Trends in Undergraduate Research (2019) 2(2): a1-12
https://doi.org/10.33736/tur.1846.2019

Brief Documentation on Selected Timber-Related Plant Species with Commercial Value in Kubah National Park, Sarawak

MOHAMAD FHAIZAL BIN MOHAMAD BUKHORI*, MOHD RIDWAN BIN ABD RAHMAN, MUHAMAD IKHWAN BIN IDRIS, MOHD AMINUDIN BIN MUSTAPHA, MOHAMAD RAZIF BIN OTHMAN & NOORAZMAN BIN SAMSUDIN

Biology Unit, Centre for Pre-University Studies, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia.
*Corresponding author: mbmfhaizal@unimas.my

ABSTRACT

A comprehensive learning experience in biodiversity-related discipline is the only practical way to introduce students into the world of flora. Thus, this study was carried out at Kubah National Park as the park is known for its flora and fauna diversity to investigate plant species which are of timber-related importance. The study identified and documented seven different plant families with 12 species of timber-related usage at Kubah National Park.

Keywords: Biology, foundation, Kubah, Sarawak, timber

INTRODUCTION

Malaysia is well known as mega diversity country particularly in East Malaysia including Sarawak and Sabah. Out of 18 million hectare (mil ha) of total forested area in the country including the 444 protected areas nationwide, five mil ha are managed by the respective federal and state forest authorities as resources for timber production (Cedergren, Falck, Garcia, Goh & Hagner, 2002; Lee & Watling, 2005; MNRE, 2015; Pakhriazad & Mohd Hasmadi, 2010; Teo, Chai & Phua, 2013). This gazetted area was known as one of the world's top timber producer in furniture and household industry-linkage (MTII, 2011; Pakhriazad & Mohd Hasmadi, 2010; SCMP, 2017).

Meanwhile, the remaining 13 mil ha areas are preserved and conserved as national parks and for wildlife including Kubah National Park (KNP). Generally, KNP in Matang, Kuching was established in 1989. The park is situated on massive sandstone ridge with its three mountain peaks, Gunung Serapi (911 meters, m high), Gunung Selang and Gunung Sendok (Abu Bakar, Radam, Samdin & Yacob, 2016; Boyce & Wong, 2008; Brahim, 2005; Das & Charles, 1993; Hanan, 2014; Kamri, 2013; Lateef, Sepiah & Bolhassan, 2015; Lesley *et al*., 2016; Meekiong, Latiff, Tawan & Miraadila, 2012; Pearce, 1992; Poulsen & Leong-Škorničková, 2017; SF, 2003; Teo *et al.*, 2013; Wahab, 2012; Yee & Chin, 1989).

To date, the country is in ongoing transition becoming a developed nation and has exerted various pressures on local biodiversity, leaving many valuable heritages vulnerable with some even facing threats of extinction and habitats that are degrading problems (Demies, Lading & Silang, 2008; Teo *et al*., 2013). As reported by the MNRE (2015), nearly half of the nation's plant diversity is facing various levels of threat. Therefore, documenting and recording the existence of timber-related plant species in KNP are timely.

In order to promote, inculcate and preserve awareness of the importance of ecology, biodiversity and environments to the students, among the action taken was by introducing the relevant topic and issues in the academic syllabus. On that note, further measures need to be adopted and practice towards the sustainable utilisation of the biodiversity resource on disseminating ecology, biodiversity and environment education such as Kubah National Park. The park was chosen since it promotes conservation education and nature study and serves as a recreational facility (Abu Bakar *et al*., 2016).

The park was opened to the public in 1995. Ever since, the park has received increasing numbers of visitors owing to its exceptional biodiversity, ecosystem, ecology, environment, habitat, species and nature including plants, fungi, mammals, amphibians, reptiles and insects. The parks cover an area of 2,230 ha and comprise of heavily forested slopes and ridges of the Serapi range (Brahim, 2005; Meekiong *et al*., 2012; Pearce, 1992; Teo *et al*., 2013).

Kubah National Park's most renowned feature is its orchids and palmetum with almost hundred different palm species found in an area of just over 22 km^2 (Boyce & Wong, 2008; Brahim, 2005; Lesley *et al*., 2016; Pearce, 1992; Poulsen & Leong-Škorničková, 2017; Vermeulen & Lamb, 2011) besides numerous other tropical rainforest trees including timber-related plant species such as mixed Dipterocarpaceae, Myrtaceae, Myristicaeae, Burseraceae and Lauraceae (Abang Bohari, 2015; Corlett & Primack, 2005; Kanzaki, Yap, Okauchi, Katsuhiko & Yamakura, 2003; Sasaki, 2006; Teo *et al*., 2013). The most dominant species in terms of importance value is *Syzygium havilandii*, followed by *Hopea dryobalanoides*, *Shorea macroptera*, *Santiria tomentosa* and *Shorea parvifolia* (Abang Bohari, 2015; Appanah & Turnbull, 1998; Kanzaki *et al*., 2003; MTII, 2011; Pakhriazad & Mohd Hasmadi, 2010; Roszaini & Salmiah, 2015; Soepadmo & Wong, 1995).

The study was conducted to achieve the learning objective of ecology, biodiversity and environmental module by addressing certain emerging environmental issues on ecosystems, natural resources and environment. The insights gained from the study can be applied to many categories of current and future environmental issues (Abu Bakar *et al*., 2016; Kamri, 2013; Kerfahi, Tripathi, Lee, Edwards & Adams, 2014; Mahidin & Sofwan, 2012; MNRE, 2006; 2015).

In order to comprehend the study, there are three components within this study, namely the acquisition of systematic understanding of environmental issues, the development of innovative study methods and tools, and the collection and dissemination of comprehensive environmental data (Appanah & Turnbull, 1998; Kanzaki *et al*., 2003; MNRE, 2006; 2015; MRPE, 2012; Sasaki, 2006; Vincent, 2002). All of the components are necessary in environmental problem-solving issue which is based on a foundation of biological study.

On that note, the objective of this study was to align with the course learning outcomes by promoting a comprehensive understanding and perception of the student's learning unit in the class towards biodiversity, ecology and environment including identification, protection, prevention and preservation of timber-related plants species in KNP. At the same time, the students observe and identify the diversity of timber-related plant species in KNP.

MATERIALS & METHODS

Study Site

The study area is located in Kubah National Park (1° 36' 45.9"N, 110° 11' 49.2"E) (Figure 1). The park is located around 25 km from Kuching city centre, and partially undisturbed natural forest situated on a small sandstone plateau which includes Gunung Selang, Sendok and Serapi. The parks is also consisting of five main vegetation types of forest namely lowland mixed dipterocarp, kerangas, alluvial, submontane and high mixed dipterocarp forest (Abang Bohari, 2015; Boyce & Wong, 2008; Brahim, 2005; Meekiong *et al*., 2012; Poulsen & Leong-Škorničková, 2017; SFD, 2003; Vermeulen & Lamb, 2011).

Field Observation

The study was conducted by observation and documentation activities on selected timber-related plant species by trekking; beginning from the Park Headquarters at 0 m up to the Waterfall Trail at 1,600 m distance (150-450 m altitude), with the detailed documentation and recording was done at every 100 m (Figure 2). The sampling method was set out to be random and all vascular trees that have closed distance observe (approximately 5 m) along the trekking were recorded and documented.

The trekking passed through kerangas forest, mixed dipterocarp forest until it reaches the riverine forest at Rayu river valley, which takes about 2 hr. The trail demonstrated some steep and slippery sections, and consistently descending into the river valley.

Timber-related plant species identified along the trekking journey were recorded. The observation and documentation were made by hand notes, taking photographs and videos. Preliminary identification was made during the visit, but the detailed analysis was made once returned to the campus.

Timber-Related Plant Species Identification

Records of timber-related plant species in KNP which were initially recorded were further categorized according to the scientific and local name, and commercial purposes. A brief analysis of the collected data was done based on reference metadata, scientific and academic books, articles, journals and official-related website revision established prior fieldwork made to finalize the obtained results.

The plant species were identified using keys including plant form or shape, size, the area or habitat of growing, area or habitat characteristics, bark characteristics, and unavailable characteristics at time of observation such as colour and size of seeds or fruit (Abang Bohari, 2015; Appanah & Turnbull, 1998; Ashton, 2004; Cedergren *et*

al., 2002; Corlett & Primack, 2005; Demies *et al*., 2008; Eyssartier, Stubbe, Walleyn & Verbeken, 2009; IUCN, 2017; Kanzaki *et al*., 2003; Lin, Chung & Peng, 2017; Meekiong *et al*., 2012; Mohd Isa, 2013; Poulsen & Leong-Škorničková, 2017; Sasaki, 2006; SF, 2003; Soepadmo & Wong, 1995; Vincent, 2002). The documents were also compared with Sarawak Herbarium, Sarawak Forestry Department.

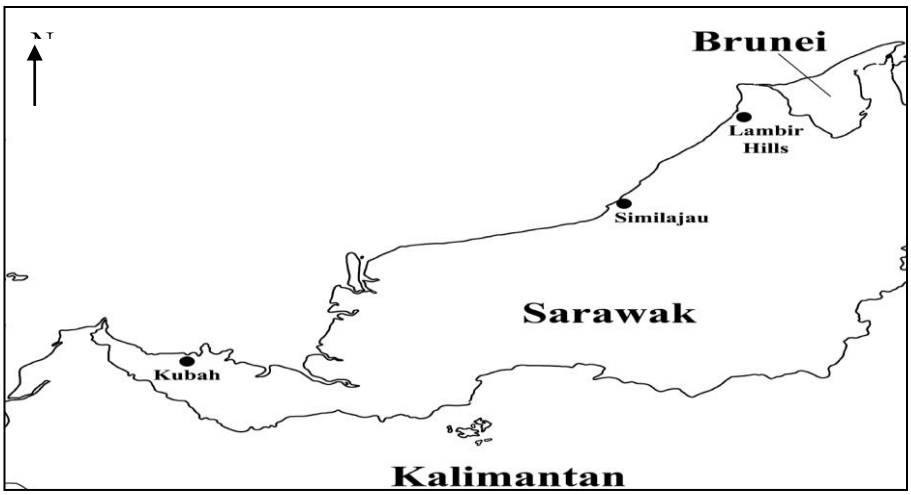

Figure 1. Aerial map of Kubah National Park, Sarawak (Dow & Reels, 2010).

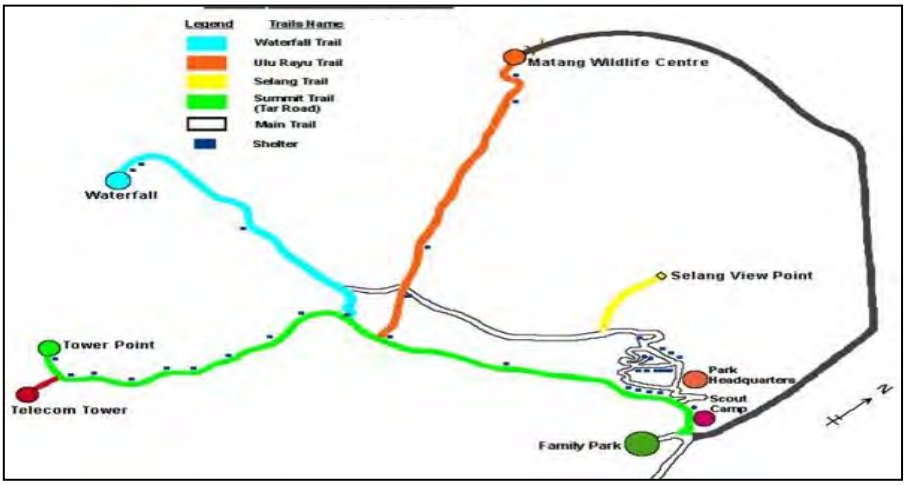

Figure 2. Trail of the study area (Sarawak Tourism, 2018).

RESULTS & DISCUSSION

Site Characteristics

The study area along the Waterfall Trail has encountered the richness of kerangas forest with the first part of the forest is a small, open and permanently saturated area providing with a fascinating patch of plants adapted to the sandy ground. The path also descends steeply into dipterocarp forest valley with a few observations and documentation of timber-related plant species (Abang Bohari, 2015; Boyce & Wong, 2008; Brahim, 2005; Pearce, 1992).

Timber-Related Plants Species

The study has discovered twelve timber-related plant species along the study area. Eight species recorded under Dipterocarpaceae family were identified as *Shorea albida*, *Shorea rugosa*, *Shorea dealbata*, *Shorea pauciflora, Dipterocarpus baudii*, *Dipterocarpus sarawakensis*, *Dryobalanops beccarii*, and *Dryobalanops oblongifolia* (Figure 3, A-H). Meanwhile, each one in four species was recorded under Euphorbiaceae family known as *Endospermum diadenum*; Ixonanthaceae family, *Ixonanthes icosandra*; Fabaceae family, which is *Koompassia malaccensis*; and Cannabaceae family, *Gironniera nervosa* (Figure 3, I-L).


A
B
C
D
E
F
G
H

Figure 3. Timber-related plant species in Kubah National Park. A. *Shorea albida*, B. *Shorea rugosa*, C. *Shorea dealbata*, D. *Shorea pauciflora*, E. *Dipterocarpus baudii*, F. *Dipterocarpus sarawakiensis*, G. *Dryobalanops beccarii*, H. *Dryobalanops oblongifolia*, I. *Endospermum diadenum*, J. *Ixonanthes icosandra*, K. *Koompassia malaccensis*, and L. *Gironniera nervosa*.

Dipterocarpaceae is the largest family dominating Sarawak's rainforest, and giving their name mixed dipterocarp rainforest. The name Dipterocarpaceae refers to the seeds, which have two or more wings, enabling them to disperse by wind and gravity, like little spinning helicopters (Ashton, 2004; Cedergren *et al*., 2002; Corlett & Primack, 2005; SF, 2003; Teo *et al*., 2013). Meanwhile, *Shorea* is the largest genus of hardwood and semi-hardwood trees within this family (Lateef, Sepiah & Bolhassan, 2016; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995). Many of these species are of significant commercial value for the timber-related industry (Ismaili, Openg, Abdul Rahim & Duju, 2016; MTII, 2011; Roszaini & Salmiah, 2015; Roszaini *et al*., 2017; Wong, 1995) (Figure 4 and Table 1).

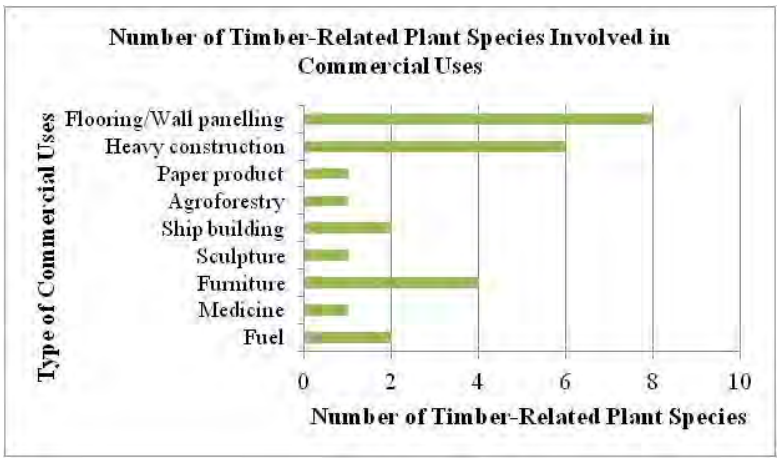

Figure 4. Number of timber-related plant species involved in commercial uses.

Table 1. List of timber-related plants species with possible value and residential observed in Kubah National Park, Sarawak (Ashton, 2004; IUCN, 2017; SF, 2003).

No.	Taxonomic Level	Habitat and Ecology	Value and Residential
1.	Order: Malvales Family: Dipterocarpaceae Genus: *Shorea* Species: *Shorea albida* Sym Local/Common Name (Vernacular): Meranti Merah Muda Alan Batu	Terrestrial Swamp forest Up to 600 m elevation	Used for timber industry. Classified as Endangered in IUCN Red List. Classified as Vulnerable in Sarawak Plant Red List. The species can range in size from a sub-canopy tree around 20 m tall to an emergent tree up to 75 m tall with a vast, diffuse canopy. The straight cylindrical bole can be 50 to 250 cm in diameter as for the larger trees with large buttresses up to 4 m high. The peat-swamp forest habitats where this tree grows are seriously threatened as regeneration is reported to be non-existent.
2.	Order: Malvales Family: Dipterocarpaceae Genus: *Shorea* Species: *Shorea rugosa* Local/Common Name (Vernacular): Meranti Merah Tua Seraya Kerukup	Terrestrial Sandy soils or heath forest Up to 400 m elevation	Used for interior and exterior panelling and joinery, light carpentry, furniture, etc. Classified as Critically Endangered in IUCN Red List. Classified as Vulnerable in Sarawak Plant Red List. The species grows on hills in lowland mixed dipterocarp forest. This huge tree has elliptical, papery, wavy leaves that measured up to 16 cm long and 7 cm wide. The bark can be almost smooth as on young trees. However, when the trees ages, furrows develop, which deepen as time passed. The inner bark is light brown to yellowish and has a slightly stringy texture.
3.	Order: Malvales Family: Dipterocarpaceae Genus: *Shorea* Species: *Shorea dealbata* Local/Common Name (Vernacular): Meranti Putih Meranti Bumbung	Terrestrial Sandy soils and on low hills forest Up to 150 m elevation	Used for timber industry. Classified as Critically Endangered in IUCN Red List. Classified as Vulnerable in Sarawak Plant Red List. Classified as Endangered in Malaysia Plant Red List. Classified as Endangered in Peninsular Malaysia Plant Red List. The species can grow up to 30 m tall. This species grows frequently in kerangas forest and sometimes in swampy land. The species is lightweight hardwood. The heartwood is yellowish white when freshly cut, but gradually become yellowish brown and slightly more distinct from the sap wood on exposure.
4.	Order: Malvales Family: Dipterocarpaceae Genus: *Shorea* Species: *Shorea pauciflora* Local/Common Name (Vernacular): Meranti Nemesu	Terrestrial Well-drained soil, lowland and hill forest Up to 700 m elevation	Used for timber industry. Found especially in riverside habitats. Classified as Endangered in IUCN Red List. Classified as Least Concern in Sarawak Plant Red List. Classified as Least Concern in Malaysia Plant Red List. The species is threatened by habitat loss. An emergent tree in undisturbed mixed dipterocarp forest at elevation up to 900 m. The species is mostly found on hillsides with clay to sandy soils. This tree can grow up to 60 to 220 cm in dia. with stout buttresses up to 4 m high.

Table 1. Cont…

5.	Order: Malvales Family: Dipterocarpaceae Genus: *Dipterocarpus* Species: *Dipterocarpus baudii* Local/Common Name (Vernacular): Keruing	Terrestrial Low-lying, well-drained or semi-swampy forests Up to 800 m elevation	Harvested for local use and trade. Yields ole-resin which is used locally for caulking boats and illumination. General construction timber such as carpentry, panelling, etc. Classified as Vulnerable in IUCN Red List. Classified as Least Concern in Malaysia Plant Red List. The species usually located commonly in tropical forest. This emergent tree species grows up to 40 m high with ovate-elliptic leaves up to 250 mm long and has seeds with winged lobes that are 150 to 180 mm long.
6.	Order: Malvales Family: Dipterocarpaceae Genus: *Dipterocarpus* Species: *Dipterocarpus sarawakensis* Local/Common Name (Vernacular): Keruing Layang	Terrestrial Alluvial soils, lowland and on hills forest Up to 1600 m elevation	Used for timber industry. Classified as Near Threatened in Sarawak Plant Red List. Classified as Critically Endangered in Malaysia Plant Red List. Classified as Critically Endangered in Peninsular Malaysia Plant Red List. The species is an emergent tree species that can attain heights above 50 m and grows to 1 m in dia. This species has a short petiole with short and dense indumentums, broadly ovate and obtuse leaves with a wavy margin that helps this species is different from the other *Dipterocarpus* species. This locally abundant tree grows in leached sandy soils on low coastal hills.
7.	Order: Malvales Family: Dipterocarpaceae Genus: *Dryobalanops* Species: *Dryobalanops beccarii* Local/Common Name (Vernacular): Kapur Bukit	Terrestrial Sandy soils, on hills and ridges forest Up to 700 m elevation	Used by the Iban for planks, beams, furniture, etc. May last for 15 years in the dry, three to four years if used for boat. Classified as Endangered in IUCN Red List. Classified as Least Concern in Sarawak Plant Red List. Classified as Endangered in Malaysia Plant Red List. Classified as Endangered in Peninsular Malaysia Plant Red List. The species is a large evergreen tree, with a large, globose crown. This species is a large emergent tree and can grow up to 65 m tall. This locally abundant tree grows in leached sandy soil and along the stream. It is a hardwood trees.
8.	Order: Malvales Family: Dipterocarpaceae Genus: *Dryobalanops* Species: *Dryobalanops oblongifolia* Local/Common Name (Vernacular): Kapur Kelansau	Terrestrial Poorly drained and on hillsides forest Up to 600 m elevation	The Iban used for planks, beams, etc. It may last for 15 years in the dry but only two years for boats. Classified as Least Concern in IUCN Red List. The species is found in at least one protected area which is KNP but is threatened elsewhere due to habitat loss. It is an emergent tree up to 60 m tall also found in mixed dipterocarp forest on sandy clay soils.

Table 1. Cont…

9.	Order: Malpighiales Family: Euphorbiaceae Genus: *Endospermum* Species: *Endospermum diadenum* Local/Common Name (Vernacular): Terbulan	Terrestrial Lowland forest Up to 1000 m elevation	Good for firewood. Classified as Not Evaluated in Malaysia Plant Red List. The species is a mid-canopy tree with a diffused, domed shaped crown. It can grow up to 36 m tall. The surface of the bark is smooth, becomes wrinkled to scaly in patches and grey fawn while the inner bark is thick and cream with orange flecks. This species grows in open places of undisturbed mixed dipterocarp and kerangas forest.
10.	Order: Malpighiales Family: Ixonanthaceae Genus: *Ixonanthes* Species: *Ixonanthes icosandra* Local/Common Name (Vernacular): Pagar Anak Inggi Burong	Terrestrial On slopes and ridges forest Up to 900 m elevation	The Iban use for building boats. It may last for about one to half years. Classified as Not Evaluated in Malaysia Plant Red List. The species is an evergreen tree with a large, spreading, moderately dense crown which can grow up to 35 m tall. The deeply fluted bole can be 50 cm in dia., usually with buttresses. The heartwood is white or reddish brown with the sapwood is white, pinkish white, honey-coloured, yellow, or brown with distinct lamination.
11.	Order: Fabales Family: Fabaceae Genus: *Koompassia* Species: *Koompassia malaccensis* Local/Common Name (Vernacular): Kempas Menggris	Terrestrial Swamp forest Up to 600 m elevation	Used for heavy construction because it is heavy hardwood timber. Classified as Least Concern in IUCN Red List. Classified as Not Evaluated in Malaysia Plant Red List. The species can grow up to 60 m tall. This species is protected in Sarawak under Wildlife Protection Bill of 1990. This species is emergent tree in freshwater peat swamp forest and also found in sub-montane forest. It also tends to be an orangish brown with an overall mahogany like appearance. This tree is durable regarding decay resistance and heavy wood timber.
12.	Order: Rosales Family: Cannabaceae Genus: *Gironniera* Species: *Gironniera nervosa* Local/Common Name (Vernacular): Medang Kasap Puloh	Terrestrial Alluvial soils, on hillsides and ridges forest Up to 500 m elevation	Used for timber industry. May last for five years if kept dry. Classified as Not Evaluated in Malaysia Plant Red List. The species usually located commonly in lowland primary and secondary evergreen forest. This evergreen tree can grow up to 40 m tall and the straight bowl can be up to 64 cm in dia., sometimes with buttresses. This tree also produced fruit which is a yellow, orange or red drupe around 5 to 10 mm long and 3 to 6 mm wide, and it is edible.

Based on Table 1, there are 12 timber-related plants species that had been observed and documented during the study. The first species is *Shorea albida* Sym (Figure 3, A). This species can range in size from a sub-canopy tree around 20 m to an emergent tree up to 75 m tall with a vast, diffuse canopy. The straight cylindrical bole can be 50 to 250 cm in diameter as for the larger trees with large buttresses up to 4 m high. This large timber tree is usually found in North-West Borneo which includes Sarawak, Brunei and Kalimantan. The peat-swamp forest habitats where the tree grows are seriously threatened as regeneration is reported to be non-existent. The plant is classified as endangered in the IUCN Red List of Threatened Species (Ashton, 2004; Corlett & Primack, 2005; Ismaili *et al*., 2016; IUCN, 2017; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995).

Shorea rugosa (Figure 3, B) can be found in Borneo and the Philippines. This species grows on hills in lowland mixed dipterocarp forest. This huge tree has elliptical, papery, wavy leaves that measured up to 16 cm long and 7 cm wide. The bark can be almost as smooth as on young trees. However, when the trees ages, furrows develop, which deepens as time passed. The inner bark is light brown to yellowish and has a slightly stringy texture. This species is classified as critically endangered in IUCN Red List of Threatened Species (Abang Bohari, 2015; Ashton, 2004; Corlett & Primack, 2005; IUCN, 2017; Meekiong *et al*., 2012; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995).

Shorea dealbata (Figure 3, C) can grow up to 30 m tall. This species can be found in Malaysia and Indonesia which grows frequently in kerangas forest and sometimes in swampy land. This species is classified as critically endangered in the IUCN Red List of Threatened Species. The species is lightweight hardwood. The heartwood is yellowish-white when freshly cut, but gradually become yellowish-brown and slightly more distinct from the sapwood on exposure (Abang Bohari, 2015; Ashton, 2004; Corlett & Primack, 2005; IUCN, 2017; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995).

Shorea pauciflora is a species found in Sumatra, Peninsular Malaysia and Singapore (Figure 3, D). It is threatened by habitat loss. This is an emergent tree in undisturbed mixed dipterocarp forest at elevation up to 900 m. The species is mostly found on hillsides with clay to sandy soils. This tree can grow up to 60 to 220 cm in dia. with stout buttresses up to 4 m high (Ashton, 2004; Corlett & Primack, 2005; Meekiong *et al*., 2012; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995; Vincent, 2002).

Dipterocarpus baudii has been recorded from Malaysia, Myanmar, Cambodia, Thailand and Vietnam (Figure 3, E). This tree is located commonly in the tropical forest. This emergent tree species grows up to 40 m high with ovate-elliptic leaves up to 250 mm long and has seeds with winged lobes that are 150 to 180 mm long (Abang Bohari, 2015; Ashton, 2004; Corlett & Primack, 2005; Soepadmo & Wong, 1995).

Dipterocarpus sarawakensis is an emergent tree species that can attain heights above 50 m and grows to 1 m in dia. (Figure 3, F). This species has a short petiole with short and dense indumentums, broadly ovate and obtuse leaves with a wavy margin that helps distinguish this species from the other *Dipterocarpus* species. It is commonly found in Peninsular Malaysia and Borneo including Sarawak, Brunei and Kalimantan. This locally abundant tree grows in leached sandy soils on low coastal hills. This species is listed critically endangered in the Malaysia Plant Red List due to habitat degradation (Abang Bohari, 2015; Ashton, 2004; Corlett & Primack, 2005; IUCN, 2017; Meekiong *et al*., 2012; Soepadmo & Wong, 1995; Vincent, 2002).

Dryobalanops beccarii is a large evergreen tree, with a large, globose crown (Figure 3, G). This species is a large emergent tree and can grow up to 65 m tall. This locally abundant tree grows in leached sandy soil and along the stream. It is commonly found in Peninsular Malaysia and Indonesia. It is a hardwood tree. This tree has been classified as endangered in the Malaysia Plant Red List (Ashton, 2004; Demies *et al*., 2008; IUCN, 2017; Pakhriazad & Mohd Hasmadi, 2010; Soepadmo & Wong, 1995; Vincent, 2002).

Drybalanops oblongifolia is native to regions of Southeast Asia and Maritime Southeast Asia (Figure 3, H). It is found in at least one protected area which is KNP but is threatened elsewhere due to habitat loss. It is an emergent tree up to 60 m tall also found in mixed dipterocarp forest on sandy clay soils (Ashton, 2004; Demies *et al*., 2008; Lateef *et al*., 2016; Soepadmo & Wong, 1995; Vincent, 2002).

Endospermum diadenum can only be found in Malaysia, Thailand and Indonesia (Figure 3, I). This tree is a mid-canopy tree with a diffused, domed-shaped crown. It can grow up to 36 m tall. The surface of the bark is smooth, becomes wrinkled to scaly in patches and grey fawn while the inner bark is thick and creamy with orange flecks. This species grows in open places of undisturbed mixed dipterocarp and kerangas forest (Appanah & Turnbull, 1998; Ashton, 2004; Meekiong *et al*., 2012; Soepadmo & Wong, 1995; Vincent, 2002).

Ixonanthes icosandra is an evergreen tree with a large, spreading, moderately dense crown which can grow up to 35 m tall (Figure 3, J). It is commonly found in Southeast Asia which includes Malaysia, Thailand, Indonesia and Philippines. The deeply fluted bole can be 50 cm in diameter, usually with buttresses. The heartwood is white or reddish-brown with the sapwood is white, pinkish-white, honey-coloured, yellow, or brown with distinct lamination (Abang Bohari, 2015; Ashton, 2004; Soepadmo & Wong, 1995; Vincent, 2002).

Koompassia malaccensis can grow up to 60 m tall (Figure 3, K). It is also found in Malaysia, Brunei, Indonesia, Singapore and Thailand. This species is protected in Sarawak under the Wildlife Protection Bill of 1990. The plant is classified as the least concern as well as critically endangered in the IUCN Red List of Threatened Species. This species is an emergent tree in freshwater peat swamp forest and also found in sub-montane forest. It also tends to be an orange-brown with an overall mahogany like appearance. This tree is durable regarding decay resistance and heavy wood timber (Appanah & Turnbull, 1998; Ashton, 2004; Cedergren *et al*., 2002; Demies *et al*., 2008; IUCN, 2017; Roszaini & Salmiah, 2015; Sarawak Government Gazette, 1998; Soepadmo & Wong, 1995; Vincent, 2002).

Finally, the *Gironniera nervosa* can also be found in Peninsular Malaysia, Borneo, Thailand, Sumatra, Moluccas and New Guinea (Figure 3, L). This tree is located commonly in lowland primary and secondary evergreen forest. This evergreen tree can grow up to 40 m tall and the straight bowe can be up to 64 cm in diameter, sometimes with buttresses. This tree also produces fruit which is a yellow, orange or red drupe around 5 to 10 mm long and 3 to 6 mm wide, and it is edible (Abang Bohari, 2015; Ashton, 2004; Soepadmo & Wong, 1995; Vincent, 2002).

CONCLUSION

Along the Waterfall Trail at Kubah National Park, 12 timber-related plants species were observed and documented including *Shorea albida* Sym, *Shorea rugosa*, *Shorea dealbata*, *Shorea pauciflora, Dipterocarpus baudii*, *Dipterocarpus sarawakensis*, *Dryobalanops beccarii*, *Dryobalanops oblongifolia*, *Endospermum diadenum*, *Ixonanthes icosandra*, *Koompassia malaccensis* and *Gironniera nervosa*. The most dominant timber family is Dipterocarpaceae in which *Shorea* is the most dominant genus.

ACKNOWLEDGEMENT

The authors would like to thank the Centre for Pre-University Studies, Universiti Malaysia Sarawak and Sarawak Forestry for the financial support and visiting permit. The authors also thank Biology Unit laboratory staff, Besar Ketol, Khairil Rizadh Razali and Norhayati Ramli for their assistance during the study.

REFERENCES

Abang Bohari, D. N. A. (2015). *Structural Characteristics of a Regrowth Forest Stand*. (Undergraduate thesis). University Malaysia Sarawak, Kota Samarahan, Malaysia.

Abu Bakar, N. A., Radam, A., Samdin, Z., & Yacob, M. R. (2016). Willingness to pay in Kubah National Park and Matang Wildlife Centre: A contingent valuation method. *International Journal of Business and Society*, *17*(1), 131–144.

Appanah, S., & Turnbull, J. W. (Eds). (1998). *A Review of Dipterocarps: Taxonomy, ecology and silviculture*. Bogor, Indonesia: Center for International Forestry Research and Forest Research Institute Malaysia.

Ashton, P. S. (2004). Dipterocarpaceae. In E. Soepadmo, L. G. Saw & R. C. K. Chung (Eds.), *Tree Flora of Sabah and Sarawak* (pp. 63-388). Forest Research Institute Malaysia, Kuala Lumpur, Sabah Forestry Department, Sandakan and Sarawak Forestry Department, Kuching.

Boyce, P. C., & Wong, S. Y. (2008). *The Aroids of the Matang Massif, Sarawak, Malaysian Borneo II: Kubah National Park*. Miami, USA: International Aroid Society.

Brahim, N. (2005). *The Diversity of Zingiberaceae in Kubah National Park (Matang Range), Sarawak.* (Undergraduate thesis). University Malaysia Sarawak, Kota Samarahan, Malaysia.

Cedergren, J., Falck, J., Garcia, A., Goh, F., & Hagner, M. (2002). Structure, Composition and Commercial Characteristics of a Primary Dipterocarp Forest in Sabah, Malaysia. *Journal of Tropical Forest Science*, *14*(3), 304–321.

Corlett, R., & Primack, R. (2005). Dipterocarps: Trees That Dominate the Asian Rain Forest. *Arnoldia*, *63*(3), 3–7.

Das, I., & Charles, J. K. (1993). A Contribution to the Herpetology of Bako National Park, Sarawak, East Malaysia. *Hamadryad*, *18*, 24–27.

Demies, M., Lading, E., & Silang, S. (2008). Biodiversity Conservation in Sarawak. In R. Manurung, Z. C. Abdullah, F. B. Ahmad & C. Kuek (Eds.), *Biodiversity-Biotechnology: Gateway to Discoveries, Sustainable Utilization and Wealth Creation* (pp. 23–32). Kuching, Malaysia: Sarawak Biodiversity Centre.

Dow, R. A., & Reels, G. T. (2010). The Odonata of three National Parks in Sarawak. *Agrion*, *14*(1), 14–19.

Eyssartier, G., Stubbe, D., Walleyn, R., & Verbeken, A. (2009). New records of Cantharellus species (Basidiomycota, Cantharellaceae) from Malaysian dipterocarp rainforest. *Fungal Diversity*, *36*, 57–67.

Hanan, H. (2014). Risk Management Awareness at Bako National Park. *Journal of Tourism, Hospitality & Culinary Arts*, *6*(1), 45–53.

Ismaili, G., Openg, I., Abdul Rahim, K. K., & Duju, A. (2016). Strength group of Aras as fast growing indigenous species timber in Sarawak. *Journal of Scientific Research and Development*, *3*(2), 28–32.

IUCN. (2017). *Malaysian Threatened and Rare Tree Identification and Landscape Guideline*. Petaling Jaya, Malaysia.

Kamri, T. (2013). Willingness to Pay for Conservation of Natural Resources in the Gunung Gading National Park, Sarawak. *Procedia - Social and Behavioral Sciences*, *101*, 506–515.

Kanzaki, M., Yap, S. K., Okauchi, Y., Katsuhiko, K., & Yamakura, T. (2003). Eco-Morphological Grouping of Non-Dipterocarp Tree Species in a Tropical Rain Forest Based on Seed and Fruit Attributes. In M. Kanzaki (Ed.), *Pasoh: Ecology of a Lowland Rain Forest in Southeast Asia* (pp. 123–135). Tokyo, Japan: Springer.

Kerfahi, D., Tripathi, B. M., Lee, J., Edwards, D. P., & Adams, J. M. (2014). The Impact of Selective-Logging and Forest Clearance for Oil Palm on Fungal Communities in Borneo. *PLoS ONE*, *9*(11), 1–8.

Lateef, A. A., Sepiah, M., & Bolhassan, M. H. (2015). Description of Pseudopestalotiopsis kubahensis sp. nov., a new species of microfungi from Kubah National Park, Sarawak, Malaysia. *Current Research in Environmental & Applied Mycology*, *5*(4), 376–381.

Lateef, A. A., Sepiah, M., & Bolhassan, M. H. (2016). Diversity and Distribution of Microfungi from Dipterocarp Forests in Sarawak, Borneo Island (Malaysia). *Malaysian Journal of Science*, *35*(2), 290–303.

Lee, S. S, & Watling, R. (2005). Macrofungal Diversity in Malaysia. In L. S. L. Chua, L. G. Kirton & L. G. Saw (Eds.), *Status of Biological Diversity in Malaysia and Threat Assessment of Plant Species in Malaysia* (pp. 169–180). Kuala Lumpur: Forest Research Institute Malaysia.

Lesley, M. B., Velnetti, L., Fazira, A. A., Kasing, A., Samuel, L., Micky, V., & Awang, A. S. A. H. (2016). Detection and antibiotic susceptibility profiles of *Listeria monocytogenes* in wildlife and water samples in Kubah National Park, Sarawak, Malaysia. *International Food Research Journal*, *23*(1), 360–365.

Lin, C. W., Chung, S. W., & Peng, C. I. (2017). Eleven new species of Begonia (Begoniaceae) from Sarawak, Borneo. *Taiwania*, *62*(3), 219–251.

Mahidin, H., & Sofwan, N. (2012). Tourist's Perception on Conservation of Natural Resources At Bako National Park, Sarawak, Malaysia. In M. E. Wasli, H. Sani, F. Badruddin Ahmad, S. Mohamad, L. Po Teen, L. Kui Soon & M. Sidi (Eds.), *4th Regional Conference on Natural Resources in the Tropics (NTrop4)* (pp. 231–241). Kuching: Universiti Malaysia Sarawak.

Meekiong, K., Latiff, A., Tawan, C. S., & Miraadila, M. I. (2012). Leaf micro-morphological characteristics of selected Vatica species (Dipterocarpaceae) from Kubah National Park, Sarawak. *Malaysian Applied Biology*, *41*(2), 41–44.

MNRE. (2006). *Biodiversity in Malaysia*. *Ministery of Natural Resources and Envirnment*. Kuala Lumpur, Malaysia: Ministry of Natural Resources and Environment.

MNRE. (2015). *National Policy on Biological Diversity*. Putrajaya, Malaysia: Ministry of Natural Resources and Environment.

Mohd Isa, M. (2013). *Leaf Micromorphological Characteristics of Selected Vatica spp. and Comparison with Other Genera in Tribe Dipterocarpeae (Dipterocarpaceae)*. (Undergraduate thesis). University Malaysia Sarawak, Kota Samarahan, Malaysia.

MRPE. (2012). *Legality Verification of Logs in Sarawak*. Kuching, Malaysia: Ministry of Resource Planning and Environment.

MTII. Malaysian Timber Industry Board (Incorporation) Act 1973, Pub. L. No. Act 105, 1 (2011). Malaysia: Laws of Malaysia.

Pakhriazad, H. Z., & Mohd Hasmadi, I. (2010). A Study on Trend of Logs Production and Export in the State of Sarawak, Malaysia. *International Journal of Marketing Studies*, *2*(1), 92–95.

Pearce, K. G. (1992). *The Palms of Kubah National Park, Matang, Kuching Division*. Kuala Lumpur, Malaysia.

Poulsen, A. D., & Leong-Škorničková, J. (2017). Two new Orchidantha species (Lowiaceae) from Borneo. *Blumea: Journal of Plant Taxonomy and Plant Geography*, *62*(2), 157–162.

Roszaini, K., & Salmiah, U. (2015). Resistance of five timber species to marine borer attack. *Journal of Tropical Forest Science*, *27*(3), 400–412.

Roszaini, K., Salmiah, U., Rahim, S., Shahlinney, L., Nor Azried, A. R., Baharudin, K., … Hamidah, A. (2017). Qualitative and Quantitative Determination of Resistance of Twenty Two Malaysian Commercial Timbers through Subterranean Termite Feeding Behavior. *Forestry Research and Engineering: International Journal*, *1*(2), 1–10.

Sarawak Government Gazette. Wild Life Protection Ordinance, 1998, Pub. L. No. Wild Life Protection Ordinance, 1998 (1998). Malaysia: Law of Sarawak.

Sarawak Tourism. (2018). *Frog Tour at Kubah National Park.* Retrieved April, 2016, from https://sarawaktourism.com/blog/frog-tour-at-kubah-national-park.html.

Sasaki, S. (2006). Ecology and Physiology of Dipterocarpaceae. In K. Suzuki, K. Ishii, S. Sakurai & S. Sasaki (Eds.), *Plantation Technology in Tropical Forest Science*. Tokyo: Springer-Verlag.

SCMP. (2017, December 22). Timber Land to Revolutionise Sarawak Retail. *South China Morning Post*, p. S10.

SFD. (2003). *Innovative Tropical Rainforest Conservation.* Kuching: Sarawak Forestry Department.

Soepadmo, E., & Wong, K. M. (Eds). (1995). *Tree Flora of Sabah and Sarawak* (1st ed.). Kepong, Malaysia: Sabah Forestry Department, Forest Research Institute Malaysia, Sarawak Forestry Department, Malaysia.

Teo, S. P., Chai, P. P. K., & Phua, M. H. (2013). Conservation Gap Analysis of Dipterocarp Hotspots in Sarawak Using GIS, Remote Sensing and Herbarium Data. *Sains Malaysiana*, *42*(9), 1237–1246.

Vermeulen, J. J., & Lamb, A. (2011). Endangered even before formally described: Bulbophyllum kubahense n.sp., a beautiful and assumedly narrowly endemic orchid from Borneo. *Plant Systematics and Evolution*, *292*, 51–53.

Vincent, A. (2002). *Studies of native and exotic tree plantations in Sarawak*. (Master thesis). University Malaysia Sarawak, Kota Samarahan, Malaysia.

Wahab, M. P. (2012). *Pitcher Plant (Nepenthes ampullaria) Choices by Frogs of the Microhyla nepenthicola and M. borneensis Complex for Breeding at Kunah National Park, Sarawak*. (Undergraduate thesis). University Malaysia Sarawak, Kota Samarahan, Malaysia.

Wong, W. S. (1995). *Timber Structures in Malaysian Architecture and Buildings*. (Master thesis). University of Tasmania at Launceston, Australia.

Yee, S. O., & Chin, L. F. (1989). Bako National Park-A natural laboratory for teaching and learning about the environment. *Teaching and Learning*, *9*(2), 82-101.

Trends in Undergraduate Research (2019) 2(2): a13-19
https://doi.org/10.33736/tur.1422.2019

Prevalence of *Escherichia coli* O157:H7 in Drink from Different Food Premises in Kota Samarahan Sarawak

NILLIAN ELEXSON*, AMIZA NUR, DIYANA NUR, ZAKIRAH AMIRAH & BEBEY GRACE

Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300, Kota Samarahan, Sarawak, Malaysia.
*Corresponding author: nelexson@unimas.my

ABSTRACT

Contamination of drinks with *Escherichia. coli* O157:H7 served in food premises such as restaurants can cause haemorrhagic colitis and haemolytic uremic syndrome to humans. The presence or absence of faecal pathogen was demonstrated using the coliform group as indicator microorganisms. Therefore, this study was conducted to detect the presence of *E. coli* O157:H7 in drinking water from food restaurant premises in Kota Samarahan and Kuching to ensure safe and potable drinking water is served to the consumer. A total of thirty (n=30) drink samples including six types of each of the samples are cold plain water, iced tea, iced milo, syrup and iced milk tea. Most Probable Number (MPN) procedure was used in this study to enumerate the MPN values of coliform bacteria in each drink collected. A total of 53.33% (16/30) of the drink samples showed positive *E. coli* detection. Then, the PCR assay showed 6.25% (one out of 16 isolates) samples were positive and carried *stx*1 gene produced by *E. coli* O157:H7 in iced milo sample types. This study showed the drinks collected from food premises were contaminated with faecal contamination, which was not safe to drink by the consumer. Therefore, preventive actions should be taken to prevent foodborne illness outbreak in the future.

Keywords: *E. coli*, *E. coli* O157:H7, food premises, MPN method, PCR assay

INTRODUCTION

Consumption of contaminated drinking water might be linked to severe life-threatening disease also known as waterborne diseases. Foodborne and waterborne diseases always become a major problem in developing countries compared to developed countries (Portier *et al*., 2013). According to the World Health Organization (WHO, 2002), contaminated drinking water may become a major burden to human health as it can contribute to waterborne disease. There are several types of waterborne outbreak agents that can survive in water including drinking water which can harm human health such as bacterial agents, viral agents, pathogenic agents and chemical agents (Leclerc, Schwartzbrod & Dei-Cas, 2002). This pathogenic agent can also be transmitted directly from dermal contact with water. Bacterial pathogens that are well known in causing waterborne diseases such as diarrheal illness, typhoid fever, cholera, cryptosporidium, giardiasis and legionellosis include *Escherichia coli*, *Salmonella typhi*, *Vibrio cholera*, *Cryptosporodium parvum, Giardia lamblia* and *Legionella pneumophila*. It has been reported that between 2011 and 2016, there are 381 cases associated with the protozoan disease worldwide (Efstratiou, Ongerth & Karanis, 2017). 7% from 6939 deaths in the United States were associated with several types of pathogens including *E. coli* transmitted by the faecal-oral route (Gargano *et al*., 2017). *E. coli* is a member of the faecal coliform group and commonly used as an indicator to monitor the microbiological quality in drinking. Their presence in drinking water indicates recent faecal contamination which means the potential presence of enteric pathogens (Figueras & Borrego, 2010). Hygiene standards of food premises may reflect the quality of food and drinks served to the customers. Restaurants and stalls with low sanitary conditions may contribute to the foodborne and waterborne outbreak and may affect many people. Golan *et al*. (2004) reported that 700 people had contracted a foodborne illness and four children died after eating contaminated meat from a restaurant, while five diners experienced diarrhoea and vomiting after eating raw oysters (Baker *et al*., 2010). As water is also used in food preparation and making drinks for the customer, it is crucial to abide by the guidelines of drinking water quality standards established by the Ministry of Health Malaysia. The maximum acceptable value for total coliform and *E. coli* is zero in 100 ml (WHO, 1997). This study is essential in terms of gauging the level of bacteria present in drinking water with *E. coli* as well as determining the quality and safety of the drinking water served from restaurants in Kota Samarahan. The risks of water-related diseases and health problems in local residents might increase by the presence of some pathogenic bacteria in drinking water. Therefore, the objective of this research is to investigate the presence of *stx*1 gene of isolated *E. coli* O157:H7 from drinks in food premises in Kota Samarahan, Sarawak.

MATERIALS & METHODS

Samples Collection

A total of 30 drinks with five flavours (iced tea, iced milk tea, syrup, iced milo and plain cold water) of each restaurant were purchased from six food premises in Kota Samarahan. Then, the samples were labelled respectively and stored at 4^0C for 2-4 hours and transported to the Microbiology Laboratory 2, Universiti Malaysia Sarawak.

Detection of Faecal Coliform Bacteria

The most probable number (MPN) procedure is used to estimate microbial densities in many matrices including foods and water. Most Probable Number (MPN) was used referring to the protocol provided by Feng *et al*. (2013). For the presumptive test, Lauryl Tryptose Broth (LTB) fermentation tubes were used to screen the presence of coliform organisms by observing the gas production within Durham tube. Nine test tube containing sterile LTB broth incorporated with Durham tube in each tube, 3 tubes double strength LTB with 10 mL portions of drinking water samples, 3 tubes single strength LTB with 1 mL portions of drinking water samples and 3 tubes single strength of LTB with 0.1 mL portions of drinking water were inoculated and incubated at 37 ^{0}C. Growth and gas production were observed after 48 hours. A loopful from each of the positive tubes of LTB was streaked onto Eosin Methylene Blue (EMB) agar and incubated for 24 hours at 37°C. Coliforms bacteria produced small colonies with dark centre while *E. coli* produced shiny metallic green colonies on the EMB agar as shown in Figure 2. The shiny metallic green colonies were isolated and stroke on the nutrient agar and incubated for 24 hours at 37^{0}C.

Gram Staining

Gram staining was carried out based on the protocol provided by Barile (2012). A single colony of bacteria from each agar plate was transferred onto a microscope slide. A drop of distilled water was added to the slide and dried up. Crystal violet was added on the slide, left for one minute and washed with distilled water. The specimen was covered with iodine solution for one minute and washed again with distilled water. Then, 90% of ethanol was run on the microscope slide until the blue staining colour no longer came off and rinsed with distilled water immediately then flooded with safranin for 45 seconds. The slide was rinsed with distilled water and dry then observed using a microscope. The result is shown in Figure 3.

Genomic DNA Extraction (Boiling Cell Method)

Extraction of DNA using the boiled-cell method provided by Chai *et al*. (2007) with minor modification was used to detect the *E. coli* O157:H7 in the water sample. A single colony from nutrient agar was transferred into the Luria Bertani (LB) broth and incubated for 24 hours at 37 ^{0}C. For each positive broth, 1.5 mL was centrifuged at 10000 rpm for five minutes to get the pellet then suspend the pellet by adding 500 µL of sterile distilled water and vortex vigorously to dissolve the pellet. Boiled for 10 minutes at 97^{0}C and allowed to cool at -20^{0}C for 5 minutes. Centrifuged the tube at 10000 rpm for 10 minutes and the supernatant was used as DNA template for PCR assay.

Polymerase Chain Reaction (PCR)

In this study, *E. coli* O157:H7 ATCC 43895 is a positive control and the PCR assay was carried out according to Chai *et al*. (2007). A total of 25 µL reaction mixtures were used to perform PCR amplification as shown in Table 1. The primers and PCR conditions as outlined in Table 2 according to Meng, Zhao, Doyle, Mitchell and Kresovich (1997) were utilized to amplify the *stx*1 gene in *E. coli* O157:H7.

Table 1. Master Mix component for PCR.

Component	1x reaction (µL)
5x Green Go Taq® Flexi Buffer	5.0
25 mM Magnesium Chloride ($MgCl_2$)	2.5
10 mM Deoxyribonucleotide oligophosphates (dNTPs)	1.25
Forward Primer	1.0
Reverse Primer	1.0
DNA template	5.0
Sterile distilled water	8.85
Go Taq® Flexi DNA Polymerase (5u/µl)	0.4
Total Volume	25.0

Table 2. Characteristic of primers to detect *stx*1 gene *E. coli* O157:H7

Primer	Nucleotide sequence	Target gene	Amplicon size (bp)
Slt1 (Forward Primer)	TGT AAC TGG AAA GGT GGA GTA TAC	*Stx*1 gene	210
Slt1 (Reverse Primer)	GCT ATT CTG AGT CAA AAA ATA AC		

Agarose Gel Electrophoresis (AGE)

A total of 5 µL of PCR products was loaded into the agar and undergo electrophoresis for 47 minutes at 90V. The PCR products were viewed using UV transilluminator after staining with ethidium bromide (EtBr). The result for PCR products for the detection of *stx* gene in *E. coli* O157:H7 with expected size 210 bp from different restaurants on gel electrophoresis was shown in Figure 4.

RESULTS

In the presumptive test, the three-tube analyses of the MPN method were used to analyse *E. coli* in drink samples collected from six restaurants in Kota Samarahan. Lauryl Tyrptose broth was used to enrich the faecal coliform in the water samples. The broth turned cloudy and the presence of gas in Durham tubes indicated that there was faecal coliform grown in water samples. MPN value was recorded based on the gas produced in each positive test tube. The averages of the MPN values between the drinks were shown in Figure 1. Drink sample A (iced milo) had the highest mean followed by drink sample E (iced milk tea), drink samples C (iced tea), drink sample B (syrup) and drink sample D (cold plain water).

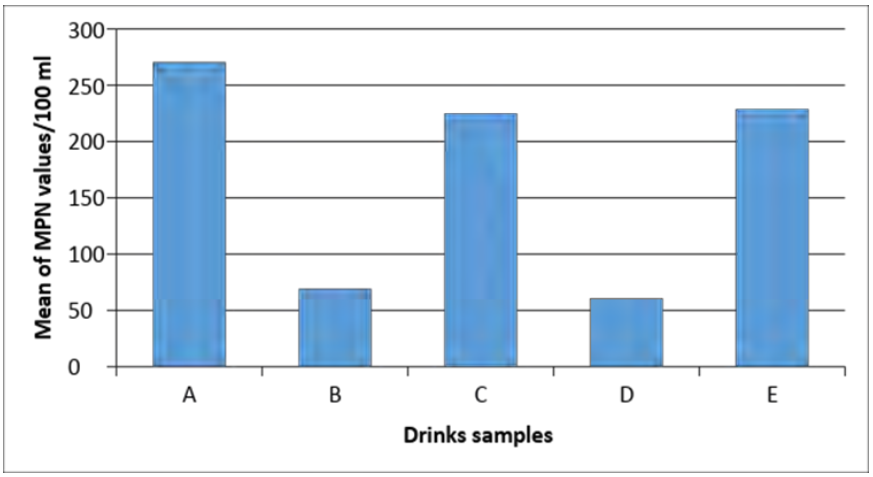

Figure 1. The average of MPN values/100 ml of drink samples recorded six restaurants based on samples types. Sample A = iced milo, Sample E = iced milk tea, Samples C = iced tea, Sample B = syrup and Sample D = cold plain water.

All 16 isolates of *E. coli* that produced green metallic sheen with dark centre colonies on EMB agar, as shown in Figure 2, after subjected for Gram staining. EMB is a selective medium for gram-negative bacteria where the combination of dyes eosin and methylene blue inhibits Gram-positive bacteria to grow on the agar. *E. coli* is able to ferment lactose rapidly and produced acid end-products. The microscopic examination results on the 16 isolates showed that 100% of the isolates were Gram-negative, rod-shaped and non-spore forming bacteria.

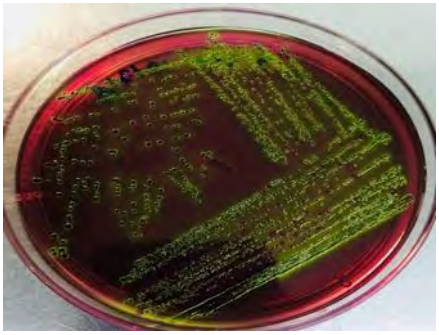

Figure 2. Green metallic sheen colonies produced on EMB agar.

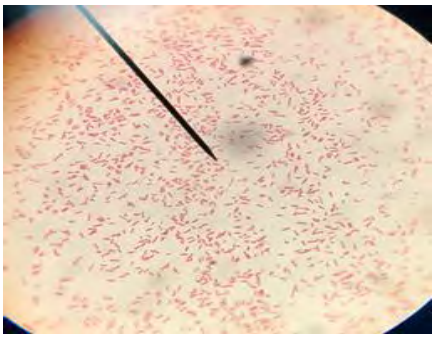

Figure 3. Gram staining result shows Gram-negative, rod-shaped and non-spore forming bacteria under Oil Immersion lens 100x.

A total of 16 isolates from drinks water samples that were identified as *E. coli* through gram staining and morphology identification were subjected to detect the *stx*1 gene using PCR assay. *E. coli* O157:H7 ATCC 43895 was used as a positive control to validate the PCR assay and to exclude the false-negative result while sterile distilled water was used as a negative control. As a result, as shown in Figure 4, there was only one (6.25%) out of 16 isolates that were positives to carry *stx*1 gene.

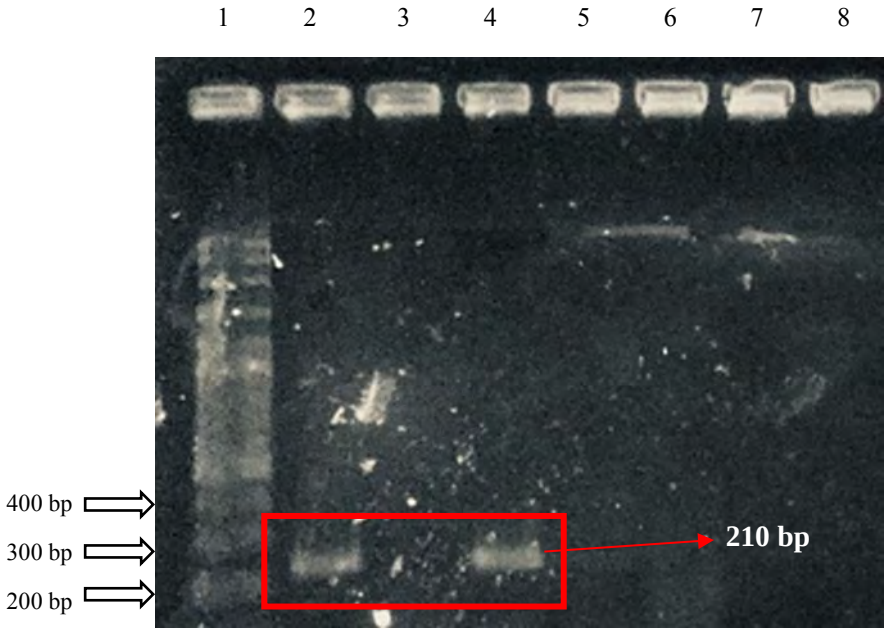

Figure 4. Lane 1: 1 kb DNA ladder (Vivantis), Lane 2: Positive control, *E. coli* O157:H7 ATCC 43895, Lane 3: Negative control, Lane 4: Ice milo drink, Lane 5: Cold plain water, Lane 6: Iced tea, Lane 7: Syrup, Lane 8: Iced milk tea.

DISCUSSION

This study was designed to determine the detection of *E. coli* O157:H7 in drinks sold from six different restaurants in Kota Samarahan. The main aim was to investigate the level of personal and environmental hygiene practices in restaurants. MPN culture method was used for enumeration of total coliform bacteria in samples collected from six restaurants in Kota Samarahan as it is commonly used by microbiologists in analysing water quality, food microbiology and public health. Drink samples from all six restaurants showed positive results in the presumptive test where it initially indicated the presence of coliform bacteria. The quality of drinking water was tested by detecting the presence of coliform bacteria because it is expensive and time-consuming to detect different pathogens as their presence is usually very low in number (Rompré, Servais, Baudart, De-Roubin and Laurent, 2002). Therefore, in this study MPN method is used to detect one of the coliform bacteria, *E. coli* which is also known as indicator microorganisms for the presence of faecal contamination in the samples. In this study, Lauryl Sulfate Tryptose broth was used as the production of gas could serve not only as a presumptive test but also a confirmation test (Gerba, 2009). Samples from CS9 showed the highest 1100 MPN value in iced tea, syrup and cold plain water which implies that drinks sold in the restaurants are generally unhygienic. The presence of coliform bacteria can be the risk factor for the transmission of disease causative agents such as *Salmonella*, *Shigella*

and other enteropathogenic *E. coli* (Newell *et al*., 2010) which are not specifically investigated in this study. The MPN values of coliform bacteria must be absence in drinking water and water that are used to prepare food which is rated excellent and potable whereas MPN value more than 50 indicates that there are pollution and contamination in water and considered as unfit for human consumption (Uyttendaele *et al*., 2015). In the completed test, all the non-potable samples were recorded based on the production of greenish metallic sheen colonies on selective media, EMB agar. The EMB agar is used to select only gram-negative bacteria to grow and to differentiate between faecal coliforms and non-faecal coliform bacteria. All the positive colonies on EMB agar were observed under the oil immersion lens and 100% of the samples were confirmed as *E. coli*. *E. coli* has an additional outer membrane which is made of phospholipids and lipopolysaccharides. (Tokuda & Matsumaya, 2004). The cell wall of the bacteria consists of negative charge due to the presence of lipopolysaccharides on the outer membrane and because of that, *E. coli* does not retain crystal violet during Gram staining.

In the detection of *stx*1 gene, the expected size of the PCR product is 210 bp as shown in Figure 4. This indicates the presence of *stx*1 gene in the samples collected. This desired gene was amplified as it is suitable for sample evaluation especially during the *E. coli* O157:H7 outbreak in food or water. From the result, it was found that out of 30 drink samples, 3.33% were positive for *stx*1 gene which is isolated from iced milo drinks. It means that the drinks contained pathogenic *E. coli* O157:H7 strain which is an indicator for faecal contamination. The massive outbreak related to *E. coli* was during 2011 in Europe, which caused by the strain of Shiga toxin-producing *E. coli* O104:H4 and the strain also frequently referred to as EHEC. Other than that, *E. coli* O157:H7 strain known to be linked with waterborne outbreaks and causing high morbidity and mortality worldwide. Besides, the pathogen can survive for long periods of time in the water mostly in cold temperatures (Wang & Doyle, 1998). The contamination of drinks served in restaurants may also come from the ice which is used to cool drinks and refrigerate foods such as fish and seafood.

Outbreaks of gastroenteritis due to contaminated ice have been reported (Khan *et al*., 1994; Pedalino *et al*., 2003) in other parts of the world. Trabulsi, Keller and Gomes (2002) concluded that typical enteropathogenic *E. coli* strain is the main cause of infantile diarrhoea in developing countries but rarely happen in industrialized countries. Other health disorders caused by *E. coli* are urinary tract infections, pulmonary infections, abscesses and skin-wound infections. Moreover, those who were infected by the pathogen might be going through the process in which the incubation period takes three to four days after exposure or worst, short as one day. The incubation might as well take as long as 10 days' time. The early symptoms often start with mild belly pain or non-bloody diarrhoea that keep worsens after several days. Elexson *et al*. (2017) stated that exposures that result in illness include the intake of food unhygienically prepared, drinking unpasteurized (raw) milk, consuming disinfected water, having contact with bovine, or contact with the excrement of infected people. Some foods are considered to carry such a high risk of infection with *E. coli* 0157including raw milk, unpasteurized apple cider, and soft cheeses made from raw dairies. Sometimes, the contact is apparent (working with bovines at a dairy or changing diapers, for example), but sometimes it is not (like the consumption of an undercooked hamburger or a contaminated piece of lettuce) (Centre for Disease Control and Prevention [CDC], 2018). Poor water quality, improper food handling practices and poor hygienic conditions of places of food premises can be a major factor in the contamination of faecal coliform to the food and drinks they served the customer (Pandey, Kass, Soupir, Biswas and Singh, 2014). Other than that, ice can also be one of the sources of contamination of pathogenic bacteria because it is directly in contact with drinks and indirectly by refrigerating foods in the freezer (Ackers *et al*., 1998). The connection between contaminated ice and enteric diseases and the connection of infections and contaminated ice used in hospitals were mainly reported regarding the issue (Banatvala *et al*., 2008).

CONCLUSION

The study has confirmed that drinks from restaurants in Kota Samarahan were contaminated with total coliform bacteria. *E. coli* O157:H7 was detected in iced milo drinks collected from one of the food premises in Kota Samarahan. Hence, strict public-health regulations regarding the sale of foods and water by restaurants, regular laboratory checking for microbiological quality of served food and drinks to the consumers must be performed by local authorities to monitor the supply of potable water, good hygiene practices, food handling and environmental sanitation in food premises in Kota Samarahan to reduce the morbidity due to gastroenteritis in the study population.

ACKNOWLEDGEMENT

Research fund was sponsored by RAGS: E14099/F07/1314/2015(08) Universiti Malaysia Sarawak (UNIMAS) Kota Samarahan Sarawak, Malaysia.

REFERENCES

Ackers, M. L., Mahon, B. E., Leahy, E., Goode, B., Damrow, T., Hayes, P. S., & Griffin, P. M. (1998). An outbreak of *Escherichia coli* O157: H7 infections associated with leaf lettuce consumption. *Journal of Infectious Diseases*, *177*(6), 1588-1593.

Baker, K., Morris, J., McCarthy, N., Saldana, L., Lowther, J., Collinson, A., & Young, M. (2010). An outbreak of norovirus infection linked to oyster consumption at a UK restaurant, February 2010. *Journal of Public Health*, *33*(2), 205-211.

Banatvala, N., Griffin, P. M., Greene, K. D., Barrett, T. J., Bibb, W. F., Green, J. H., & Wells, J. G. (2001). The United States national prospective hemolytic uremic syndrome study: microbiologic, serologic, clinical, and epidemiologic findings. *The Journal of Infectious Diseases*, *183*(7), 1063-1070.

Barile, M. F. (2012). Gram staining technique. *Methods in Mycoplasmology V1: Mycoplasma Characterization*, 1, 39.

Chai, L. C., Robin, T., Ragavan, U. M., Gunsalam, J. W., Bakar, F. A., Ghazali, F. M., Radu, S., & Kumar, M. P. (2007). Thermophilic *Campylobacter* spp. in salad vegetables in Malaysia. *International Journal of Food Microbiology*, *117*(1), 106-111.

Centre for Disease Control and Prevention, CDC. (2018). Multistate Outbreak of *E. coli* O157:H7 Infections Linked to Romaine Lettuce. Retrieved May 2, 2018, from https://www.cdc.gov/ecoli/2018/o157h7-04-18/index.html.

Efstratiou, A., Ongerth, J. E., & Karanis, P. (2017). Waterborne transmission of protozoan parasites: review of worldwide outbreaks-an update 2011–2016. *Water research*, *114*, 14-22.

Elexson, N., Nik Yuhanis, F. N., Malcolm, T. T. H., New, C. Y., Chang, W. S., Ubong, A., Kuan, C. H., Loo, Y. Y., Thung, T. Y., & Son, R. (2017). Occurrence of *Escherichia coli* harbouring *stx* genes in popiah, a Malaysian street food. *Food Research*, *1*(1), 29-32.

Feng, P., Weagant, S. D., Grant, M. A., Burkhardt, W., Shellfish, M., & Water, B. (2002). BAM: Enumeration of *Escherichia coli* and the Coliform Bacteria. *Bacteriological Analytical Manual*, 13-19.

Figueras, M., & Borrego, J. J. (2010). New perspectives in monitoring drinking water microbial quality. *International Journal of Environmental Research and Public Health*, *7*(12), 4179-4202.

Golan, E. H., Roberts, T., Salay, E., Caswell, J. A., Ollinger, M., & Moore, D. L. (2004). Food safety innovation in the United States: evidence from the meat industry (No. 34083). *United States Department of Agriculture, Economic Research Service*.

Gerba, C. P. (2009). Indicator microorganisms. In *Environmental Microbiology* (2nd ed.). (pp. 485-499). Academic Press.

Gargano, J. W., Adam, E. A., Collier, S. A., Fullerton, K. E., Feinman, S. J., & Beach, M. J. (2017). Mortality from selected diseases that can be transmitted by water–United States, 2003–2009. *Journal of Water and Health*, *15*(3), 438-450.

Khan, A. S., Moe, C. L., Glass, R. I., Monroe, S. S., Estes, M. K., Chapman, L. E., & Iskander, J. K. (1994). Norwalk virus-associated gastroenteritis traced to ice consumption aboard a cruise ship in Hawaii: comparison and application of molecular method-based assays. *Journal of Clinical Microbiology*, *32*(2), 318-322.

Leclerc, H., Schwartzbrod, L., & Dei-Cas, E. (2002). Microbial agents associated with waterborne diseases. *Critical reviews in microbiology*, *28*(4), 371-409.

Meng, J., Zhao, S., Doyle, M. P., Mitchell, S. E., & Kresovich, S. (1997). A multiplex PCR for identifying Shiga-like toxin-producing *Escherichia coli* O157: H7. *Letters in Applied Microbiology*, *24*(3), 172-176.

Newell, D. G., Koopmans, M., Verhoef, L., Duizer, E., Aidara-Kane, A., Sprong, H., ... & van der Giessen, J. (2010). Food-borne diseases—the challenges of 20 years ago still persist while new ones continue to emerge. *International Journal of Food Microbiology*, *139*, S3-S15.

Pedalino, B., Feely, E., McKeown, P., Foley, B., Smyth, B., & Moren, A. (2003). An outbreak of Norwalk-like viral gastroenteritis in holidaymakers travelling to Andorra, January-February 2002. *Euro surveillance: bulletin européen sur les maladies transmissibles= European communicable disease bulletin*.

Portier, C. J., Tart, K. T., Carter, S. R., Dilworth, C. H., Grambsch, A. E., Gohlke, J., & Maslak, T. (2013). A human health perspective on climate change: a report outlining the research needs on the human health effects of climate change. *Journal of Current Issues in Globalization*, *6*(4), 621.

Pandey, P. K., Kass, P. H., Soupir, M. L., Biswas, S., & Singh, V. P. (2014). Contamination of water resources by pathogenic bacteria. *AMB Express*, *4*, 51.

Rompré, A., Servais, P., Baudart, J., De-Roubin, M. R., & Laurent, P. (2002). Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. *Journal of Microbiological Methods*, *49*(1), 31-54.

Tokuda, H., & Matsuyama, S. I. (2004). Sorting of lipoproteins to the outer membrane in *E. coli*. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1693*(1), 5-13.

Trabulsi, L. R., Keller, R., & Gomes, T. A. T. (2002). Typical and Atypical Enteropathogenic *Escherichia coli*. *Emerging infectious diseases*, *8*(5), 508-513.

Uyttendaele, M., Jaykus, L. A., Amoah, P., Chiodini, A., Cunliffe, D., Jacxsens, L., Holvoet, K., Korsten, L., Lau, M., McClure, P., Medema, G., Sampers, I., & Jasti, P. R. (2015). Microbial hazards in irrigation water: Standards, norms, and testing to manage use of water in fresh produce primary production. *Comprehensive Reviews in Food Science and Food Safety*, *14*(4), 336-356.

Wang, G., & Doyle, M. P. (1998). Survival of enterohemorrhagic *Escherichia coli* O157: H7 in water. *Journal of food Protection*, *61*(6), 662-667.

World Health Organization, WHO. (1997). Guidelines for drinking-water quality. Volume 3, Surveillance and control of community supplies. *World Health Organization*.

World Health Organization, WHO. (2002). The world health report 2002: reducing risks, promoting healthy life. *World Health Organization*.

Trends in Undergraduate Research (2019) 2(2): c1-7
https://doi.org/10.33736/tur.1981.2019

3D Face Recognition Analysis Using Random Forest

NURUL ATIFAH ROSLAN, HAMIMAH UJIR* & IRWANDI HIPNI MOHAMAD HIPINY

Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
*Corresponding author: uhamimah@unimas.my

ABSTRACT

Face recognition is an emerging field due to the technological advances in camera hardware and for its application in various fields such as the commercial and security sector. Although the existing works in 3D face recognition perform well, a similar experiment setting across classifiers is hard to find, which includes the Random Forest classifier. The aggregations of the classification from each decision tree are the outcome of Random Forest. This paper presents 3D facial recognition using the Random Forest method using the BU-3DFE database, which consists of basic facial expressions. The work using other classifiers such as Neural Network (NN) and Support Vector Machine (SVM) using a similar experiment setting also presented. As for the results, the Random Forest approach has yield 94.71% of recognition rate, which is an encouraging result compared to NN and SVM. In addition, the experiment also yields that fear expression is unique to each human due to a high confidence rate (82%) of subjects with fear expression. Therefore, a lower chance to be mistakenly recognized someone with a fear expression.

Keywords: 3D face recognition, Neural Network, Random Forest, Support Vector Machine

INTRODUCTION

3D face recognition is the enhancement of face recognition technology in a few aspects. It has the possibility to overcome feature localization, pose and illumination problems. 3D face recognition is known for its easy and non-intrusive acquisition of face information compared to the other biometrics. Moreover, 3D face recognition has the low privacy of information compared to signatures and fingerprints (Gökberk *et al*., 2008). In recent years, 3D face recognition system has become a popular biometric system due to its higher capability and accuracy, and for its application in various fields such as in commercial, security, health and banking sector.

Existing works in 3D face recognition have been using various kinds of classifiers, for example, Neural Networks (NN) (Hu *et al*., 2013; Kim *et al*., 2017), Support Vector Machine (SVM) (Mousavi *et al*., 2008), Principal Component Analysis (PCA) (Lee & Han, 2006) and Linear Discriminate Analysis (Drira *et al*., 2012; Hiremath, 2013). These classifiers have a few issues when dealing with high dimensional data and facing an overfitting problem. Works using Random Forest also have been discussed in Salhi *et al*. (2012) and, Kremic and Subasi, (2016). However, the comparison with other classifiers using a similar experimental setting has not been presented in any work.

Our objectives are to develop 3D facial recognition using Random Forest and to compare Random Forest with two existing methods, which are Probabilistic Neural Network (PNN) and SVM using a similar experiment setting. In this paper, the analyses of the 3D face recognition experiments conducted are presented. Section 2 describes the existing-related Random Forest works in this study and followed by an explanation of our work in section 3. Section 4 discusses the results and analysis of the experiments. Finally, the conclusions and future works are drawn.

RELATED WORKS

Random Forest

Random forest is one of the most advanced ensembles learning algorithms available, and it runs efficiently on large databases (Kulkarni & Sinha, 2013). In addition, Random Forest is also an effective method to estimate the missing data and capable of maintaining accuracy when a huge proportion of data is missing (Breiman, 2001). Random Forest besides able to generate forest, which can be used on additional data in the future. Despite having those positive traits, Random Forest has its own drawbacks where it can be over fitted for certain data sets, which include data sets used in classification and regression tasks. Plus, data with categorical variables as well as various numbers at a level will make the random forest become more biased towards the attributes which have more levels.

Kremic and Subasi (2016) has proposed a 2D facial recognition with Random Forest method using International Burch University (IBU) database. The dataset consists of face images, training data for training classifier and test data for the evaluation of the final methods. They compared two methods, which are SVM and Random Forest. It begins with reading the image from a database, perform skin colour detection, then convert RGB to greyscale, optimize the histogram value of the methods and used it to classify and retrieve the accuracy of the images. During the testing phase, RF has built 30 trees where each of the trees has considered nine random features. As the results, Random Forest obtained 97.17% inaccuracy. The number of cross-validation folds used in Random Forest evaluation is k = 10. The F-measure and receiver operating characteristic (ROC) area of Random Forest is 97% and 0.998, respectively; which are better than SVM with F-measure of 95% and ROC of 0.95. Therefore, it shows that Random Forest has good performance and gave high accuracy as well as had less time taken to build the model compared to the SVM approach.

Bayramoglu *et al*. (2013) have proven that Random Forest performed better in 3D Facial Action Unit detection by combining the person independent geometric features and descriptor based on Local Binary Pattern (LBP) approach. They proposed the combining descriptors for detecting 3D facial action unit (AU) by using the Random Forest as the classifier. The reason they used person independent geometric feature is that they want to overcome the diversity between different persons and different in ages by distance-based features. Therefore, they proposed in using ratios of distances and areas as well as angles on a single 3D face data. Besides, the Bosphorus database is used in their experiment where they tested their 3D AU detection methods. This database consists of 105 subjects with 24 facial AUs which result in 4666 face scans. Plus, it also has six basic expressions (anger, disgust, fear, happy, sad and surprised), occlusions, fixed rotation and image's intensity. A Gaussian filter is used to smooth the raw 3D data that contains spikes and noises. The performance of the 3D facial AU detection using Random Forest is evaluated by obtaining the mean percentage of ROC of 97.7%.

On the other hand, Drira *et al*. (2012) propose 3D Dynamic expressions' recognition based on novel deformation vector field and Random Forest. In this work, they used BU-4DFE database and Deformation Vector Field (DVF) approach where it is basically a Riemannian facial shape analysis dynamic information of the whole face where the resulting temporary vector field will be used to create a feature vector for recognition of expression from 3D dynamic faces. Besides, a multi-class Random Forest is used with LDA-based feature space transformation to achieve the average recognition rate. The experiment was conducted with 60 subjects, including six basic expressions. Based on this experiment, they obtained 93.21% accuracy for the average recognition rate where it has 19% difference from Wang *et al*. (2006) where they only achieved 73.61% with their proposed method.

An iterative Multi-Output Random Forest (iMORF) algorithm for analysis of the unified face which estimating the head poses, facial expression and landmark position is proposed by Zhao *et al*. (2014). Three databases were used, which were 300-W, Bosphorus and CK+. For 300-W, it is used to automatically detect facial landmarks where 6193 images re-annotated with 68 landmark points and three basic expressions (neutral, happy and others), while Bosphorus is used for face image processing with 105 subjects and 4666 faces with various expressions as well as a variety of poses. The CK+ database was used for research in automatic recognition of expressions where it has eight facial expressions (neutral, anger, contempt, disgust, fear, happiness, sadness and surprised) with 68 landmark points. The performance is evaluated by obtaining an accuracy of a head pose estimation and expression recognition. For head

pose estimation using 300-W database, their method achieved the highest accuracy with 86.40%. Meanwhile, for the accuracy of expression recognition, the method yet achieved another highest percentage compared to the other method with 90.04%.

Probabilistic Neural Network

Generally, Probabilistic Neural Network (PNN) is a multilayer feedforward network which consists of four layers, which are input, hidden, summation and output layers. The reason why PNN is chosen is that it is suitable and often used for classification and pattern recognition problems. The process flow is close with the basic NN as the input in the first layer will compute the distance from input vector to training input vector which then produces a vector that the elements will indicate how close the input with training input. For the second layer, the contribution for each class of inputs will be summed together and produce output as a vector of probabilities. Finally, the maximum of probabilities is picked from the complete for transfer function on the output of the second layer which results in producing 1 for positive identification for the class and 0 for negative identification for the non-targeted class (Rutkowski, 2004).

Support Vector Machine

Support Vector Machine (SVM) can be defined as a dimensional hyperplane that separates a set of positive examples from a set of negative examples with maximum margin. Tang and Huang (2008) proposed a 3D 25 facial expression recognition based on the properties of line segments connecting with facial feature points. In their work, they used multi-class SVM and perform expression recognition for person and gender independent using the properties of line segments that connect with certain 3D facial feature points. About 96-dimensional features comprised of the normalized distances and slopes of the line segments are used to recognize the six basic expressions which are anger, disgust, fear, happiness, sadness, and surprised. The result of the recognition rate obtained from this work is 87.1% with the highest average recognition rate which is 99.2% for recognition of surprise expression. The results in this work have outperformed the other related work which they compared.

OUR PROPOSED METHOD

One of the standard phases in the face processing field is to perform facial feature extraction. Between facial features used in face processing are Local Binary Pattern (Bayramoglu *et al*., 2013; Sandbach *et al*., 2012), surface normal (Ujir, 2014), facial feature distances (Soyel & Demirel (2008); Ying *et al*., 2017) et cetera. Certain fiducial points are used as the reference for the facial feature in the recognition process. Facial feature distance is the characteristic of distances extracted from the facial feature points. Facial feature distance can be obtained by various types of distance metric measurements such as Euclidean distance, Geodesic, and others. In our work, 83 facial feature points are used to calculate the facial feature Euclidean distance in the facial extraction phase, refer to Figure 1. The Euclidean distance equation is as in equation 1.

The fundamental element of the Random Forest is to build a small decision tree with few other features where it can lead to a computationally cheap process. In addition, the main principle of Random Forest is a few groups of weak trees (learner) are combined to form a strong learner by obtaining the prediction data and result through averaging of all reached terminal nodes which are known as regression or taking the majority vote where it is based on categorical variables, which are called as classification. Furthermore, in a decision tree of the random forest, the input data will be entered from the top and while it is transverse down the tree, the data will be bucketed into smaller sets.

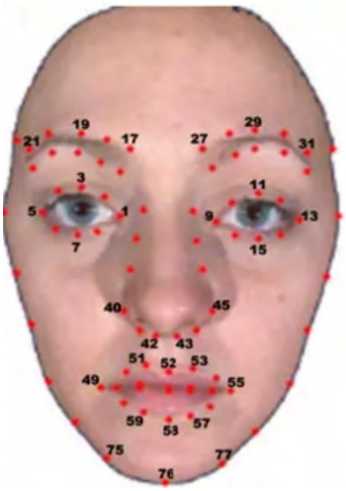

Figure 1: 83 facial feature points on BU-3DFE data (Yin *et al*. 2006)

$$d(x,y,z) = \sqrt{|x_1 - x_2|^2 + |y_1 - y_2|^2 + |z_1 - z|^2} \quad \text{(Equation 1)}$$

The framework of the Random Forest used is as illustrated in Figure 2. According to Radenkovic (2015), in order to create a training model for the random forest, there are many decision trees required. Therefore, based on the figure above, for each of the decision trees, the number of features and N samples are chosen. Then, m will be extracted from the M features and should be much less than. A training set for this tree will be randomly chosen by using the estimated prediction error of the tree. Later, to choose the split feature, we will choose only among the features, and the best split is calculated based on the in the training set. Now, the decision tree will be created until its fully grown, and the grown tree is split as a maximum as it can, which means until the stopping criterion is achieved and no pruning of the tree. Therefore, when many trees have grown, it will create a random forest model where it is a combination of all the decision trees.

During the training phase for the Random Forest method, the data was divided into two parts, which were training and validation data set. In order to partition the data; hold-out validation was used due to its suitability for a very large data set while preventing the model from overfitting. A hold-out method was used to predict the output values for the unseen data (test/validation data). Then, it will estimate the error for the test set which was used to evaluate the model. In this project, 50% of a hold-out method was used.

Seven basic expressions with four expression levels of intensities (excluding neutral expression, which is only one intensity) were used. During the holdout validation, the data set was divided equally into training and testing data set. After partitioning the data sets, training was proceeded using classification by the Random Forest method. During the classification, the data sets were trained with grown trees, which were equal to 100 trees. According to Breiman (2001), the more trees used in Random Forest, the better the accuracy will be obtained for prediction.

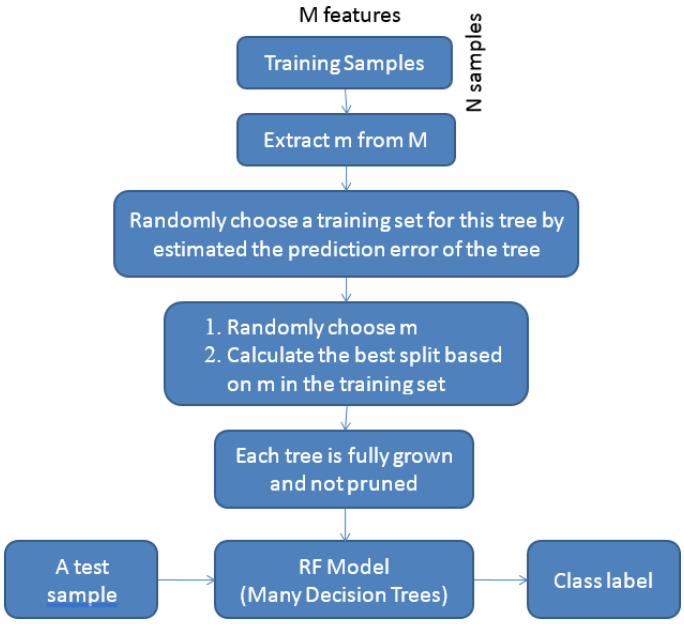

Figure 2: Random Forest process (Radenkovic, 2015).

For the validation test, the validation data sets were used to determine the performance of the Random Forest method in terms of predicting the true class of the subjects. One of the parameters to know whether Random Forest can make a better prediction is by measuring the out-of-bag (OOB) error, which also known as a probability of misclassification. Each tree has its own OOB data set, which is used for error estimation of the individual trees in the forest (Kulkarni & Sinha, 2013). We obtained a 7.14% OOB error. The lower the percentage of error rate, the better the performance of the method used for classification technique. In addition, OOB error for each grown tree also has been calculated and the graph for those errors for each of the grown trees is shown in Figure 3. Based on figure 3, 100 number of grown trees will give the lower value of OOB and that justifies the 100 grown trees used.

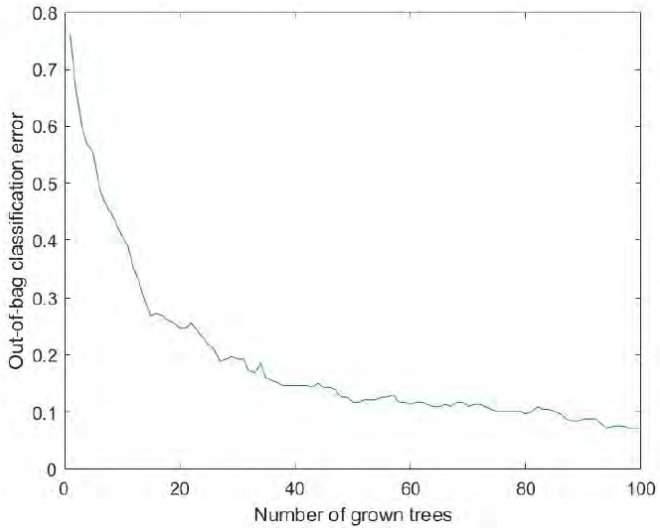

Figure 3: OOB error against number of grown trees

RESULTS AND DISCUSSION

Table 1 shows the comparison between Random Forest, PNN and SVM based on their recognition rate. The result for Random Forest obtained 94.71% of recognition rate and followed by work by Soyel and Demirel (2008) yielded 87.88% of the recognition rate. For the implementation of SVM in work from Tang and Huang (2008), they produced 87.81% of the recognition rate. In this comparison, BU-3DFE database is used, which contains basic expressions.

Table 1. Comparison of error rate.

Work	Method	Database	Emotions	Recognition rate (%)
Our work	Random Forest	BU-3DFE	7 basic emotions	94.7
Soyel and Demirel (2008)	Probabilistic Neural Network (PNN)	BU-3DFE	7 basic emotions	87.88
Tang and Huang (2008)	Support Vector Machine (SVM)	BU-3DFE	6 basics emotions	87.81

Table 2 shows the average confidence rate for the subject with expressions using Random Forest. The confidence rate is the percentage of probability of a subject is recognized based on their facial expressions. The higher the confidence rate obtained, the higher the probability of the expressions belongs to the subject tested, which results in the correct recognition of the subjects. For instance, on average, 75.4% of the subjects with anger expression could be accurately recognized. There were 24.6% of them are mistakenly recognized as other subjects' anger face. The lowest recognition rate is neutral, 58.2%. In other word, when a subject shows a neutral/expressionless face, it can be easily mistaken as other persons. The highest recognition is for fear with 82%. For fear, the facial features for each of the subjects are significant. Therefore, a subject with fear expression is easy to recognize as their fear faces are different from one another.

Table 2. Average of confidence rate (%) for subject with expressions.

Anger	Disgust	Fear	Happy	Neutral	Sad	Surprised
75.4	76.1	82	65.8	58.2	76.7	79

CONCLUSION

Our goal is to compare the Random Forest method with PNN and SVM. Based on the results obtained, Random Forest has given a good performance where it has obtained a 94.71% recognition rate, which is better than PNN and SVM obtained 18.18% of the misclassification rate. The experiment setting used in this work is similar to the existing works in Soyel and Demirel (2008) and, Tang and Huang (2008).

With these promising preliminary results, our future work will be on using Random Forest on real-time 3D face data, specifically on the face-recognition mobile application developed. In addition, a further test of the classification abilities of Random Forest for 3D facial expression classifications in consumer setting as well as data that consist of the non-frontal pose will be carried out.

ACKNOWLEDGEMENTS

This research is fully supported by Malaysian Ministry of Higher Education (MOHE) through Race Acculturation Collaborative Efforts RACE/1331/2016(4). The authors fully acknowledged MOHE and Universiti Malaysia Sarawak for the approved fund which makes this important research viable and effective.

REFERENCES

Bayramoglu, N., Zhao, G., & Pietikäinen, M. (2013). CS-3DLBP and geometry-based person independent 3D facial action unit detection. *Proceedings of International Conference on Biometrics* (ICB), pp.1-6.

Breiman, L. (2001). Random Forests, *Machine Learning*, *45*(1), 5-32.

Drira, H., Ben Amor, B., Srivastava, A., & Berretti, S. (2012). 3D dynamic expression recognition based on a novel deformation vector field and Random Forest. *Proceedings of the 21st International Conference on Pattern Recognition* (ICPR2012), pp. 1104-1107.

Gökberk, B. Salah, A.A., Akarun, L., Etheve, R., Riccio, D., & Dugelay, J.L. (2008). 3D face recognition. In D. Petrovska-Delacretaz, G. Chollet, B. Dorizzi (Eds.), *Guide to biometric reference systems and performance evaluation*, London: Springer-Verlag.

Hiremath, P.S., & Hiremath, M. (2013). Linear discriminant analysis for 3D face recognition using radon transform. In: Swamy P., & Guru D. (Eds.), *Multimedia processing, communication and computing applications*. Lecture Notes in Electrical Engineering, Vol 213. New Delhi: Springer.

Hu, H., Shah, S.A.A., Bennamoun, M., & Molton, M. (2013). 2D and 3D face recognition using Convolutional Neural Network, *TENCON 2017 - 2017 IEEE Region 10 Conference*, pp. 133-132.

Kim, D. Hernandez, M., Choi, J., & Medioni, G. (2017). Deep 3D face identification, *IEEE International Joint Conference on Biometrics (IJCB)* pp. 133-142.

Kremic, E., & Subasi, A. (2016). Performance of Random Forest and SVM in face recognition. *International Arab Journal of Information Technology*, *13* (2), 287-293.

Kulkarni, V.Y., & Sinha, P.K. (2013). Efficient learning of Random Forest classifier using disjoint partitioning approach. *Proceedings of the World Congress on Engineering* 2013 Vol II.

Lee, Y.H., & Han, C.W. (2006). 3D facial recognition using Eigenface and Cascade Fuzzy Neural Networks: normalized facial image approach. In: Grigoriev D., Harrison J., Hirsch E.A. (Eds.), *Computer Science – Theory and Applications*. CSR 2006. Lecture Notes in Computer Science, Vol 3967. Berlin, Heidelberg: Springer.

Mousavi, M.H., Faez K., & Asghar, A. (2008). Three-dimensional face recognition using SVM classifier, *Seventh IEEE/ACIS International Conference on Computer and Information Science* (ICIS 2008), pp. 208-213.

Radenkovic, P. (n.d.). RandomForest. Retrieved from home.etf.rs/~vm/os/dmsw/Random%20Forest.pptx

Rutkowski, L. (2004). Introduction to Probabilistic Neural Networks. In: *New Soft Computing Techniques for System Modeling, Pattern Classification and Image Processing*. Studies in Fuzziness and Soft Computing, Vol 143. , Berlin, Heidelberg: Springer.

Sandbach, G., Zafeiriou, S., & Pantic, M. (2012). Local normal binary patterns for 3D facial action unit detection. *Proceedings of* 19th *IEEE International Conference on Image Processing*, pp. 1813-1816.

Salhi, A.I., Kardouchi, M., & Belacel, N. (2012). Fast and efficient face recognition system using Random Forest and histograms of oriented gradients. *Proceedings of the International Conference of Biometrics Special Interest Group*.

Soyel, H., & Demirel, H. (2008). 3D facial expression recognition with geometrically localized facial features. *23rd International Symposium on Computer and Information Sciences*, pp. 1-4.

Tang, H., & Huang, T.S. (2008). 3D facial expression recognition based on properties of line segments connecting facial feature points. *Proceedings of 8th IEEE International Conference on Automatic Face Gesture Recognition*, pp. 1-6.

Ujir, H., Spann, M., & Hipiny, I. (2014). 3D facial expression classification using 3D facial surface normals. *In* H.A. Mat Sakim & M.T. Mustaffa (Eds.), *The 8th International Conference on Robotic, Vision, Signal Processing & Power Applications*, Lecture Notes in Electrical Engineering 291, pp. 245-253.

Wang, J., Yin, L., Wei, X., & Sun, Y. (2006). 3D facial expression recognition based on primitive surface feature distribution. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1399-1406.

Yin, L., Wei, X., Sun, Y., Wang, J., & Rosato, M.J. (2006). A 3D facial expression database for facial behavior research, *7th International Conference on Automatic Face and Gesture Recognition*, pp. 211 – 216.

Ying, A.C.C., Ujir, H., & Hipiny, I. (2017). 3D facial expression intensity measurement analysis. *In* Zulikha, J. & N.H. Zakaria (Eds.), *Proceedings of the 6th International Conference of Computing & Informatics*, pp. 43-48. Sintok: School of Computing.

Zhao, X., Kim, T., & Luo, W. (2014). Unified face analysis by Iterative Multi-Output Random Forests. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition,* pp. 1765-1772.

Trends in Undergraduate Research (2019) 2(2): e1-6
https://doi.org/10.33736/tur.1972.2019

Understanding Academic Performance based on Gender, Race, Stress and Sleep Quality

EUGENIE MOK YEN LING, TAN KOCK WAH*

Faculty of Cognitive Science & Human Development, University Malaysia Sarawak (UNIMAS), Kota Samarahan, 94300 Sarawak, Malaysia.
*Corresponding author: Tan Kock Wah (email: kwtan@unimas.my)

ABSTRACT

This study aims to explore gender, race, stress and sleep quality as factors which influence the academic performance of undergraduate students. The samples of study involved 80 students from the Faculty of Cognitive Science and Human Development, Universiti Malaysia Sarawak. Questionnaire used in the data collection comprising demographic information, stress questionnaire (Student Stress Inventory (SSI)) and sleep quality questionnaire (Pittsburgh Sleep Quality Index (PSQI)). The academic performance of the undergraduate students in this study was measured by their cumulative grade point average (CGPA). The result of the study showed that female students performed better than male students; whilst non-Bumiputera students performed better than Bumiputera students. The study also found that the higher the stress level of a student (SSI score), the lower his or her CGPA results. Moreover, the findings showed that there is a significant difference in FCSHD undergraduates' academic performance based on their sleep quality, in which the mean of CGPA for students with good sleep quality is higher than that of students with bad sleep quality. Ultimately, the study reveals that gender and race do play a part in the academic performance of students and having good sleep quality and a little bit of stress is equally important in learning and helpful in attaining good grades and hence, should be leveraged by students.

Keywords: gender, race, stress, sleep quality, students, academic performance

INTRODUCTION

University life is often deemed as one of the most stressful periods in a student's life. As students struggle to cope with exams, social pressures and future careers, many find themselves unable to deal with the situation. Some of them seek the easiest way out by ending their lives without thinking of the consequences. In a survey conducted by the American College Health Association (ACHA), it was reported that the suicide rate among young adults, whose age ranged between 15 to 24 years old, has tripled since the 1950s and currently, suicide is known to be the second most common cause of death among college students (Henriques, 2014).

In Malaysia, 7% of teenagers were reported to have suicide thoughts and the percentage of those who have turned the thoughts into actions were about half the percentage of those who have thought of committing suicide (Kok & Goh, 2011). In a study by Mustaffa, Aziz, Mahmood and Shuib (2013), the statistics released by the National Suicide Registry Malaysia (NRSM) suggested that the number of suicides had peaked dramatically in the last three years (1,156 victims) whereas the national statistics showed that females' suicidal rate was higher than males by the ratio of 3:1. Hence, the rate of suicides was indeed alarming. There were many factors that lead to suicide including physical illness, unwillingness to seek help, isolation and feelings of hopelessness. Nonetheless, it is vital to understand that academic performance is also one of the factors that can lead to suicides among students. According to Regier (2015), academic success was vital due to it being closely linked to the positive outcomes we value. For instance, those who did well in their studies are more likely to get hired and less likely to get involved in criminal activity and were happier than those who did not do well academically. Moreover, students who performed better in school were better able to make the transition into adulthood and to achieve occupational and economic success. On the other hand, according to Lageborn, Ljung, Vaez and Dahlin (2017), low educational attainment was correlated with a higher risk of suicide. According to Sin Chew daily, a suicide case was reported which involved a female student from Tunku Abdul Rahman University College (TARUC). It was found that the pressure the student experienced from daily life and studies were too much for her to handle, hence these were the main reasons the suicide occurred (Kwan, 2017).

Previous studies had been conducted to identify some of the factors that may have a huge impact on students' academic performance. According to Houston (2016), past research tend to focus more on factors such as personality traits, cognitive ability and motivational factors such as academic locus of control and self-efficacy which affect academic performance, however, little has been known whether demographic factors such as gender and race will have an impact on student's academic performance or not. Moreover, it is known that factors such as stress and sleep quality play a prominent role in determining students' academic performance, however the results obtained from past research were contradictory. Hence, this study aims to determine if there are differences in students' academic performance based on gender, race, stress and sleep quality respectively. This study will focus on undergraduates from the Faculty of Cognitive Sciences and Human Development, UNIMAS.

Research Hypothesis

Test Performed	Null Hypothesis	
Mann-Whitney U	$H_0 1$	There is no significant difference in FCSHD undergraduates' academic performance based on race.
	$H_0 2$	There is no significant difference in FCSHD undergraduates' academic performance based on gender.
	$H_0 3$	There is no significant difference in FCSHD undergraduates' academic performance based on sleep quality.
Spearman's Correlation	$H_0 4$	There is no relationship between stress and academic performance among FCSHD undergraduates.

MATERIALS & METHODS

This is a quantitative study with the aim to investigate the impact of gender, race, sleep quality and stress on academic performance. Convenience sampling technique was used to survey a sample of 80 students from the Faculty of Cognitive Science and Human Development (FCSHD), UNIMAS. The research instrument used in the data collection was questionnaire. The questionnaire consisted 3 sections where section A was concerned with collecting basic demographic information on gender, race and academic performance (CGPA), section B and section C comprised of stress questionnaire (Student Stress Inventory (SSI) and sleep quality questionnaire (Pittsburgh Sleep Quality Index (PSQI)) respectively.

The stress questionnaire consisted of 40 questions in which a Likert four points scale was used in the measurement where 1 corresponded to 'Never', 2 corresponded to 'Somewhat frequent', 3 corresponded to 'Frequent' and 4 corresponded to 'Always'. In terms of scoring, a total score of 40-80 meant mild stress, 81-121 meant moderate stress whereas 122-160 meant severe stress. On the other hand, the PSQI questionnaire consisted of 19 self-rated questions, taking into consideration the seven components which were subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medications, and daytime dysfunction over the last month where each component scores were weighted equally on a 0-3 scale, with 3 indicating the greatest dysfunction. The seven component scores were totalled to obtain a global PSQI score, ranging from 0-21. Participants who scored above 5 indicated that they had poor sleep quality whereas participants who scored 5 and below were considered to have good sleep quality. The SSI questionnaire was obtained from the Faculty of Education and Human Development, Sultan Idris Education University (Arip, 2016) whereas PSQI questionnaire was obtained from the Sleep and Chronobiology Center, Department of Psychiatry, University of Pittsburgh (Buysse, Reynolds III, Monk, Berman and Kupfer, 1988).

The questionnaires were pre-tested on a sample of 30 students to test the data-gathering mechanisms and to assess the validity and reliability of the questionnaire and the Cronbach's Alphas obtained for both Student Stress Inventory (SSI) and Pittsburgh Sleep Quality Index (PSQI) were 0.832 and 0.827 respectively. Moreover, students' self-reported Cumulative Grade Points Average (CGPA) were used to measure academic performance. Statistical Package for the Social Sciences (SPSS) would then be used to analyse the data collected.

RESULTS AND DISCUSSIONS

$H_0$1: There is no significant difference in FCSHD undergraduates' academic performance based on race.

Table 1. Mean rank of academic performance (CGPA) based on race

	Academic Performance (CGPA)	
	N	**Mean Rank**
Bumiputera	27	28.17
Non-Bumiputera	53	46.78

Table 2. Mann-Whitney U test

	Academic Performance (CGPA)		
	Z	**U**	**Sig. (2-tailed)**
Race	-3.390	382.5	0.001

A Kolmogrov-Sminov test indicated that the CGPA performances for the two groups did not follow a normal distribution. As such, a non-parametric test was performed. Based on Table 2, Mann-Whitney test indicated that the test was statistically significant in which the CGPA was greater for Non-Bumiputra (Mean Rank=46.78) than for Bumiputra (Mean Rank=28.17), U=382.5, p=0.001. Hence, H_{01} was rejected.

Similar result was shown in a study by Tang and Daneil (2017), in which their findings indicated there was a significant difference in the mean CGPA among different races of students. Thus, the findings of this study indicated non-Bumiputera students performed better than Bumiputera students academically. Moreover, findings from studies in Western countries indicated there was a correlation between academic performance and race. In a study by Battle and Lewis (2008), African Americans was found to exhibit lower educational achievement. In another study by Che Azmi and Harith (2012), it was found that non-Bumiputera Malaysians performed academically better than Bumiputera Malaysians. Hence, by exploring the cultural backgrounds of different races when it comes to academic performance, educators will be more sensitive in terms of teaching as they are able to explore the needs of students of different races.

$H_0$2: There is no significant difference in FCSHD undergraduates' academic performance based on gender.

Table 3. Mean rank of academic performance (CGPA) based on gender

	Academic Performance (CGPA)	
	N	**Mean Rank**
Male	40	34.79
Female	40	46.21

Table 4. Mann-Whitney U test

	Academic Performance (CGPA)		
	Z	**U**	**Sig. (2-tailed)**
Gender	-2.2	571.5	0.028

A Kolmogrov-Sminov test suggested that the CGPA did not follow a normal distribution for the two groups. Hence, a non-parametric test was performed. Based on Table 4, the Mann-Whitney test indicated that there was a statistically significant difference in CGPA between gender. The CGPA was greater for Female (Mean Rank=46.21) compared to Male (Mean Rank=34.79), U=571.5, p = 0.028. Hence, H_{02} was rejected.

The result was in line with the findings by Dayioğlu and Türüt-Aşik (2007) who reported that gender played a role in determining the academic performance of students and found that females performed better than males in

higher education in terms of their CGPA. Besides, similar result was shown in a study by Khwaileh and Zaza (2010) where females tend to do better academically than males. This could be explained by Gnaulati (2014) where it was found that girls had far better self-regulation than boys and self-regulation was important when it came to academic performances. This, however, was in contrast with the result obtained from Leander (2016) where it was reported that male students outperformed females in exams and one of the possible explanations may be due to the way the exams were designed rather than students' academic ability. Nonetheless, generally, Voyer and Voyer (2014) speculated that social and cultural factors could be a possible explanation why females performed better than males academically. Hence, this study provided evidence that gender influenced academic performance of students and this was largely supported by past studies discussed in this paper.

$H_0$3: There is no significant difference in FCSHD undergraduates' academic performance based on their sleep quality.

Table 5. Mean rank of academic performance (CGPA) based on sleep quality

	Academic Performance (CGPA)	
	N	**Mean Rank**
Good Sleep Quality	42	45.8
Bad Sleep Quality	38	34.6

Table 6. Mann-Whitney U test

	Academic Performance (CGPA)		
	Z	**U**	**Sig. (2-tailed)**
Sleep Quality	-2.145	575.5	0.032

A Kolmogrov-Sminov test indicated that the CGPA performances did not follow a normal distribution. Hence, a non-parametric test was performed. Based on Table 6, it was found that the test was statistically significant. Hence, H_{03} was rejected. It appeared that students with Bad Sleep Quality tend to score lower compared to those who could sleep well (Good Sleep Quality Mean Rank=45.8, Bad Sleep Quality Mean Rank=34.6), U=575.5, p=0.032.

Thus, as opposed to Nihayah et al. (2011) findings which concluded that sleeping hours did not affect students' academic performance (CGPA), the result obtained highlighted the importance of students having good sleep quality as it would be beneficial in the improvement of their academic achievements. This result was consistent with Curcio et al. (2006), where it was found that students who had poor sleep quality had a lower academic performance and students who had a more regular sleep-wake pattern tend to have a higher CGPA. These findings were supported by another study conducted by Kelly, Kelly and Clanton (2001), where it was shown that students who slept approximately 6 hours or less had a lower academic performance compared to students who slept for 9 or more hours per night. Hence, insufficient amount of sleep appeared to have a great impact on students' academic performance. One of the reasons was because sleepiness caused problems with concentration and mood, which made it harder for students to stay awake in class. Hence, the inability to concentrate and think clearly was likely to lead to poor academic performance.

Thus, students with good sleep quality achieved better academically than students with bad sleep quality. In sum, the result of the study provided evidence that students having good sleep quality are more likely to improve their academic achievement as they are more focused and alert.

$H_0$4: There is no relationship between stress and academic performance among FCSHD undergraduates.

Table 7. Spearman's correlation table

Stress	r_s	**Sig**
Academic Performance (CGPA)	-0.259	0.020

* Correlation is significant at the 0.05 level (2-tailed).

Based on the results in Table 7, the correlation coefficient, r_s was -0.259, meaning the higher the stress level of a student (SSI score), the lower his or her CGPA results. Since the significant value p = 0.020 where p < 0.05, thus

H_{04} is rejected. Therefore, there is a significant relationship between stress and academic performance among FCSHD undergraduates, r_s = -0.259, N = 80, p = 0.020.

According to Akgun and Ciarrochi (2003), respondents with a severe stress level were noticed to obtain higher CGPA. However, the results obtained in this study showed otherwise, in which students who had higher stress level did not perform as good as those who had lower stress level in academic terms. Furthermore, in a study by Sloboda (1990), the results showed that stress led to a detrimental academic performance in university. In another study by Khan and Chaudary (2014), it was found that students of a public university experienced stress which resulted in their inability to meet academic goals, poor academic performances and continuous decline in grades. One possible explanation is that stress negatively impact on students' ability to concentrate. When a student was not able to focus and pay attention in his or her studies, the amount of information absorbed and stored in the memory would also decrease, leading to poor grades (Lumontod, 2017). Hence, this study provides evidence that stress had an impact on the academic performance of students. In all, this study provided some insights into how different stress levels could impact differently on students' achievements: exposure to a moderate level of stress could have the potential to push an individual to perform better but excessive stress could be detrimental.

CONCLUSION

In conclusion, findings from the research revealed that there wa a significant difference in FCSHD undergraduates' academic performance based on gender, where female students performed better than male students; whilst non-Bumiputera students performed better than Bumiputera students. The study also found that the higher the stress level of a student (SSI score), the lower his or her CGPA results. Moreover, the findings showed that there was a significant difference in FCSHD undergraduates' academic performance based on their sleep quality, in which the mean of CGPA for students with good sleep quality was higher than that of students with bad sleep quality. As student life revolved around academic related matters, therefore its importance should not be overlooked as low educational attainment was often associated with a higher risk of suicide. Thus, this study provided more insight for undergraduate students regarding the factors affecting academic performance where a lower stress level and good sleep quality could be important in achieving good grades.

ACKNOWLEDGEMENTS

We would like to thank all parties who had assisted in making this project possible, especially to the respondents who had volunteered to participate in the study

REFERENCES

Akgun, S. & Ciarrochi J. (2003). Learned resourcefulness moderates the relationship between academic stress and academic performance. *Educational Psychology*, *23*(3), 287-294.

Arip, M. (2016). *Manual of Student Stress Inventory (SSI) development, validity and reliability of Student Stress Inventory (SSI).* Retrieved from https://www.researchgate.net/publication/316662054_Manual_of_Student_Stress_Inventory_SSI_Development_Validity_And_Reliability_of_Student_Stress_Inventory_SSI

Battle, J. & Lewis, M. (2008). The increasing significance of class: The relative effects of race and socioeconomic status on academic achievement. *Journal of Poverty, 6*(2), 21-35.

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1988). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*, 193-213.

Che Azmi, A., & Harith, S. (2012). Higher order cognitive skills and performance of undergraduate students: Do these skills differ among Asian groups? *Far East Journal of Psychology and Business, 9*(3), 19-28.

Curcio, G., Ferraraa, M., & Gennaro, L. D. (2006). Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews, 10*(5), 323-337.

Dayioğlu, M., & Türüt-Aşik, S. (2007). Gender differences in academic performance in a large public university in Turkey. *The International Journal of Higher Education and Educational Planning, 53*(2), 255-277.

Gnaulati, E. (2014). Why girls tend to get better grades than boys do. Retrieved from https://www.theatlantic.com/education/archive/2014/09/why-girls-get-better-grades-than-boys-do/380318/

Henriques, G. (2014). *The college student mental health crisis*. Retrieved from https://www.psychologytoday.com/us/blog/theory-knowledge/201402/the-college-student-mental-health-crisis

Houston, D. C. (2016). *Explaining race differences in academic performance: The role of perceived expectations and outcome valence*. (Unpublished master's thesis) Winston-Salem State University, North Carolina, United States of America.

Kelly, W. E., Kelly, K. E., & Clanton, R. C. (2001). The relationship between sleep length and grade-point average among college students. *College Student Journal, 35*(1), 84–86.

Khan, A. M., & Chaudary, A. M. (2014). Impact of stress among students of a public sector university. *Journal of Research and Reflections in Education, 8*(1), 48-54.

Khwaileh, F., & Zaza, H. I. (2010). Gender differences in academic performance among undergraduates at the university of Jordan: Are they real or stereotyping? *College Student Journal, 45,* 633-648.

Kok, J. K., & Goh, L. Y. (2011). *Young people and suicide issue*. Paper presented at International Conference on Humanities, Society and Culture, Singapore.

Kwan, L. (2017). *20-year-old TARUC student commits suicide due to stress from work and studies.* Retrieved from https://www.worldofbuzz.com/20-year-old-taruc-student-commits-suicide-due-stress-work-studies/

Lageborn, C. T., Ljung, R., Vaez, M., & Dahlin, M. (2017). Ongoing university studies and the risk of suicide: A register-based nationwide cohort study of 5 million young and middle-aged individuals in Sweden, 1993–2011. *BMJ Journals, 7*(3), 1-8.

Leander, S. (2016). *Gender gap discovered in science-exam performance*. Retrieved from https://asunow.asu.edu/20160602-gender-gap-discovered-science-exam-performance

Lumontod, R. (2017). The effect of stress on students' academic performance. Retrieved from https://psychlens.com/effect-of-stress-on-students/

Mustaffa, S., Aziz, R., Mahmood, M. N., & Shuib, S. (2013). Depression and suicidal ideation among university students. *Journal of Social and Behavioural Sciences, 116*, 4205-4208.

Nihayah, M., Ismarualyusda, I., Syarif, H. L., NurZakiah, M. S., Baharudin, O., & Fadzil, M. H. (2011). Sleeping hours and academic achievements: A study among biomedical science students. *Procedia Social and Behavioral Sciences, 18,* 617-621.

Regier, J. (2015). Why is academic success important? Retrieved from https://saskschoolboards.ca/wp-content/uploads/2015/08/2011siast.pdf

Sloboda, J. A. (1990). Combating examination stress among university students: Action research in an institutional context. *British Journal of Guidance & Counselling, 18* (2), 124 – 136.

Tang, H. E., & Daneil, I. L. (2017). Demographic background, gender and race: Their impacts on students' academic success. *Pertanika Journal of Social Sciences and Humanities, 25*, 105-114.

Voyer, D. & Voyer, S. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140(*4), 1174-1204.

Trends in Undergraduate Research (2019) 2(2): f1-8
https://doi.org/10.33736/tur.1395.2019

Implementation of Guided Group Academic Writing Using Online Learning Tools

MUHAMMAD JAZLAN BIN AHMAD KHIRI

Faculty of Language and Communication, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia
Corresponding author: akmjazlan@unimas.my

ABSTRACT

Writing is an important skill necessary for students pursuing a higher level of education. The challenge for second language learners would be to write accurate academic essays as required by the course for students that has been aligned to the Common European Framework (CEFR) B1 standard since 2015 at a public institution of higher learning. The academic English course offered at a public institution of higher learning was subjected to a guided academic writing using online tools by Harju and Åkerblom (2017) description of the teacher's role in facilitating learning to achieve the intended learning outcome which is to write an academic essay focusing on its features. The participants in this study were second year undergraduates in the second semester of their study who struggled to write an essay early in the semester. This study used a quantitative analysis collected from 10 participating students (*n=10*). The results found that the students made improvements to the feature of their essay and their feedback of the intervention were processed using frequency analysis indicated a hundred percent preference of the students to learn by their own discretion guided by the instructor and supported with online learning tools.

Keywords: Group, Guided, Online learning tools, Writing

INTRODUCTION

Learners today collaborate to produce an academic writing. Yarrow and Topping (2001) found in their comparison that learners who wrote as a group has better quality compared to individual writing. In addition, they also found that there was improvement to the students' positive self-esteem in writing with a partner or a group. Vanderveen and Wells (2012) stated that collaborative learning using online tools such as Google Docs for students essay writing was useful to monitor the progress and changes made to improve their writing.

As an effort to corroborate with previous researches on the benefits of using peer writing (Harju & Åkerblom, 2017; Huskin, n.d.; Vanderveen & Wells, 2012; Yarrow & Topping, 2001), this study validated the findings made through the use of qualitative approach where the data were analysed through the students' overall writing performance through assessment. In addition, group writing should be continued for teaching and learning as the tasks provided to the learners would be meaningful for them to actively seek for information compared to the all-knowing teacher.

This study involved a group of second year students taking an academic English course. They were involved in the guided group academic writing using online learning tools through three objectives. The first objective was to determine whether the students writing features had improved through the instructor's intervention through guided learning activity when they wrote an argumentative essay. This was due to the problem with their first essay on Colours and Cultures in Malaysia that had incorrect features ranging from a weak hook, inaccurate thesis statement, and ambiguous supporting details. These issues were largely attributed to the poor use of transitional words and weak vocabulary.

The second objective aimed to identify pedagogical factors that influenced the guided group academic writing that focused on outcome. This is in line with Harju and Åkerblom (2017) who argued that learning in the 21st century is accomplished by giving a greater autonomy for learners to achieve the learning outcome by their own initiative over a teacher centred classroom.

Lastly, the study aimed to use an online learning tool such as E-learning Enrichment and Advance Platform (eLEAP) and Google Docs to see if the use of these tools can make significant improvement to the students

learning. The intervention made focused on using selected videos from Youtube to supplement or used as a trigger to give a clearer idea on how to write accurate features of an essay.

MATERIALS AND METHODS

The subject in this academic English course in this study was CEFR aligned to B1. This level referred to learners who "Can write straightforward connected texts on a range of familiar subjects within his field of interest, by linking a series of shorter discrete elements into a linear sequence"(Council of Europe, 2018, p. 23). The marking scheme used to grade students' assessment for this writing course was based on the Cambridge CEFR B1 Assessment in terms of language, content, and organization. Communicative achievement was not applicable as this particular descriptor is for assessing speaking assessment, not writing in relation to this study. The textbook used for this course is Weaving it Together 3 (4th Edition) published by National Geographic Learning with Cengage Learning and the material was benchmarked to be suitable for B1 learners. The book contains a diverse set of topics ranging from food, customs, environment that are supplemented with reading comprehension exercises and writing of genres such as expository, argumentative and cause and effect. Over the course of 14 weeks in a semester, the students were also taught on American Psychology Association (APA) style 6th edition citation and referencing.

The student-centred approach is important and needs to be applied across all levels of learning institutions. However, there is real concern whether the educators know what student-centred approach is, so as to apply it (Estes, 2004). Cuban (2001) even asserted the use of the student-centred teaching approach as an ideal that eventually will weaken over time despite its benefit to student learning. The implementation of the student-centred approach is a crucial and part of government policy (Malaysia Ministry of Education [MOE], 2013). In addition, the major benefit of this approach is that it would allow the instructor to have greater freedom in designing tasks to achieve the necessary learning outcomes and could elicit diverse set of responses from the students (Porath, 2014). Soo (2015) pointed out that a set of clear rules in completing a task and maintaining group harmony can make the learners work together even though they might have peers who have lower proficiency compared to the rest. Phillipakos (2017) highlighted that effective learning can be achieved through three steps. The first is in making it clear to students about the graded criteria for the writing assessment. Second, is genre evaluation where the features of a specific essay that describes its purpose are made clear while the third step is maintaining students' motivation to keep them focused in completing the objective of the task. Berling (1991) also pointed out the need for instructors to "decentralise' their roles as the instructor to the students. Harju and Åkerblom (2017) also concurred with a similar finding to emphasise that instructors should not place a rigid emphasis on the framework and structure in the learning of writing, but give the students a weaker framing of writing knowledge for them to develop the product as a group. One common agreement among scholars is the need to explain the standards of assessing to the learners to fulfil the expected outcome of the course in order for them to function effectively as a group to produce a quality piece of writing (Cuban, 2001; Estes, 2004; Harju & Åkerblom, 2017; Phillipakos, 2017; Soo, 2015). This shared perspective features the need for a more progressive, and efficient role for the instructor to implement guided group academic writing activities where learners expectation of learning is different due to the use of technology. There is therefore, a need to improve current practice and delivery of the academic English course.

There are a myriad of online learning tools available of the internet for students and instructors such as blogs, and Wikipedia (Velasco, 2018). Such variety poses a challenge in the selection of tools that fit the requirement of the prescribed task. Google docs was found to the appropriate tool for writing as it encourage seamless group learning that does not overwhelm learners compared to discussion threads (Morrison, 2014). Vanderveen and Wells (2012) too supported the use of Google Doc for peer writing activities, as it is user friendly due to the familiarity of the feature of the program similar to Microsoft Word. Moreover, such online learning tools support learning outside the classroom and improve the students' accessibility to information relevant to their learning for a course (Palak & Walls, 2009). Besides, they are familiar with asynchronous communication tools such as Facebook and Twitter to get the latest information to follow the current trend and happenings around them (Black, 2012). Bryant and Bates (2015) concluded that the use of online learning tools encouraged student enquiry, and create space for learners to challenge each other and innovate knowledge in the process. This study used Google Docs and the instructor participated to monitor and give minimal advice through comments for improvement. Additionally, the students also used eLEAP, the university's online learning platform for the students to share their files, and chat to exchange information with their peers.

Guided Group Writing

The grading criteria for their assessment and the features that must be included in a five-paragraph essay were explained to the students at the first week of the course. These features included the hook, background information, thesis statement, topic sentences, supporting details, and concluding paragraph. The following week focused on developing the thesis statements. Students were taught to write simple sentences to define their topic sentence ideas before rewriting them into one compound sentence. At week three, they wrote an expository essay as a group on the topic of Colours and Cultures in Malaysia. They were required to submit it via eLEAP and for future reference to use as trigger to write the five-paragraph essay. The quality of the group writing after week three was found to be weak in terms of its features. The thesis statement was unclear, the topic sentence was not relevant to the essay, and supporting details were poorly connected due to limited use of transitional words and limited vocabulary.

Intervention in the form of a guided group academic writing started in week four. Each student was assigned a paragraph of an argumentative essay in class. They were asked to enquire one each other to create a complete puzzle of the fragmented argumentative essay. They were then instructed to label the features of the puzzle they had matched to its content. The next step was to divide them to their respective groups for the graded argumentative essay which must be submitted in three weeks. The following weeks were on APA style 6th edition citation and referencing followed by presentation skills where they had to present the graded argumentative essay after week seven. The detailed progress of the group for the first seven weeks during the course activity is shown in Table 1.

Table 1. Course activity for the first seven weeks

Week	Course activity
1	Topic sentence, supporting details and concluding sentence based on the theme in Unit 1: Symbols
2	Thesis statement writing
3	Write a hook and background information and concluding paragraph and complete Colours and Cultures in Malaysia. Submit on eLEAP.
4	Writing argumentative essay for assessment
5	Citation and referencing
6	Presentation skills
7	Submission of argumentative essay assessment

The argumentative essay for the academic English language course has a maximum number of three students per group. The group comprised of the same gender that was coincidental and this gender-based composition is not relevant to the purpose of the study. The three groups for the argumentative essay assessment and the topics assigned to the students were as follows:

1. Superstitions in our culture have positive impacts on our lives,
2. Everyone should travel by public transport to reduce air pollution.
3. Superstitions in our culture have negative impacts on our lives.

Google random number generator was used to randomly assign the topics. The male students ($n=6$) were divided into two groups ($n=3$) and one was assigned to write topic no. 1 and the other topic no. 2 as stated above while the female students ($n=4$) wrote on topic no. 3. All groups used Google Doc to add and edit the content of their argumentative essay. The marking scheme used to assess the students' essay was based on the Cambridge English CEFR B1 achievement scale. The assessed features included content (21%), organization (10%), language (10%) citation and references (9%). The students were given a period of three weeks to complete their essay and submit a print copy to the instructor. The completed and submitted essays were compared with their previous essay on Colours and Cultures in Malaysia (CCM) to determine if there were any improvements to these students' writing.

A questionnaire was distributed and the data collected included the students' personal information such as their Malaysian University English Test (MUET) band result, the programme they are studying at university, and most importantly, how the guided group writing and technology had affected their writing. Four point Likert scale was used instead of five because neutral or unsure or do not know response has no significance in acquiring the students' feedback (Krosnick & Presser, 2012). As observed by Krosnick and Presser (2012), the options for neutral, unsure or do not know are unnecessary where respondents' biasness or poor motivation in participating can affect the study. The instrument used to collect the data from the students in this study required their views

on an item in the questionnaire (*strongly disagree=1, disagree=2, agree=3, strongly agree=4*). Data on the implementation of the course instructor lesson and online learning environment data were analysed using frequency analysis and presented as Table 1 and 2, respectively.

RESULTS AND DISCUSSION

The data presented in a column chart from the questionnaire in Figure 1 revealed that seven of the students obtained a Band 3 MUET result while the remainder were Band 2 with three students. These students' proficiency was a mix of below average and average in the mastery of the English language for higher education.

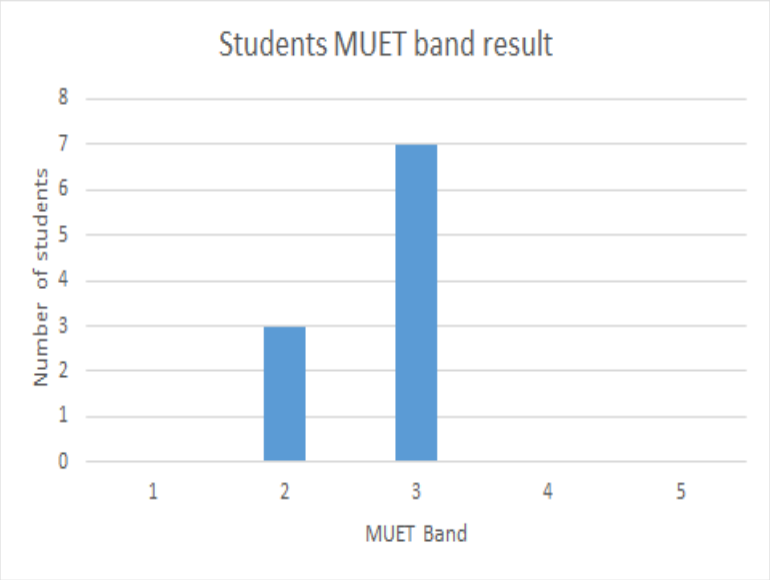

Figure 1. Participating students MUET band results used in the study

This study suggested that peer guided learning had improved the students writing as stated in Figure 2. The marks for these groups of students' essay were compared, which were Colours and Cultures in Malaysia (CCM), the expository essay and the student argumentative essay in which the selected topics were "Superstitions in our culture have positive impacts on our lives (SSC[+])", "Everyone should travel by public transport to reduce air pollution (EPT)" and Superstitions in our culture have negative impacts on our lives (SSC[-])". The percentage of comparison of the students' previous essay to the argumentative essay are shown in Table 2.

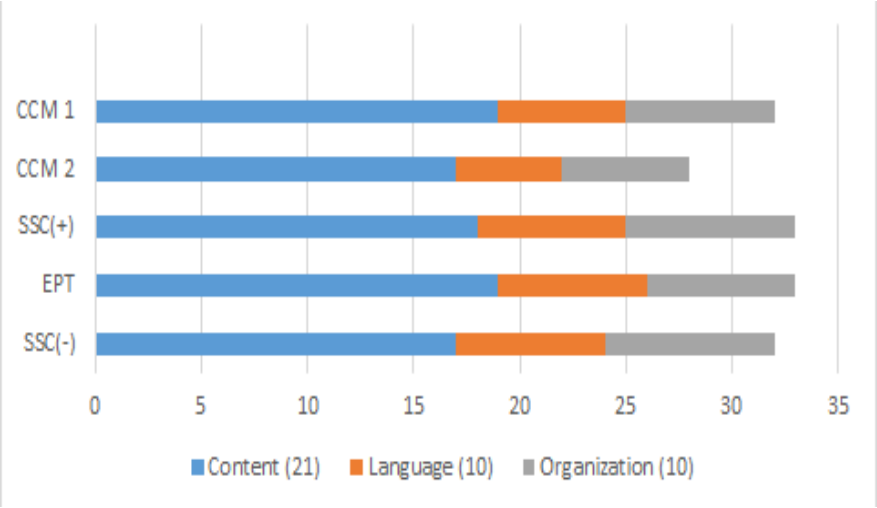

Figure 2. Comparison of students group writing marks according to assigned topics

Table 2. Comparison of the students' previous essay to the argumentative essay

Feature	Content	Language	Organization	Total
CCM 1 vs SSC(+)	94.74%	116.67%	114.29%	103.13%
CCM 1 vs EPT	100.00%	116.67%	100.00%	103.13%
CCM 2 VS SSC (-)	89.47%	116.67%	114.29%	100.00%

A comparison between Figure 1 and Table 2 showed that the students' language and organization improved after the intervention through the guided writing activity over the course of two weeks. The marks awarded for language and organisation showed an increase of at least 16.67% and 14.29%, respectively. In addition the total marks the student acquired too had increased by at least 3.13%. The marks for EPT content indicated no difference when comparing their previous essay to the latest. However, two groups of students ($n=7$) showed a decrease in terms of their content with a reduction of 5.26% for the group assigned on CCM 1 and SSC (+) while CCM 2 and SSC (-) fared worse with a reduction on 10.53%

The likely reason why the group assigned with the SSC (+) and SSC (-) topic showed a decline in the content is likely because of the subject matter in question of the assessment. The topic on superstition may not be a familiar topic because of the culture and taboo associated to the topic limits the group from expressing their ideas to establish a common ground. This is different for the group assigned on the topic EPT because public transport is a familiar topic for students in public institution of higher learning where they commute using the bus regularly or use any forms of public transport.

The question items in Table 3 investigated on factors in the teaching approach that contributed to the student writing. The data which were processed using frequency analysis indicated the students' opinion that they have control over the materials, learning process, and able to produce an academic essay where all students strongly agree to all the items in the questionnaire. The findings in this table clearly indicate that these learners preferred to be in control of their learning. This finding is reflected in the data where the mean (4.00), median (4.00) and std. deviation (.000) are equal indicating that they strongly agreed to a guided group writing activity. This finding justifies Harju and Åkerblom (2017) view on the process of implementing a group writing activity for students. They need to be guided in terms of the course purpose, course outcome, assessment criteria, enquiry based learning between their peers and instructor, and a promotion of self-directed learning ultimately benefits the students in the learning process (Harju & Åkerblom, 2017).

The data as shown in Table 4 enquired whether online learning tools such as ELEAP and Google Docs helped to improve these students' writing as a group. The mean and standard deviation on two items in the questionnaire on the use of ELEAP and Google Doc were reported at 3.80 with a standard deviation of .422. There is a need to note that online learning tools are meant to support or supplement the activity. Furthermore, the instructor does not edit the Google Doc frequently because it is the learners' responsibility to take control and manage their own learning while the instructor advises them on the Google Doc to improve their writing based on the CEFR B1 writing rubric. The learners' however do strongly agree that the use of Google Doc did help with the students writing. The data on setting specific dates over the course of two weeks with an interval of reviewing the students writing every five days because the argumentative essay must be completed within 2 weeks (week 7-8) was rated with a mean, median and standard deviation of 4.00, 4.00 and .000 respectively. These findings indicated the instructor's effort and agreement with the students that they needed to show progress at specific dates in weeks seven and eight of the course. Similarly, the instructor uploaded relevant videos to supplement their learning for the course in week two until six as stated in Table 1.

Further, the findings suggested that using online learning tools gave these students space for interaction (i.e. eLEAP) and to explore new information. The viewing of the suggested videos set by the instructor as a start enabled them to explore new ideas to develop more complex ones when writing and interacting with one another as they utilised Google Doc to write their essay. Additionally, the small sized group (n=10) students in this class which formed smaller groups for the argument essay were able to align their interest in accomplishing the task towards a common learning goal (Huskin, n.d.). This finding showed that these learners preferred guidance as opposed to instruction in their learning.

Table 3. Students perception of the student-centred lesson for academic writing

	My course instructor provides clear explanations of the grading criteria for the argumentative essay assessment.	My course instructor outlines the lesson for argumentative essay and gave me tasks to better understand the essay.	My course instructor uses clear examples to help me understand argumentative essay.	My course instructor allows me to ask questions and answers satisfactorily in class.	My course instructor asks me questions relevant to the argumentative essay.
N	10	10	10	10	10
Mean	4.00	4.00	4.00	4.00	4.00
Median	4.00	4.00	4.00	4.00	4.00
Std. Deviation	.000	.000	.000	.000	.000
Range	0	0	0	0	0
Minimum	4	4	4	4	4

Table 4. Students perception of the online learning tools used

	My course instructor's use of ELEAP is helpful for my learning.	My course instructor's use of Google Doc to help my writing is helpful.	My course instructor edit my Google Doc writing frequently.	My course instructor set dates for reviewing our essay in Google Doc.	My course instructor's uploaded video on argumentative essay is helpful.
N	10	10	10	10	10
Mean	3.80	3.90	3.80	4.00	4.00
Median	4.00	4.00	4.00	4.00	4.00
Std. Deviation	.422	.316	.422	.000	.000
Range	1	1	1	0	0
Minimum	3	3	3	4	4

CONCLUSION

The instructor does play a role to guide students' understanding and give them opportunities to apply what they had learnt into practice in every class. Additionally, the instructor too must make the additional effort of using online learning tools to guide the students learning beyond the classroom. Ultimately, the combination of the instructor's effort and use of online learning tools would ensure the achievement of a course's outcome.

This study, which focused on writing, offered insights into improving the students' overall writing skill as aligned to the CEFR descriptors that measures mastery of the language. Students would prefer the opportunity for activities that guide their learning over time to develop the skills needed for the course. Such guidance involves description of the grading criteria, task relevance, instructor's knowledge and the opportunity to ask and respond to questions. Additionally, it should be noted that the topic to assign the students must be a subject they are familiar with to improve the reliability of the intervention to validate the effectiveness of guided group academic writing.

Nevertheless, there is a need to increase the validity of this study through increasing the number of participants to improve the overall reliability of the guided group academic writing approach implementation in classroom. The findings from a small number of participants (*n=10*) cannot be generalised to indicate the effectiveness of online learning tools in guided group academic writing. However, this study can provide a basis for future research in the area.

REFERENCES

Berling, J. A. (1991). Getting out of the way: A strategy for engaging students in a collaborative way. *Teaching Theology and Religion, 1*(1), 31-35.

Black, L. W. (2012). Blog, chat, edit, text, or tweet? Using online tools to advance adult civic engagement. *New Directions for Adult & Continuing Education, 2012*(135), 71-79.

Bryant, J., & Bates, A. J. (2015). Creating a constructivist online instructional environment. *TechTrends, 59*(2), 17-22.

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment*. Retrieved from https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989

Cuban, L. (2001). How did teachers teach, 1890-1980. *Theory Into Practice, 22*(3), 159-165.

Estes, C. A. (2004). Promoting student-centered learning in experiential education. *Journal of Experiential Education, 27*(2), 141-160.

Gamill, D. M. (2006). Learning the write way. *The Reading Association*, *59*(8), 754-762.

Harju, A., & Åkerblom, A. (2017). Colliding collaboration in student-centred learning in higher education. *Studies in Higher Education, 42*(8), 1532-1544.

Huskin, P. R. (n.d.). Engagement strategies for increasing student writing success. *Education, 136*(3), 283-290.

Krosnick, J. A., & Presser, S. (2012). Question and questionnaire design. In P. V. Marsden, & J. D. Wright (Eds.), *Handbook of Survey Research* (pp. 263-313). Bingley, UK: Emerald Group Publishing Limited.

Malaysia Ministry of Education [MOE] (2013). *Malaysia education blueprint 2013-2025: Preschool to post-secondary education*. Retrieved from https://www.moe.gov.my/images/dasar-kpm/articlefile_file_003108.pdf

Morrison, D. (2014). *Students perceptions of online group work: What they really think and how to make it work*. Retrieved from https://onlinelearninginsights.wordpress.com/2014/02/25/student-perceptions-of-online-group-work-what-they-really-think-and-how-to-make-it-work/

Palak, D., & Walls, R. T. (2009). Teacher's beliefs and technology practices: A mixed-methods approach. *Journal of Research on Technology in Education, 41*(4), 417-441.

Phillipakos, Z. A. (2017). Giving feedback: Preparing students for peer review and self-evaluation. *The Reading Teacher, 71*(1), 13-22.

Porath, S. (2014). Talk less, listen more: Conferring in the reader's workshop. *The Reading Teacher, 67*(8), 627-635.

Soo, H. K. (2015). Preparing English learners for effective peer review in the writers' workshop. *The Reading Teacher, 68*(8), 599-603.

Vanderveen, J. M., & Wells, J. W. (2012). Group work online. In R. K. Morgan, K. T. Olivares, M. D. Dixson, A. D. Gavrin, M. C. Morrone, J. E. Lafuze, & A. S. Morrone (Eds.), *Quick Hits for teaching with technology: Successful strategies by award-winning teachers* (pp. 90-91). Bloomington, IND: Indiana University Press.

Velasco, R. C. L. (2018). Infusing classrooms with web 2.0 technologies. *Technology and Engineering Teacher, 77*(6), 36-39.

Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology, 71*, 261-282.

Trends in Undergraduate Research (2019) 2(2): g1-7
https://doi.org/10.33736/tur.1466.2019

Productivity Convergence of High-Tech Automotive Industry in Asian

AZIRA RUSU & DAYANG AFFIZZAH AWANG MARIKAN*

Faculty of Economics and Business, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

*Corresponding author: amdaffizah@unimas.my

ABSTRACT

This study is attempting to study the convergence of high-technology productivity of the automotive industry in Asia. This study uses time series data collected from 10 Asian countries from 2002-2016. In addition, this study was tested using the panel method proposed by Phillips and Sul (2007a), to identify whether the convergence of automotive technology high-tech productivity convergence or divergence. The results for full panel convergence in this study show divergence. However, in this study there were three convergence clubs found. The first clubs are Japan, South Korea and Thailand, the second club is Indonesia and Iran while the third club is Malaysia, Vietnam and the Philippines. This study suggests that governments need to play an important role to implement good policies to attract more Asian countries to work with each other. This study can be summarized that economies in Asian countries experience different levels of development and the shift in economic behaviour is very different among Asian countries.

Keywords: Convergence, Divergence, Asian, Automotive industry

INTRODUCTION

High-tech can be defined as a high-technology industry where a group of companies involved in high-technology products in terms of research, development, production and technical services. Most technology products of a country must be included in high-tech processing (Ding, 2016). High-tech can characterize by knowledge of e-intensity and technology-intensity which represent the comprehensive strength and the overall competitiveness of a region or country. Moreover, in the economic perspective which to use or develop advanced technologies known to the public is potentially leading to higher investment and resulting for future growth of a country. The productivity convergence in high-tech automotive industry is depending on the request from the customer. Therefore, future innovation will focus on technology trends to the customers. Automobile manufacturers will look in terms of technology trends such as electrification, autonomous driving, diverse mobility, and connectivity.

According to Ding (2016) the most important of the high-tech industry in Asian is in the development of the high-tech industry that has growth of national economics and the sector has become an area of focus for many countries as it adds significant value to their economic. Where since the year 1980s, the diffusion of the technology brought on by globalization, so the high-tech industry has become an important area international economic competition in the world (Liu & Tsai, 2007). The importance of high-technology industries, especially automotive that enables a country to be more developed on the value chain. This enables the spread of advanced technology in the industry and moves towards acquiring more technology-based firms than those employing the workforce. This will increase the reputation of the market and the firm's performance will become more efficient and increase the country's revenue. Furthermore, this will make a country specialize in high-tech automotive and increases the growth domestic product (GDP). Developing countries have problems with the productivity convergence of high-tech where the failure new technology to boost productivity that knows as Solow Paradox. It is that technological advances increase productivity only after a long lag and this might lead to fear of the job-destroying effect. According to Bai and Ng (2004) a test that distinguishes between non-stations, which is not convergence, which is derived from one or the same source or variable. For the whole economy, the overall indication of convergence cannot be traced back to both factors indicating that not only the country's specific arguments need to be considered. On the other hand, developments or technology stocks that are often experienced have led to more differences across the Europe region and overall economic perspective are ensured when looking at the major sectors of the Europe region. The problem is Japan use many more industrial automotive than in emerging countries, though China is beginning to invest heavily in automotive as its labour costs rises. So, often branded as

the lump of labour fallacy, where belief that there is only so much work to go around so if the machines do more of it and less is left for another. This is fallacious because as technology displaces workers from an occupation, it enriches other who spend their gains on good and services that create new employment for the worker whose job have been automatic away with machines. On the other hand, high-tech is often viewed as high risk, but offering the opportunity for high profit to countries. So, if the countries have low technology can give impact to the productivity. Hence, Japan has made substantial contributions to various fields such as electronics, machinery, earthquake engineering, automotive, optical, industrial robotics, metal and semiconductors. Furthermore, this study is intended to find the key competence factors that can contribute to the high-tech productivity convergence in the automotive industry in Asia. Besides that, it is also important to find out the affect productivity convergence of high-tech in Asian as objective of this study.

Convergence Theory

According to Crossman (2017), convergence theory is a theory that assumes that countries are in the early stages of industrialization and resemble other industrialized nations. Money from other countries can be poured out for developing countries and can take advantage of these opportunities. These countries will become more vulnerable to the international market. Not only that, convergence in the economy or better known as the catch-up effect has hypothesized that the poor country has per capita economic income will be more likely to grow at a faster pace than the rich country. This is because, the decline in returns, especially capital that is not comparable to the capital-rich nation. Therefore, poor or developing countries can follow the methods of production, technology and institutions of developed countries. In addition, convergence theory is an economic theory that assumes that the concept of development takes place is a good thing universally and determined by economic growth. This will be able to unify the developing countries with advanced countries. This is because, the economies of developing countries are increasingly expanding from most developed industrial countries. That is, all nations must unite to achieve the same level. Furthermore, divergence is different from convergence because divergence happens when two points are not united to one another. Divergence in the economy occurs when capital invested in poor countries and the international market does not know or do not know that there are investment opportunities in the country then catch-ups will not happen. This will be the economic divergence. The unstable countries will be more likely to divergence because there are social and political factors such as lack of job opportunities or infrastructure.

MATERIALS & METHODS

Data Description

This study is about productivity convergence of high-tech automotive industry in Asian which selected several countries as a study case such as Japan, China, South Korea, India, Thailand, Indonesia, Iran, Malaysia, Vietnam, and Philippines. In this study, the ranking of countries will be determined based on higher of productivity convergence of high-tech. The variable is total of productivity after total of production (TOP) divided by wage or salaries (WG). All the data used for this empirical analysis study obtained from the OICA production statistics and World Bank Data. In this study, all the data used are annual data, from the year 2002 to 2016.

The Non-linear Factor Model

This study applies a nonlinear time varying factor model from Phillips and Sul (2007a), this has some advantages to studying the shifting of high-tech automotive industry in Asia. This is because, it includes simple linear regression and single-sided regression coefficient tests with standard normal critical values. This method is useful for observing and measuring the transition to long-term growth paths and individual transitions over time, over the common trend, representation or aggregate variables. Therefore, to investigate the convergence of high-technology in the automotive industry in the Asian it is necessary to use the appliances applied by Phillips and Sul (2007a). Panel data are usually decomposed by:

$$X_{it} = g_{it} + a_{it} \quad (1)$$

In equation (1) x_{it} is a panel log income per capita for nation i, (i = 1,…,N) and at time t = 1,…,N. It is common to decompose x_{it} into two components as systematic, git and transitory, a_{it}. At the point, we do not assume any parametric assumptions of g_{it} and a_{it}, it is because the framework may include linear, nonlinear, stationary and non-stationary processes. In equation (1) may contain both common and idiosyncratic components in g_{it} and a_{it}.

$$X_{it} = \left(\frac{g_{it} + a_{it}}{\mu_{it}}\right) \mu_t = \delta_{it}\ \mu_t \ \ for\ all\ country, i\ and\ time, t \tag{2}$$

Using equation (2), the common and idiosyncratic factors in the panel can be separated by Phillips and Sul (2007a) through factoring the common stochastic trend component. Equation (2) state that x_{it} is decomposed into two components: common μ_{it} and idiosyncratic δ_{it}. The component δ_{it} it is a measure of distance between x_{it} the common component, μ_{it} . This is absorbing the error term and the unit specific component and hence serves as the idiosyncratic component which is changing over time. The common trend component in the panel denoted by μ_{it} , is assumed to have some deterministic or stochastic trend attitude that influences the transitory component a_{it} as $t \to \infty$. In order to specify the null hypothesis of convergence, the non-stationary transitional nature of factor loading is proposed in semi parametric form, so that each coefficient converges to some unit specific constant:

$$\delta_{it} = \delta_i + \left(\frac{\sigma_i + \xi_{it}}{L\,(t)t^{\alpha}}\right) \tag{3}$$

Where it δ_{it} is fixed, it ξ_{it} is iid (0,1) across i, I σ_i are idiosyncratic scale parameters, slowly varying function is represented by L(t) and L(t)=log t so that $L(t) \to \infty$ as $t \to \infty$. The rates at which the cross-sectional variation decaying to 0 is denoted the parameter α. The formulation above ensures that ξ_{it} it converges to ξ for all $\alpha \geq 0$.

The Transition Path

Estimation of the time varying factor loading δ_{it} it is a central issue of the approach proposed by Phillips and Sul (2007a), since the estimates deliver information about transition behaviour of panel units. So, by applying its corresponding form, a smooth and effective method to obtain fact about the δ it is as regard:

$$h_{it} = \frac{x_{it}}{\frac{1}{N}\sum_{i=1}^{N} x_{it}} = \frac{\delta_{it}}{\frac{1}{N}\sum_{i=1}^{N} \delta_{it}} \tag{4}$$

In equation (4), measure the loading coefficient δ_{it} it in relation to the panel average. For economy i, alike to δ_{it}, h_{it} is still traces out a transition path through present produces one is in association to panel average. Over time, the variable h_{it}, trace out an individual trajectory for each i relative to the average that why it is called as transition path. Together, from the common steady state growth path μ_t of the country i's relevant deviation is as well measure by h_{it}. Thus, any divergences from μ_t are reflected from the transition path h_{it}. By the forming, the average of cross-section of the corresponding path of transition of i equal unity. Addition, the relative transition path hit converges to unity and the cross-sectional variation (Ht) of the relative transition path converges zeroes, if the panel units converge and all the factor loading δ_{it} it approximates to fixed δ, which is as follows:

$$H_t = \frac{1}{N}\sum_{i=1}^{N}(h_{it} - 1)^2 \to 0, t \to \infty \tag{5}$$

These properties are used in testing the convergence null hypothesis. Therefore, the null hypothesis can be determined by $H_0 : \delta_i = \delta$ and $\alpha \geq 0$ for all i, and the alternative hypothesis is $H_0 : \delta_i \neq \delta$ for some i, or $a < 0$. so, in some countries the alternative hypothesis does not show convergence, but in the null hypothesis can indicate convergence for all countries. To analyze convergence concepts, can indicate that the app follows the long-term behaviour in macroeconomic data. Thus, it is usually desirable to eliminate business cycle factor using smoothing technique to obtain hit from X_{it}. So, according by extending (4) to Incorporated a business cycle effect k_{it}, it can be written as:

$$X_{it} = \delta_{it}\,\mu_{it} + k_{it} \tag{6}$$

Due to the adaptability and the points that Hodrick and Prescott (1997) smoothing filter quest simply the addition of smoothing series and bot looking for preceding particularization of the characteristics of the common trend μ_t in X_{it} in this analysis that Hodrick and Prescott (1997) smoothing filter is adopted. That having the computed the HP estimate:

$$\hat{X}_{it} = \hat{\delta}_{it}\,\hat{\mu}_{it} \tag{7}$$

Extending the above, the cross-sectional averages in (4) lead to the estimated transition path computed as:

$$\hat{h}_{it} = \frac{\hat{X}_{it}}{\frac{1}{N}\sum_{i=1}^{N}\hat{N}_{it}} \quad (8)$$

Where $\hat{X}_{it}$ are the filtered income per capita series. Within the expectation, in samples, the panel average $N^{-1}\sum_{i=1}^{N}X_{it}$ is positive also asymptotically that is performed for many related economic time series for instance, price, gross domestic product or other aggregates.

The Log t Test Regression

Base on the time varying the factor in the equation (2) and depending on the log t convergence test that is depending on a simplistic time series regression of Phillips and Sul (2007a) proposed a unique convergence test and clustering algorithm, which involves a one-sided t-test. The test is known as t-test as the t-statistic refers to the coefficient of log t regression in the equation. After estimating the transition path, the variance ratio of cross section H_1/H_i is to be computed by acknowledging X_t as:

$$H_t = \frac{1}{N}\sum_{i=1}^{N}(h_{i1} - 1)^2 \quad (9)$$

The transition distance of Ht has a limiting form as shown below:

$$H_t \sim \frac{A}{L(t)^2\, t^{2a}}, t \to \infty \quad (10)$$

Where A is a positive constant, L(t)=log (t+1) is a slowly varying function and the speed of convergence is α. Usually, after removing a fraction (r) of the sample, in equation (11) is tested. In addition, it is suggested that some point i, become (rT) where (rT) represents the integer art of (rT) and r=0.3. For examining the convergence null hypothesis discussed above, the log t test is carried out as regards by:

$$\log H_1/H_i - 2\log L(t) = \hat{c} + \hat{b}\log t + \hat{\mu}_t \quad t = (rT), \ldots . T \quad (11)$$

Where H_t is the cross-sectional variation H_1/H_i is the ratio of the cross-sectional variation at the beginning of the sample, H1 (i.e Ht at t = 1) over the respective variation for every point in time t, that is Ht (t,…,T). The ratio H_1/H_i measures the distances of the panel from the common limit. At the same time, L(t) = log (t) and r > 0. Thus, the regression presented in equation (6) is known as log t regression due to the log t regressor. The value of $\hat{b}$ is higher or greater, as the rate of convergence is faster. It can be a conditional build up as it tests whether heterogeneous varied idiosyncratic components converge along with the continuous time after controlling growth components within the country. By applying the transitional t-statistic, if t_b < - 1.65 (5% significance level), we reject the H 0 of convergence. It can be concluded panel convergence when the statistics, t_b recommends that $\hat{b}$ is else positive otherwise equal to 0. To reject convergence as a whole, the test procedures applied to subgroups are in accordance with clustering mechanism testing procedures in Phillips and Sul (2007b). On the other hand, we reject H 0 convergence, when the statistics, t_b recommends that $\hat{b}$ negative and significant.

RESULTS AND DISCUSSION

Full Panel Convergence

The productivity convergence of high-tech automotive in Asian countries was analyzed using full panel convergence. The rank of the countries needed to be rank based on higher value and followed by others. Then, by using log t-test, the productivity level was determined by full panel convergence. Finally, this result can be determined either productivity convergence or divergence as shown in Table 1.

Table 1: Full Panel Convergence (Productivity Convergence).

Club	Country	T-Statistic	Remarks
Full Sample	Selected Asian Country	-12.64849**	Divergence

Notes: Asterisk (**) denoted statistically significance at the 5% level. The 5% critical value is -1.65.

The t-statistics of -12.64849 indicating the rejection of the null hypothesis of convergence at the 5% significance level or t-statistic less than -1.65. This is because if the result is more than -1.65 (t_b > -1.65) that means the result are converging but, if the result is less than -1.65 (tb < -1.65) so, that means the result are diverging. Hence, the null hypothesis of full panel convergence is rejected for the period specified from 2002-2016 when the data cropping takes place focusing on the final part of the sample data. This is because, in Phillips and Sul (2007a, 2007b, 2007c) theory, it is believed that the regression empirical log is based on the current data where the first r% or 0.3 data is released prior to analysis. For selected Asian countries, the period is from 2007 to 2016, to reject the null hypothesis for absolute convergence. This is because it can show the difference between a group of developed countries and less developed countries in high-tech automotive industry. Furthermore, these methods are based on factors that vary over time and data do not need to carry out stationary data to meet the problem of sample data stations. This indicates that the entire selected Asian country is heterogeneous in terms of productivity convergence. However, this does not mean that there is no convergence in Asian subgroups that carry further analysis for the period of data that has been cut or ignored 0.3 from the initial data. In the theory of Phillips and Sul (2007c) cointegration and convergence is very relevant, but both have different characteristics and the cointegration test does not work as a test to test the next convergence. Not only that, they also believe that there are many possibilities if full panel convergence is removed. Therefore, the Philips and Sul method are highly adaptable for use in the next part.

Cross Sectional Variation

The transition path productivity convergence in Asian from 2007 to 2016 was illustrated in Figure 1. This is due to avoid the initial impact of the beginning of the base year, the first 5 years of yearly observation out of 5 are discarded and only 10 filtered of observation is being used in this analysis. This is because refer to the r= 0.3.

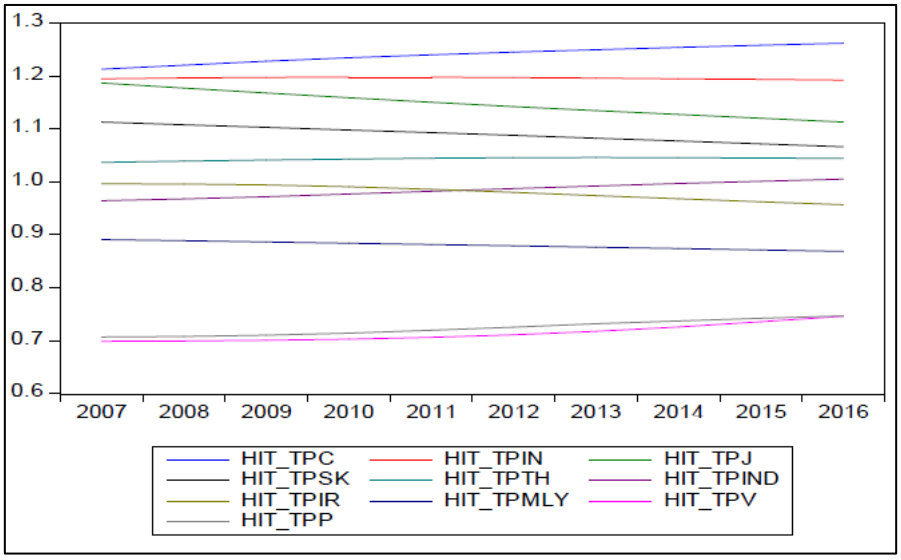

Figure 1: Transition Path for Productivity Convergence in Asian.

The behaviour of convergence productivity is closely related to the panel. The path of transition hit, occupies a growth path for each country and is related to the average sample and can show convergence productivity in the automotive industry that exceeds average cross-section or otherwise. Furthermore, the relevant transition path leads to unity for all countries by assuming convergence across the country panel. In addition, the slope of each curve is represented by the convergence productivity growth rate for the relevant country along with the average cross section. Through Figure 1 shows, a whole panel of different convergence, so it can be inclined to unity the transition path. In Figure 1 shows that there are several countries, crossing or equilibrium points. For example, the country of Indonesia (HIT_TPIND) and Iran (HIT_TPIR) and Vietnam (HIT_TPV) and Philippine (HIT_TPP). These four countries show clearly that have convergence between them. Not only that, looking at the final year in

this data shows that Vietnam can catch up the Philippine country in the coming year as this figure of productivity for Vietnam is increasing with each year. Furthermore, China (HIT_TPC) total productivity is the first ranking in this group. This is because it has the highest total productivity among other countries. In the diagram shows that China is constantly increasing with each year. Not only that, the China country freely moves in the transition path and does not achieve equilibrium with other countries. This shows that a China country is likely to achieve divergence where rejecting the null hypothesis in convergence. While, based on this figure that show Malaysia (HIT_TPMLY) from the beginning around the year 2007 had shown transition parameter highest but slowly goes down in along the period and show the transition parameter is lowest. This show that Malaysia country have achieve convergence in this analysis.

From Table 2, China is the first rank and the base country for this analysis. The second country will be added to run log t regression. Next, computed the t-statistic until the value of the t-statistic is less or more than -1.65. So, continuing this method by adding India to China and get the t-statistic is -24.7503. This result is -24.7503 is less than -1.65 so other country is stop added. Next, India was taken by to be a base country and Japan is added to India while the t-statistic is -8.240443. Because of the value is less than -1.75 so stop added. Japan was taken as a base country and South Korea is added to Japan and the t-statistic is 0.336316. The result is more than -1.65 continuing to add. Then, added Thailand to Japan and South Korea with the t-statistic is 6.337347 is continues added. Indonesia added to Japan, South Korea, and Thailand. The t-statistic is 4.167087 stops added. This is because the value is lower than value 6.337347. There are two rules in club convergence that is less or more than -1.65 and compare with the previous value. If the present value is lower than the previous value, then it stops adding a new country. Indonesia was taken as a base country and Iran added to Indonesia. The t-statistic is -0.746012 continues added. Malaysia added to Indonesia and Iran. The t-statistic is -38.74240 stop added. Malaysia was taken by to be a base-country and Vietnam added to Malaysia. The t-statistic is 0.250618 continues added. Philippines added to Malaysia and Vietnam. The t-statistic is 0.450611 and this result is more than -1.65. The convergence exists in most Asian countries which are the club 1 includes Japan, South Korea, Thailand, club 2 consists of Indonesia and Iran and the club 3 involves of Malaysia, Vietnam and Philippine. For club 1, Japan will be a heading for South Korea and Thailand. While club 2, Indonesia will be a heading for Iran and club 3, Malaysia will be a heading for Vietnam and Philippine. The formation of this clubs has reached the objective of this study where there is clubs' convergence in high-tech automotive industry in Asia. This is determined by comparing the t-statistics given in the critical value results. So, to know whether the subgroup is converging is through a t-statistic value greater than the critical value of -1.65. This study shows that the club 1, club 2, and club 3 are converging. If the critical value is less than the critical value of -1.65 then it is considered as divergence. Thus, the country that divergence are China and India.

Table 2: Results of Clubs Convergence in Asian.

Last T Order	Country	Step 1	Step 2	Step 1	Step 2	Step 1	Step 2	Club	Remarks
1	TPC	Base							Divergence
2	TPIN	-24.75030*	Base						Divergence
3	TPJ		-8.240443	Base				1	Convergence
4	TPSK			0.336313				1	Convergence
5	TPTH			6.337347				1	Convergence
6	TPIND			4.167087*	Base			2	Convergence
7	TPIR				-0.746012			2	Convergence
8	TPMLY				-38.74240*	Base		3	Convergence
9	TPV					0.250618		3	Convergence
10	TPP					0.450611*		3	Convergence

Notes: Asterisk (*) denoted statistically significant at 5% level. The 5% critical value is -1.65
Symbols Represent: TPC-China, TPIN-India, TPJ-Japan, TPSK-South Korea, TPTH-Thailand, TPIND-Indonesia, TPIR-Iran, TPMLY-Malaysia, TPV-Vietnam and TPP-Philippine.

CONCLUSION

Asian has a steady increase in automotive productivity and there is a divergence or no convergence that occurs in the entire panel convergence. However, rich Asian countries have proven to dominate the base group of convergence clubs in high tech automotive driven by automotive wage and production in Asian countries. This study uses Phillips and Sul (2007a) test, to analyze the possibility of convergence attitudes in selected data. Subsequent to the selection of state clusters and to use China (last highest observation) as a benchmark country and can form five different subgroups. In this study it is shown that Malaysia is the lowest convergence in the

group. This is because, there is a shortage of projects in increasing productivity in high tech automotive industries in the country. When lack of improvement in high technology will result in cumulative regional divergence. This is the same as the Krugman (1991) which states that when the lack of improvements in infrastructure will lead to the process of regional deviations. This may indicate that selected Asian countries can catch up with each other.

In conclusion, this study has studied the convergence of productivity of the high-tech automotive industry in Asian countries. Full panel convergence results in this study are diverging but at the convergence club level shows the selected Asian country, there are three clubs that are experiencing convergence and the other two suffer from divergence. In this study have three club convergence, which is club 1 consist of Japan, South Korea and Thailand, club 2 consist to Indonesia and Iran and Club 3 consist of Malaysia, Vietnam and Philippines. Productivity characteristics in the study are to have reasonable wages for employees, having good economic development to increase employee productivity. Secondly, the level of high-tech automotive productivity is to see the skills of workers in the field, productivity development in the country and productivity efficiency in the automotive field.

The automotive industry plays an important role in contributing to improving economic growth. The high-tech automotive industry can benefit the nation by providing high employment opportunities and improving the standard of living, especially for top-class countries like China, Japan and South Korea. Therefore, a divergence state should implement continuous development planning and R&D and innovation. This is because, it can increase exports in the automotive industry and productivity growth of the country. Countries that carry out R&D will continue to gain the edge in improving the high-tech automotive industry. Other than that, Asian countries also need to form regional groups to produce more advanced automotive products. Hence the formation of regional clusters in the high-tech automotive industry will further expand output and be able to accumulate more capital. This group of countries should cooperate with each other in terms of investigation, trade, policy and liberalization. This could create a series of automotive productivity expenditures that can boost economic growth to become more competitive in the international market.

REFERENCES

Bai, J. & Ng, S. (2004). Panic attack on Unit Roots and Coinegration. *Econometrica, 72*(4), *1127-1177.*

Crossman, A. (2017). *What is Convergence Theory?* Retrieved from Thoughtco: https://www.thoughtco.com/convergence-theory-3026158

Ding, R. (2016). *Innovation Efficiency of High-tech Industries in China.* United Kingdom: University of Nottingham. Retrieved February 21, 2018, from https://core.ac.uk/download/pdf/42492866.pdf

Hodrick, R. & Prescott, E. (1997). Postwar Business Cycle: An Empirical Investigation. *Journal of Money, Credit and Banking*, *29*(1), *1-16.*

Krugman, P. (1991). *Geography and Trade.* United States of America: The MIT Press.

Liu, P. L. & Tsai, C. H. (2007). Effect of Knowledge Management Sysrems on Operating Performances: An Empirical Study of High-tech Companies Using the Balanced Scorecard Approach. *International Journal of Management, 24*(4), *734-743.*

Phillips, P. C. B. & Sul, D. (2007a). Transition Modeling and Econometric Convergence. *Econometrica, 75*, *1771-1855.*

Phillips, P. C. B. & Sul, D. (2007b). Some Empirics on Economic Growth Under Heterogeneous Technology. *Journal Macroeconomic, 29*(3), *455-469.*

Phillips, P. C. B. & Sul, D. (2007c). Economic Transition and Growth. *Journal of Applied Econometrics, 24*(7), *1153-1185.*

Trends in Undergraduate Research (2019) 2(2): g8-14
https://doi.org/10.33736/tur.1467.2019

Modelling Tax-Consumption Relationship: An Empirical Evidence from Asia

TAN SHU HUI & MUHAMMAD ASRAF ABDULLAH*

Faculty of Economics and Business, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia.

*Corresponding author: amasraf@unimas.my

ABSTRACT

This study examines the impact of taxation on consumption behaviour by focusing on selected countries from Asia. The study adopts panel fixed effect and random estimators to gauge the influence of tax on consumption expenditures. Findings of the study support negative influence of tax on household consumption. The finding is parallel with the absolute income hypothesis that high volume of tax collected from the public results in falling households' disposable incomes, hence downwardly affecting households' consumption levels. The study's finding implies the importance of carefully observing an appropriate tax policy that suit the country's level of development in order to promote high rates of economic growth and consumption.

Keywords: Tax revenue; Fiscal policy; Government revenues; Household consumption; Public Economics

INTRODUCTION

Over the past decades, household consumption expenditure remained the largest contributor to GDP for many countries in Asia. According to the World Bank (2017), household consumption expenditures of most selected Asian countries have contributed to more than 50% of GDP from 2012 to 2016. There is a high percentage of contribution of consumption to GDP although some of the countries like Japan and Thailand recorded a decreasing trend in household consumption expenditure. According to the World Bank (2018), consumption expenditure forms the largest component of the aggregate demand for some Asian countries. Consumption plays an important role in the economy since it has a significant effect on economic growth of a country (Kim, 2017). Therefore, this study is carried out to further investigate the nexus between tax revenue and households' consumption patterns in the Asian's region.

The significant contribution of consumption expenditure to GDP in Asia has motivated this study to examine the impact of tax on consumption level since most Asian countries rely heavily on income taxes as the main source of government revenue for decades (The World Bank, 2018). According to Keynes (1936), the propensity to consume is a fairly stable function where the aggregate consumption depends on aggregate income of the household. The consumption expenditures of households rise as their income levels increase. Tax may affect the disposable incomes of households and their purchasing power, thus negatively affecting their aggregate consumption level. A lower tax rate increases the disposable income and spending power of consumers, thus stimulating consumption expenditure as consumers have more money to spend and vice versa. With a lower tax rate, consumption expenditure of household would rise, resulting in an increase in the aggregate demand which promotes a higher economic growth and increases social welfare of a country (Pettinger, 2017).

In recent years, the relationship between tax and consumption expenditure had received growing attention among researchers. Some of the studies were highlighted by Ajibola and Segun (2017), Leka (2015), and Eusebio and Franchesca (2013) who suggested a positive impact of tax on households' consumption expenditure. Other group of researchers namely Varlamova and Larionova (2015), Tamasauskiene and Opulskyte (2012), Baker and Kueng (2016), and Kolahi *et al.* (2016) emphasised on negative relationships between tax and consumption expenditure. These mixed

results suggested the need to reinvestigate the issue of tax-consumption relationship in order to suggest appropriate policy implication to the policy makers.

Tax is a compulsory contribution to state revenue. It is levied by the government on individual income and organization profits, or added to the cost of some goods, services, and certain activities (Hyman, 2011). Taxes affect both consumers and producers as it has the tendency to change both demands and supplies of goods and services through changes in prices, hence, influencing resource use patterns (Hyman, 2014). Tax affects economic growth via changes in households' consumption behavior (Ebiringa & Yadirichukwu, 2012; Unegbu & Irefin, 2011; William & Benjamin, 2010). Therefore, this study focuses on the impact of tax from consumer's context.

The general idea of investigating the relationship between tax and consumption was based on the absolute income hypothesis. This hypothesis was proposed by Keynes (1936) to identify the relationship between disposable income and consumption in his general theory. Keynes asserted that real consumption is a function of real disposable income that is net income after deducting taxes. Consumption will decrease with the rise in tax that reduces the households' disposable income and vice versa. The simple Keynesian consumption function is written as follows:

$$C_t = \alpha + \beta\,(Y_d) \qquad (1)$$

Where,
C_t = Consumption at the time t
α = Autonomous consumption
β = Marginal propensity to consume
Y_d = Disposable income at time t

According to simple Keynesian model in equation 1, marginal propensity to consume (MPC) is the ratio between the change in consumption expenditure and the change in disposable income which is income after taxes paid. It measures the change in consumption expenditure as a result of changes in every unit of disposable income. The MPC is assumed to be positive and constant while the autonomous consumption is expected to be positive and small.

This hypothesis postulates some important characteristics. These characteristics are including non-proportionality where consumption expenditure increases or decreases with the rise or the fall in income. The non-proportional consumption function stated that MPC is smaller than average propensity to consume ($MPC < APC$) in the short run due to non-changing autonomous consumption function with income in the short run. However, MPC is equal to APC in the long run. This is because there is a proportional relationship between income and consumption.

The Keynesian consumption function is written as $C = A + B\,(Y^d)$. Where C is the consumption expenditure, A represents autonomous spending, B is the marginal propensity to consume, and Y^d represents real disposable income. Theoretically, tax has a significant and negative impact on the disposable income of household, thus downwardly affecting household consumption expenditure.

There is a growing body of literature that examine the association between tax and consumption focusing on different regions and countries across the globe over the past decades. Among these literatures are that associated to investigation of relationship between varying taxes types and household consumption in different countries settings namely countries of the OECD including United States, Europe and Asia. Most of the research findings suggest negative impact of different types of taxes on household consumption while others opposed.

Among the strand of research that support positive impact of taxes on consumption are highlighted by Eusebio and Franchesca (2013), Leka (2015) and Ajibola and Segun (2017). The study's findings suggest that tax is one of the factors that influence consumption expenditure of household. In a way, these research suggests that a rise in net tax by one percent will increase the households' consumption expenditures.

Eusebio and Franchesca (2013) conducted a study to investigate the impact of macroeconomic variables on consumption expenditure in the Philippines using quarterly time series data based on OLS and Maximum Likelihood estimation techniques over the period 1947 to 2007. Next, by employing annual time series regression analysis for the period from 1985 to 2013 using sample data from Albania, Leka (2015) extends the study of relationship between tax and consumption. Subsequently, the study by Ajibola and Segun (2017) who applies the OLS regression technique

on annual economic dataset from Nigeria for the period 1994 to 2014 shows positive influences of value added tax and company income tax on household consumption expenditure. The study's results postulate that an increase in the rates of indirect taxes is associated to the increase in household consumption expenditure.

In contrast to above studies, Romer and Romer (2010), Alm and Ganainy (2012), Tamasauskiene and Opulskyte (2012), Varlamova and Larionova (2015), Strilkova and Siroky (2015), Baker and Kueng (2016), Kolahi *et al*. (2016), and Bartkus (2017) who apply panel data estimation approach argue negative relationships between tax and consumption expenditure of household. Applying the panel OLS to analyse the sample countries from OECD, Varlamova and Larionova (2015) revealed negative relationship between taxes and household expenditure. Subsequently, by extending the panel estimation to dynamic panel data based on generalized method of moments (GMM) system estimators to different countries set, Alm and Ganainy (2012) and Kolahi *et al.* (2016) confirm that VAT (value added tax) has a significant negative effect on consumption. An increased VAT would result in a reduction of consumption. Similarly, another study done by Tamasauskiene and Opulskyte (2012) using multiple regression analysis also portrays a negative relationship between indirect taxed and household consumption expenditure.

Another branch of literature leading by Strilkova and Siroky (2015), Baker and Kueng (2016), and Romer and Romer (2010) who employed time series data in their analyses also support a negative effect of tax on consumption. The finding by Strilkova and Siroky (2015) indicates that lower VAT rate would result in a higher net income, hence pushing the consumer expenditures up. The study collects data using questionnaires and adopts analytic-synthetic and comparative methods to analyze the sample data from Czech Republic. Baker and Kueng (2016) and Romer and Romer (2010) observe a strong correlation between household spending and changes in taxes as consumers tend to change their consumption pattern following an increase in tax rate in the United States. Another literature by Bartkus (2017) who applies time series analysis using unit root, VAR and Granger causality also convey a negative consumption impact resulting from high tax rate in Lithuania.

Meanwhile, another strand of literature for instance, Miki (2011), Tochukwu *et al.* (2015), and Hussain and Malik (2016) confirm mixed results with some literature suggest a positive nexus while others indicate an opposite nexus between the tax and consumption expenditure. In the studies by Hussain and Malik (2016) and Tochukwu *et al.* (2015) who apply different estimation approaches where the former adopts censored regression model while the later applies OLS estimation technique arrive at similar conclusion of a positive nexus between tax and household consumption expenditure on non-durable goods and negative impact of the tax and household consumption expenditure on durable goods.

Miki (2011) who investigates the effect of the VAT rate on aggregate consumption in 14 developing countries using fixed and random effect estimation method supports similar mixed result. The sample data for the study are separated into two groups, one group consists of data before implementation and the other consists of data after implementation of VAT. The finding for first group of data implies a positive tax-aggregate consumption relationship. Next finding for second group indicates a negative effect of VAT reduction on aggregate consumption.

The above review of tax impact on consumption highlighted mixed results where the tax and consumption expenditure may have a positive, a negative or both negative and positive nexus across different settings of studies. These inconsistent findings signal the need for re-estimation of the relationship between tax and consumption by focusing on different context of study.

This paper is dissected into five sections. Section one discusses the motivation of the study by emphasizing on the significance contribution of consumption expenditure on GDP. The section also discusses how tax influences disposable incomes which finally changes the consumption expenditure of households. In addition, this section explicates the conventional theory of consumption based on Keynes's (1936) view. In discussing the consumption theory, review of selected past literatures is carried out in order to explain why tax and consumption relationship merits further investigation. Section two explores the data and methodology that are applied in the study's analysis. The results and discussion of the study are presented in section four. Lastly, section five provides the conclusion of the study.

MATERIALS & METHODS

This study uses panel data that consists a sample of seven countries for five years spanning between 2012 and 2016. The study period is chosen to avoid fluctuations in economic data due economic crisis. The study employs annual data with a total of 35 observations. The independent variable of our interest for this study is tax. Tax variable in our context of study comprises of total tax revenues including income tax, consumption tax and indirect tax revenues. It can affect the distribution of income and wealth, mode of consumption as well as the pattern of production. Tax data in this study is proxy by total tax revenue collected by the government for a particular year. In this study, tax is taken as the percentage of GDP.

Meanwhile, GDP per capita and gross savings are served as the control variables as commonly found in other literature regarding the relationship between tax and consumption. GDP per capita is an index that measured the degree of total economic output relative to the population of a nation in a particular year. It is calculated by dividing GDP at the current market prices with the number of population of a country before adjusting for inflation (The United Nations, 2017a). GDP per capita is a common indicator of the economic well-being of a citizen. Theoretically, the nexus between GDP per capita and household consumption expenditure should be positive. A rise in GDP per capita results in an increase of purchasing power of household.

Next, gross savings is an indicator that measures the level of resources or funds that are available for investment in capital asset (The United Nations, 2017b). Savings could contribute to sustainable future economic growth by increasing the amount of capital available for investment. Gross savings in this context of study are calculated by gross disposable income less consumption. Saving has a negative relationship with consumption. On the other hand, household consumption expenditure is the dependent variable for this study.

In this study, consumption is proxied by household consumption expenditure since it consists of the spending on goods and services that are incurred by households. Data of household consumption expenditure in this study is measured by taking the ratio of consumption to GDP. All the raw data for the variables are standardized in USD before conversion into percentage or ratio. All the data for this study are procured from the World Bank.

This study's analysis adopts panel random effect estimator based the studies by Burge and Rogers (2011) and Miki (2011) with arguments that the models suit short time periods (small *T*) and more countries (large *N*). According to Williams (2018), the random effects estimator is more efficient than the pooled OLS because the errors are classical. Nevertheless, it is more efficient compared with fixed effect estimator as it does not use a large number of dummy variable and loss of the degrees of freedom can be avoided if the μ_i can be assumed random. Furthermore, panel data is applied in this study since it provides rich data with both time series and cross section dimensions that account for heterogeneity effect of countries. In order to recognize the best panel estimator for analysis of the study, two tests have been applied to identify suitability of pooled OLS, random effect and fixed effect models namely Breusch-Pagan Lagrange Multiplier (LM) test and the Hausman test. The relationship between tax and household consumption expenditure is modeled as follows:

$$HCE_{it} = \beta_0 - \beta_1(T_{it}) - \beta_2 log(GDPPC_{it}) - \beta_3(GS_{it}) + \varepsilon_{it} \quad (2)$$

where,
HCE = Household consumption expenditure
T = Tax
GDPPC = Gross Domestic Product per capita
GS = Gross savings
β_0 = Constant
$\beta_1, \beta_2, \beta_3$ = Coefficients of independent variables
ε = Error term

The expected impact of tax on household consumption expenditure is negative according to the Keynesian absolute income hypothesis. Nonetheless, there might be a positive impact of tax on consumption if the types of consumption expenditure are different. The consumption expenditure of household is separated into three broad categories which are expenditures for durable goods, non-durable goods, and services. The tax might reduce the consumption of durable good as durable goods are defined as the goods those whose expected lifetime is more than three years. Meanwhile,

there are positive nexus between the tax and household consumption expenditure on non-durable goods as non-durable goods are something that can be expected to last for a substantial time period (Hussain & Malik, 2016; and Tochukwu *et al.*, 2015). Two economics variables namely GDP per capita and savings which have incredible influence on consumption according to theory are introduced to ensure correct specification of the model. GDP per capita should have a negative association with consumption because GDP per capita represents household's income. In other word, it means that an increase in GDP per capita would results in rising household income. The increase in household income would suggest more purchasing power for households, thus raising the household consumption expenditure. On the other hand, the relationship between gross savings and household consumption expenditure is anticipated to be negative. Savings is defined by the difference between a person's consumption expenditure and disposable income. Savings are expected to reduce as the consumption increase with the assumption that income is constant in the short run.

RESULTS AND DISCUSSION

The static panel data regression requires two basic tests to test for suitability between pooled OLS and random effect and fixed effect models. This econometric requirement is satisfied by conducting the Breusch-Pagan LM test. The result of the Breusch-Pagan LM test is depicted in Table 1. The probability statistics obtained from the test is 0.0038, which is smaller than significance level of 1% or 0.01. Hence, we have enough statistical evidence to reject the null hypothesis at 1% significant level. We conclude the presence of individual specific effect in the model. In other words, the random effect model is more appropriate than the pooled OLS model. This is because the individually specific effects are completely ignored in the pooled OLS model since it is a simple OLS technique of panel data.

Table 1: Result of Breusch-Pagan Lagrangian Multiplier Test.

Chi2(1)	Prob>Chi(2)
8.39	0.0038***

Notes: The dependent variable is household consumption expenditure. Independent variables include tax, GDP per capita, and gross savings. Asterisks (***) indicate statistically significant at 1% significance level.

Next, to treat for country specific effects we proceed with the validation of suitability of fixed effect and random effect models to be used for estimating tax impact on consumption. The validation of correct model is done by employing the Hausman test to test for suitability between random effect and fixed effect estimators. The result of the Hausman test is reported in Table 2. As indicated in the results from Table 2, the Hausman value which is given by the Chi Square statistics is 5. The small Hausman value couple with the large probability value of 0.1719 suggests that we accept the null hypothesis which indicates no correlation between country specific effect and the independent variables. Therefore, we conclude that the random effect is consistent and efficient. In other words, the random effect model is more appropriate than the fixed effects model to estimate the impact of tax on consumption in this study's context. In this study, the fixed effect model has too many parameters, thus results in an enormous loss of degree of freedom. Hence, the individual effects are categorized as random. Besides, random effects model allows for modelling the heterogeneity across units since there are different individual effects in this study but fixed effect model only makes an assumption of homogeneity. Therefore, the random effect model is more appropriate than the fixed effects model in this study.

Table 2: Result of Hausman Test

Test Summary	Chi-Square Statistics	Probability
Cross-section random	5.00	0.1719

Notes: The dependent variable is household consumption expenditure. Independent variables include tax, GDP per capita, and gross savings.

The impact of tax on household consumption expenditure in selected Asian countries based on the random effect model analysis is demonstrated in Table 3. The results indicate that tax is statistically significant at 1% significance level since the p-value of tax is 0.004 which is less than 0.01. Furthermore, the t-statistic value is sufficiently high at

2.87 which suggests that the tax has a highly significant and negative influence on consumption. This finding is parallel to findings of Romer and Romer (2010), Burge and Rogers (2011), Alm and Ganainy (2012), Bartkus (2017), Varlamova and Larionova (2015), Leka (2015), Strilkova and Siroky (2015), Baker and Kueng (2016), and Kolahi *et al.* (2016), who propose a negative impact of tax on the household consumption expenditure. The study's finding is in line with absolute income hypothesis's proposition that an increase in tax results in a reduction of income and household consumption expenditure.

Result of control variable, GDP per capita is statistically significant at 10% significance level with the p-value of 0.053 and moderate t-statistic value of 1.93. This values indicate that GDP per capita has a moderately significant influence on the household consumption expenditure. Nonetheless, the sign of the GDP per capita was negative which is in line with our prior expectation of negative relationship between GDP per capita and consumption. This finding is support by Leka (2015). The increase in the rate of national savings for Asian country results in a decline in the household consumption expenditure despite an increase in GDP per capita.

As for savings variable, the p-value is 0.000 which is less than 0.01 coupled with a strong t-statistic value of 16.46 implies a significant and negative impact of saving on household consumption expenditure. This study's finding is supported by Leka (2015) that a reduction in gross savings results in an increase in household consumption expenditure.

Table 3: Random Effect Model Result

Variables	Coefficient	Standard Error	T-statistic
T	-0.7652***	0.2663	-2.87
LGDPPC	-1.0290*	0.5329	-1.93
GS	-0.7911***	0.0481	-16.46

Notes: The dependent variable is household consumption expenditure. Independent variables include tax, GDP per capita, and gross savings. Asterisks (***, *) indicate statistically significant at 1% and 10% significance level.

CONCLUSION

This research explores the impact of tax on household consumption expenditure in selected Asian countries. A data set consists of selected seven countries in Asia namely Malaysia, Singapore, Indonesia, Thailand, Japan, Bhutan, and South Korea covering the period 2012 to 2016 is employed in analyzing the tax-consumption relationship. The data set is analysed using panel random effect model. The findings from this study postulates a negative influence of tax on household consumption expenditure. This finding is in agreement with the absolute income hypothesis that a high tax rate results in falling income and consumption levels. In addition, there is a negative impact of GDP per capita and gross savings on household consumption expenditure in selected Asian countries.

A high rate of tax negatively affects the household consumption expenditure through falling disposable income. The study's result suggests that policy makers should implement a suitable tax system if the country would like to foresee a high economic growth in the future as reasonable tax rate would encourage consumer spending. This objective can be accomplished through a favourable fiscal policy which encourage positive growth of business and consumption. There is a need for the policy makers to review existing fiscal policy that encourage business activities and consumption whilst at the same time maintaining sustainable public revenues to implement government projects and policies.

The government may identify other sources of revenues if tax rates were to reduce. Another alternative to ensure sufficient public revenues is to practice prudent spending and prioritizing projects that have the potential to uplift the rate of economic growth of a country. These practices would ensure sustainability of the country's debt.

ACKNOWLEDGEMENTS
Constructive comments on this paper by the reviewers are gratefully acknowledged.

REFERENCES

Alm, J., & Ganainy, A. E. (2012). Value-added taxation and consumption. *International Tax and Public Finance*, *20*, *105-128*.

Ajibola, J. O., & Segun, O. P. (2017). Taxation and its influence on Household consumption: The Nigerian experience. *International Journal of Economics and Business Management*, *3*(2), *108-128*.

Bartkus, A. (2017). The impact of taxes on the consumption to income ratio. *Ekonomika, 96*(2), *7-26.*

Baker, S, R., & Kueng, L. (2016). *Shopping for lower sales tax rates* (NBER Paper No. 23665).

Burge, G., & Rogers, C. (2011). Local option sales taxes and consumer spending patterns: fiscal interdependence under multi-tiered local taxation. *Journal Regional Science and Urban Economics, 41, 46-58.*

Ebiringa, O, T., & Yadirichukwu, E. (2012). Analysis of tax formation and impact on economic growth in Nigeria. *International Journal of Accounting and Finance Reporting, 2*(2), *367-385.*

Eusebio & Franchesca, N. M. (2013). *A regression analysis on the relationship of final consumption expenditure (US Dollars) with GDP per capita, Gross savings, and net taxes on products (1978-2010).* Retrieved from https://www.academia.edu/4428396/A_Regression_Analysis_on_the_Relationship_of_Final_Consumption_Expenditure_US_Dollars_with_GDP_per_Capita_Gross_Savings_and_Net_Taxes_on_Products_1978-2010_

Hussain, S. M., & Malik, S. (2016). Asymmetric effects of exogenous tax changes. *Journal of Economic Dynamics & Control, 69*, *268-300.*

Hyman, D, N. (2011). *Public Finance: A contemporary application of theory to policy*. Tenth edition. USA: Cengage Learning.

Hyman, D, N. (2014). *Public Finance: A contemporary application of theory to policy*. Eleventh edition. USA: Cengage Learning.

Keynes, J. M. (1936). *The general theory of employment, interest and money*. New York: Harcourt, Brace.

Kim, H. (2017). The effect of consumption on economic growth in Asia. *Journal of Global Economics, 5*(3).

Kolahi, S. H. G., Noor, Z. B. M., & Kashmari, A. (2016). Effects of value add tax on consumption in developing countries. *Applied Economics and Finance. 3*(2), *186-193.*

Leka, R. (2015). *Final consumption expenditure in Albania*. Retrieved from http://fdut.edu.al/wp-content/uploads/2015/11/Rovena-Leka-relationship-of-final-cons.expenditure-with-GDP-per-capita-gross-savings-and-net-taxes-on-product.1985-2013.pdf

Miki, B. (2011). *The effect of the VAT rate change on aggregate consumption and economic growth*. Columbia University, Centre of Japanese Economy and Business (Paper No. 297). Working Paper Series

Pettinger, T. (2017). *The effect of tax cuts*. Retrieved from https://www.economicshelp.org/blog/13566/economics/the-effect-of-tax-cuts/

Romer, C. D., & Romer, D. H. (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *The American Economic Review, 100, 763-801.*

Strilkova, R., & Siroky, J. (2015). Changes in the VAT burden on expenses of selected households in the Czech Republic. *Law and Economics Review, 6*(3), *189-202.*

Tamasauskience, Z., & Opulskyte, R. (2012). Assessment of impacts of changes in indirect taxes on macroeconomic indicators of Lithuania in context of other EU member States. *Social Research*, *2*(27), *5-18.*

The United Nations (2017a). Gross Domestic Product per capita. Retrieved from http://www.un.org/esa/sustdev/natlinfo/ indicators/methodology_sheets/econ_development/gdp_percapita.pdf

The United Nations (2017b). Gross Savings. Retrieved from http://www.un.org/esa/sustdev/natlinfo/indicators/methodologysheets/econdevelopment/grosssaving.pdf

The World Bank (2017). *World Development Indicators*. Retrieved from http://databank.worldbank.org/data/reports.aspx?source=2&country

The World Bank (2018). *World Development indicator*s. Retrieved from http://databank.worldbank.org/data/source/world-development-indicators/preview/on

Tochukwu, O. R., Jerry, K., & Titus, O. A. (2015). Value added tax and consumption expenditure behaviour of households in Nigeria: An empirical investigation. *International Review of Social Sciences, 3*(6), *236-248.*

Unegbu, A. O. & Irefin, D. (2011). Impact of VAT on economic development of emerging nations. *Journal of Economics and International finance, 3*(8), *492-503.*

Varlamova, J., & Larionova, N. (2015). Macroeconomic and demographic determinants of household expenditures in OECD countries. *Procedia Economics and Finance, 24*, *727 – 733.*

William, G. G. & Benjamin, H. H. (2010). *A value-added tax for the United States: Part of the solution*. Retrieved from https://www.brookings.edu/research/a-value-added-tax-for-the-united-states-part-of-the-solution/

Williams, R. (2018). *Panel Data 4: Fixed effects VS random effects models*. Retrieved from https://www3.nd.edu/~rwilliam/stats3/panel04-fixedvsrandom.pdf