/j.sna.2012.07.003

Daliri, A., Galehdar, A., John, S., Wang, C. H., Rowe, W. S. T., & Ghorbani, K. (2012c). Wireless strain sensors using electromagnetic resonators, ASME Conference on Smart Materials. *Adaptive Structures and Intelligent Systems (SMASIS2012)*.

Daliri, A., Galehdar, A., Rowe, W. S. T., Ghorbani, K., & John, S. (2012a). Utilising microstrip patch antenna strain sensors for structural health monitoring. *Journal of Intelligent Material Systems and Structures, 23*(2), 169-182. http://dx.doi.org/10.1177/1045389X11432655

Daliri, A., John, S., Galehdar, A., Rowe, W. S. T., & Ghorbani, K. (2010b). Strain measurement in composite materials using microstrip patch antennas. *ASME Conference on Smart Materials, Adaptive Structures and Intelligent Systems (SMASIS2010)*. pp. 591-598. http://dx.doi.org/10.1115/SMASIS2010-3703

Das, N. K., Khorrami, F., & Nourbakhsh, S. (1998). A new integrated piezoelectric-dielectric microstrip antenna for dual wireless actuation and sensing functions. *SPIE conference on Smart Electronics and MEMS*, San Diego, California. http://dx.doi.org/10.1117/12.320163

Deshmukh, S., & Huang, H. (2010). Wireless interrogation of passive antenna sensors. *Measurement Science and Technology, 21*, 035201. http://dx.doi.org/10.1088/0957-0233/21/3/035201

Deshmukh, S., Mohammed, I., Tentzeris, M., Wu, T., & Huang, H. (2009). Crack detection and monitoring using passive wireless sensor. *ASME conference on Smart Materials, Adaptive Structures and Intelligent Systems,*

pp. 511-516. http://dx.doi.org/10.1115/SMASIS2009-1326

Ekmekci, E., & Turhan-Sayan, G. (2011). Metamaterial sensor applications based on broadside-coupled SRR and V-Shaped resonator structures. *IEEE International Symposium on Antennas and Propagation (APSURSI)*, USA, 1170-1172. http://dx.doi.org/10.1109/APS.2011.5996492

Erdmann, J., Deshmukh, S., & Huang, H. (2010). Microwave antenna sensors for fatigue crack monitoring under lap joints. *ASME conference on Smart Materials, Adaptive Structures and Intelligent Systems (SMASIS)*. http://dx.doi.org/10.1115/SMASIS2010-3744

Housner, G. W., Bergman, L. A., Caughey, T. K., Chassiakos, A. G., Claus, R. O., Masri, S. F., … Yao, J. T. P. (1997). Structural control: past, present and future. *Journal of Engineering Mechanics*, 897-971. http://dx.doi.org/10.1061/(ASCE)0733-9399(1997)123:9(897)

Jang, S. D., & Kim J. (2012). Passive wireless structural health monitoring sensor made with a flexible planar dipole antenna. *Smart Materials and Structures, 21*, 027001. http://dx.doi.org/10.1088/0964-1726/21/2/027001

Jia, Y., & Sun, K. (2006b). Thick film wireless and powerless strain sensor. *SPIE Smart Structures and Materials, 6174*, 1-11. http://dx.doi.org/10.1117/12.660365

Jia, Y., Sun, K., Agosto, F. J., & Quinones, M. T. (2006a). Design and characterization of a passive wireless strain sensor. *Measurement Science and Technology, 17*, 2869-2876. http://dx.doi.org/10.1088/0957-0233/17/11/002

Li, J., Withayachumnankul, W., Chang, S., & Abbott, D. (2011). Metamaterial-based strain sensors. *Proceedings of the 7th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, Australia, 30-32. http://dx.doi.org/10.1109/ISSNIP.2011.6146571

Loh, K. J., Lynch, J. P., & Kotov, N. A. (2007). Passive wireless strain and pH sensing using carbon nanotube-gold nanocomposite thin films. *Proceedings of SPIE*, 6529. http://dx.doi.org/10.1117/12.715826

Lynch, J. P., & Loh, K. J. (2006). A summary review of wireless sensors and sensor networks for structural health monitoring. *The Shock and Vibration Digest, 38*(91). http://dx.doi.org/10.1177/0583102406061499

Mandel, C., Schussler, M., & Jakoby, R. (2011). A wireless passive strain sensor. *IEEE Sensors*, 207-210. http://dx.doi.org/10.1109/ICSENS.2011.6126942

Matsuzaki, R., Melnykowycz, M., & Todoroki, A. (2009). Antenna/sensor multifunctional composites for the wireless detection of damage. *Composites Science and Technology, 69*, 2507-2513. http://dx.doi.org/10.1016/j.compscitech.2009.07.002

Matsuzaki, R., & Todoroki, A. (2005). Wireless strain monitoring of tires using electrical capacitance changes with an oscillating circuit. *Sensors and Actuators A: Physical, 119*, 323-331. http://dx.doi.org/10.1016/j.sna.2004.10.014

Matsuzaki, R., & Todoroki, A. (2006). Wireless detection of internal delamination cracks in CFRP laminates using oscillating frequency changes. *Composites Science and Technology, 66*, 407-416. http://dx.doi.org/10.1016/j.compscitech.2005.07.016

Melik, R., Pergoz, N. K., Unal, E., Puttlitz, C., & Demir, H. V. (2008). Bio implantable passive on-chip RF-MEMS strain sensing resonators for orthopedic applications. *Journal of Micromechanics and Microengineering, 18*, 115017. http://dx.doi.org/10.1088/0960-1317/18/11/115017

Melik, R., Pergoz, N. K., Unal, E., Puttlitz, C., & Demir, H. V. (2009b). Flexible metamaterials for wireless strain sensing. *Applied Physics Letters, 95*, 181105. http://dx.doi.org/10.1063/1.3250175

Melik, R., Unal, E., Perkgoz, N. K., Puttlitz, C., & Demir, H. V. (2009a). Metamaterial-based wireless strain sensors. *Applied Physics Letters, 95*, 011106. http://dx.doi.org/10.1063/1.3162336

Melik, R., Unal, E., Perkgoz, N. K., Puttlitz, C., & Demir, H. V. (2010). Metamaterial based telemetric strain sensing in different materials. *Optics Express, 18*(5). http://dx.doi.org/10.1364/OE.18.005000

Mohammad, I., Gowda, V., Zhai, H., & Huang H. (2012). Detecting crack orientation using patch antenna sensors. *Measurement Science and Technology, 23*, 015102. http://dx.doi.org/10.1088/0957-0233/23/1/015102

Mohammad, I., & Huang, H. (2010). Monitoring fatigue crack growth and opening using antenna sensors. *Smart Materials and Structures, 19*, 055023. http://dx.doi.org/10.1088/0964-1726/19/5/055023

Naqui, J., Duran-Sindreu, M., & Martin, F. (2011). Novel sensors based on the symmetry properties of split ring resonators (SRRs). *Sensors, 11*(8), 7545-7553. http://dx.doi.org/10.3390/s110807545

Occhiuzzi, C., Paggi, C., & Marrocco G. (2011). Passive RFID strain-sensor based on meander-line antennas. *IEEE Transactions on Antennas and Propagation, 59*(12), 4836-4840. http://dx.doi.org/10.1109/TAP.2011.2165517

Qian, Z., Tang, Q., Li, J., Zhao, H., & Zhang, W. (2012). Analysis and design of a strain sensor based on a microstrip patch antenna. *International Conference on Microwave and Millimetre Wave Technology (ICMMT)*. http://dx.doi.org/10.1109/ICMMT.2012.6230473

Salmani, Z., Xie, Y., Zheng, G., Zhang, H., & Zhang, H. (2011). Application of antenna in strain measurement, *International Workshop on Antenna Technology (iWAT): Small Antenna, Novel Structures and Innovative Metamaterials, Hong Kong: IEEE*, 336-339. http://dx.doi.org/10.1109/IWAT.2011.5752311

Spencer, Jr. B. F., Ruiz-Sandoval, M. E., & Kurata, N. (2004). Smart sensing technologies: opportunities and challenges. *Structural Control and Health Monitoring, 11*(4), 349-368. http://dx.doi.org/10.1002/stc.48

Tan, L. E., Pereles, B. D., Shao, R., Ong, J., & Ong, K. G. (2008). A wireless, passive strain sensor based on the harmonic response of magnetically soft materials. *Smart Materials and Structures, 17*(2), 025015. http://dx.doi.org/10.1088/0964-1726/17/2/025015

Tata, U., Deshmukh, S., Chiao, J. C., Carter, R., & Huang, H. (2009b). Bio-inspired sensor skins for structural health monitoring. *Smart Materials and Structures, 18*(10), 104026. http://dx.doi.org/10.1088/0964-1726/18/10/104026

Tata, U., Huang, H., Carter, R. L., & Chiao, J. C. (2009a). Exploiting a patch antenna for strain measurements. *Measurement Science and Technology, 20*, 015201. http://dx.doi.org/10.1088/0957-0233/20/1/015201

Thai, T. T., Aubert, H., Pons, P., Plana, R., Tentzeris, M. M., & DeJean, G. R. (2011). A newly developed radio frequency wireless passive highly sensitive strain transducer. *IEEE Sensors*, 211-214. http://dx.doi.org/10.1109/ICSENS.2011.6127239

Umbrecht, F., Wendlandt, M., Juncker, D., Hierold, C., & Neuenschwander, J. (2005). A wireless implantable passive strain sensor system. *IEEE Sensors*. http://dx.doi.org/10.1109/ICSENS.2005.1597627

Wang, C. H., & Chang, F. K. (2005). Scattering of plane waves by a cylindrical inhomogeneity. *Journal of Sound and Vibration, 282*, 429-451. http://dx.doi.org/10.1016/j.jsv.2004.02.023

Wang, C. H., & Rose, L. R. F. (2003). Wave reflection and transmission in beams containing delamination and inhomogeneity. *Journal of Sound and Vibration, 264*, 851-872. http://dx.doi.org/10.1016/S0022-460X(02)01193-8

Watters, D. G., Jayaweera, P., Bahr, A. J., & Huestis, D. L. (2001). Design and performance of wireless sensors for structural health monitoring. *Proceedings of Review of Progress in Quantitative Nondestructive Evaluation, 21A*, 969-976. http://dx.doi.org /10.1063/1.1472901

Xu, X., & Huang, H. (2012). Multiplexing passive wireless antenna sensors for multi-site crack detection and monitoring. *Smart Materials and Structures, 21*, 015004. http://dx.doi.org/10.1088/0964-1726/21/1/015004

Yi, X., Cho, C., Fang, C. H., Cooper, J., Lakafosis, V., Vyas, R., … Tentzeris, M. M. (2012a). Wireless strain and crack sensing using a folded patch antenna. *6th European Conference on Antennas and Propagation (EUCAP)*, 1678-1681. http://dx.doi.org /10.1109/EuCAP.2012.6206690

Yi, X., Fang, C. H., Cooper, J., Cho, C., Vyas, R., Wang, Y., … Tentzeris, M. M. (2012b). Strain sensing through a passive wireless sensor array. *20th Analysis and Computation Specialty Conference*, 117-126. http://dx.doi.org/10.1061/9780784412374.011

Yi, X., Wu, T., Lantz, G., Wang, Y., Leon, R. T., & Tentzeris, M. M. (2011b). Thickness variation study of RFID-based folded patch antennas for strain sensing. *Proceedings of SPIE Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*. http://dx.doi.org/10.1117/12.879868

Yi, X., Wu, T., Wang, Y., Leon, R. T., Tentzeris, M. M., & Lantz, G. (2011a). Passive wireless smart-skin sensor using RFID-based folded patch antennas. *International Journal of Smart and Nano Materials, 2*(1), 22-38. http://dx.doi.org/10.1080/19475411.2010.545450

Call for Manuscripts

Modern Applied Science (MAS) is an international, double-blind peer-reviewed, open-access journal, published by the Canadian Center of Science and Education. It publishes original research, applied, and educational articles in all areas of applied science. It provides an academic platform for professionals and researchers to contribute innovative work in the field. The scopes of the journal include, but are not limited to, the following fields: agricultural and biological engineering, applied mathematics and statistics, applied physics and engineering, chemistry and materials sciences, civil engineering and architecture, computer and information sciences, energy, environmental science and engineering, mechanics. The journal is published in both print and online versions. The online version is free access and download.

We are seeking submissions for forthcoming issues. All manuscripts should be written in English. Manuscripts from 3000–8000 words in length are preferred. All manuscripts should be prepared in MS-Word format, and submitted online, or sent to: mas@ccsenet.org

Paper Selection and Publishing Process

a) Upon receipt of a submission, the editor sends an e-mail of confirmation to the submission's author within one to three working days. If you fail to receive this confirmation, your submission e-mail may have been missed.

b) Peer review. We use a double-blind system for peer review; both reviewers' and authors' identities remain anonymous. The paper will be reviewed by at least two experts: one editorial staff member and at least one external reviewer. The review process may take two to four weeks.

c) Notification of the result of review by e-mail.

d) If the submission is accepted, the authors revise paper and pay the publication fee.

e) After publication, the corresponding author will receive two hard copies of the journal, free of charge. If you want to keep more copies, please contact the editor before making an order.

f) A PDF version of the journal is available for download on the journal's website, free of charge.

Requirements and Copyrights

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the authorities responsible where the work was carried out, and that, if accepted, the article will not be published elsewhere in the same form, in English or in any other language, without the written consent of the publisher. The editors reserve the right to edit or otherwise alter all contributions, but authors will receive proofs for approval before publication.

Copyrights for articles are retained by the authors, with first publication rights granted to the journal. The journal/publisher is not responsible for subsequent uses of the work. It is the author's responsibility to bring an infringement action if so desired by the author.

More Information

E-mail: mas@ccsenet.org

Website: www.ccsenet.org/mas

Paper Submission Guide: www.ccsenet.org/submission

Recruitment for Reviewers: www.ccsenet.org/reviewer

The journal is peer-reviewed
The journal is open-access to the full text
The journal is included in:

CABI
Chemical Abstracts database
DOAJ
EBSCOhost
Excellence in Research Australia (ERA)
Google Scholar
LOCKSS
Open J-Gate
ProQuest
Scopus
Standard Periodical Directory
Ulrich's
Universe Digital Library

Modern Applied Science
Monthly

Publisher Canadian Center of Science and Education
Address 1120 Finch Avenue West, Suite 701-309, Toronto, ON., M3J 3H7, Canada
Telephone 1-416-642-2606
Fax 1-416-642-2608
E-mail mas@ccsenet.org
Website www.ccsenet.org/mas

ISSN 1913-1844