SKIN HYDRATION AND TRANSEPIDERMAL WATER LOSS MEASUREMENT USING VIS/NIR SPECTROSCOPY AND FEED-FORWARD BACKPROPAGATION NEURAL NETWORK

TAN CHUN HO

UNIVERSITI SAINS MALAYSIA

2022

SKIN HYDRATION AND TRANSEPIDERMAL WATER LOSS MEASUREMENT USING VIS/NIR SPECTROSCOPY AND FEED-FORWARD BACKPROPAGATION NEURAL NETWORK

by

TAN CHUN HO

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

June 2022

ACKNOWLEDGEMENT

First and foremost, I would like to appreciate and acknowledge all the related parties or individuals for their help and advice to finish this study. I would like to express my deepest gratitude to both of my supervisors, Professor Mohd. Zubir Mat Jafri and Associate Professor Dr. Ahmad Fairuz Omar. Their guidance and support have helped me especially during the most challenging times. Thanks to their encouragement, attention and patience that has contributed to the success of this research. Their constant motivation and support have promoted my dedication to solve all obstacles and challenges occurred in my experimental works and thesis writing.

My sincere appreciation to both of my beloved parents for their unconditional support and encouragement has given me strength to complete this project in Universiti Sains Malaysia. Their love and support catalyse and encouraged me to pursue my PhD' study all the way through the project.

I would like to take this opportunity to thank the lecturers and lab assistants at School of Physics, Universiti Sains Malaysia and the staffs and workers from Top Glove R&D Centre (F25), Sky Resources Sdn. Bhd, especially for their help and cooperative, who participates in this project directly or indirectly from initial till the end of the project. In addition, this project was financially supported by the Top Glove Sdn. Bhd., Malaysia [grant numbers 304/PFIZIK/650956/T142]; Sky Resources Sdn. Bhd., Malaysia, USM bridging research grant [number 304/PFIZIK/6316173]; human ethics consent number JEPeM17010005, and MyBrain15. Their help and participation are very essential to conduct the project smoothly and able to finish this study.

Finally, I am so grateful for my lab mates, friends, and lecturers for their moral support while we were conducting our research together. I want to thank them for their invaluable experience and exchanging their knowledge with me throughout the project. Besides, it is a tough time when the COVID-19 pandemic hits Malaysia, and I am so thankful for the helps from Stephenie, Ben, Beh, Syahidah, and others when I am under self-quarantine.

TAN CHUN HO
2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS

$\boldsymbol{A}$	Collection of the a before it is summed together and activated
a	Elements in matrix $\boldsymbol{A}$
$\boldsymbol{B}$	Collection of biases for the current layer
b	Bias
c	Constant that controls the decay rate
f	Upper limit of the activation function, which is 1 for hyperbolic tangent
g	Lower limit of the activation function, which is -1 for hyperbolic tangent
m	Position of node in vertical direction
N	Layer's numbers
n	Position of previous node in vertical direction
S	Sample size
$\boldsymbol{W}$	Collection of weights for the current layer
w	Weight
$\boldsymbol{X}$	Collection of normalized input, or the values in the previous node
x	Data
y	The normalized predicted output
z	The summation of every element in $\boldsymbol{A}$
Δ	Gradient of the *SSR*
η	Learning rate
γ	Momentum
α	Step size
θ	Phase angle

LIST OF ABBREVIATIONS

AI	Artificial intelligence
ATR-FTIR	Attenuated total reflectance Fourier transform infrared spectroscopy
BPNN	Backpropagation neural network
CNN	Convolutional neural network
DT	Decision tree
EEMCO	European Group on Efficacy Measurement and Evaluation of Cosmetics and other Products
FFBPNN	Feed-forward backpropagation neural network
FFNN	Feed-forward neural network
FTIR	Fourier transform infrared spectroscopy
KNN	K-nearest neighbour
LDA	Linear discriminant analysis
LR	Logistic regression
NB	Gaussian naive byes classifier
NFC	Near field communication
NMF	Natural moisturizing factor
NIR	Near infrared
RMSE	Root mean square error
RNN	Recurrent neural network
S	Standard error
SSR	Sum of squared residual
SVC	Support vector machine based classifier
SWIR	Short-wave infrared
TEWL	Transepidermal water loss
UV	Ultraviolet

VIS	Visible
µS	microSiemens

LIST OF APPENDICES

PENGUKURAN PENGHIDRATAN KULIT DAN KEHILANGAN AIR TRANSEPIDERMA MENGGUNAKAN SPEKTROSKOPI VIS/NIR DAN RANGKAIAN NEURO RAMBATAN BELAKANG SUAPAN-HADAPAN

ABSTRAK

Kelembapan kulit dan kehilangan air pada transepidermal (TEWL) biasanya boleh diukur dengan menggunakan kaedah elektrik. Namun begitu, kaedah elektrik boleh dipengaruhi oleh pelbagai faktor seperti kakisan dan pengoksidaan elektrod, suhu dan kelembapan bilik, dan kepekatan ion di bawah kulit. Oleh itu, kajian ini menggunakan spektroskopi cahaya tampak/inframerah jarak dekat (VIS/NIR), khususnya pada panjang gelombang 698 nm, 970 nm, 1200 nm, 1450 nm, dan 1950 nm untuk meramalkan kelembapan kulit dan TEWL. Dalam kajian ini, model ramalan dibina dengan menggunakan analisis regresi linear, analisis regresi berganda, dan Rangkaian Neuro Rambatan Belakang Suapan-Hadapan (FFBPNN) dengan algoritma penurunan gradien. FFBPNN yang digunakan dalam kajian ini terdiri daripada satu lapisan masukan, tiga lapisan tersembunyi (setiap lapisan tersembunyi mempunyai lima neuron), dan satu lapisan luaran. Kelembapan kulit, TEWL, dan specktrum kulit diambil dari tapak tangan and lengan subjek. Kajian ini merangkumi dua puluh empat subjek yang sehat tanpa penyakit kulit and berumur dari 22 hingga 32 tahun. Selain itu, subjek dipilih secara rawak and terdiri daripada calon dengan pelbagai etnik, warna kulit, dan jantina. Kajian ini dijalankan dalam bilik berhawa dingin dan suhu bilik dikawal dengan menggunakan penghawa dingin. Tetapan penghawa dingin ditetapkan pada suhu 24 °C dan tetapan ini kekal sepanjang proses pengambilan data. Kemudiannya, data dibahagikan kepada tiga set secara rawak; set pertama digunakan untuk melatih FFBPNN, set kedua and set ketiga digunakan untuk pemilihan dan pengesahan modal.

Set data yang sama juga digunakan untuk menbina dan mengesahkan model regresi. R^2 bagi model FFBPNN adalah dari 79.1% hingga 89.1% dan 41.0% hingga 68.0% untuk kelembapan kulit dan TEWL masing-masing. Manakala, R^2 bagi model regresi berganda adalah 24.8% dan 19.8% untuk kelembapan kulit dan TEWL masing-masing. Selain itu, kajian ini juga mendapati bahawa hubungan antara kelembapan kulit dan TEWL mempunyai korelasi positif R^2 = 47.51%. Keputusan ini mempunyai kesamaan dengan kajian lain yang dapat korelasi positif, iaitu cahaya matahari yang mencukupi, dan persekitran yang panas dan lembap. Kesimpulannya, FFBPNN dan spektroskopi VIS/NIR dapat meramalkan kelembapan kulit dan TEWL dengan ketepatan yang tinggi. Kajian ini juga menunjukkan potensi rangkaian neuro dan spektroskopi untuk aplikasi analisis kuantitatif.

SKIN HYDRATION AND TRANSEPIDERMAL WATER LOSS MEASUREMENT USING VIS/NIR SPECTROSCOPY AND FEED-FORWARD BACKPROPAGATION NEURAL NETWORK

ABSTRACT

Skin hydration and transepidermal water loss (TEWL) are commonly measured using electrical method. However, electrical method can be affected by numerous factors such as the physical condition of the probe due to corrosion and oxidation, ambient temperature and humidity, and the underlying ions concentration under the skin. Thus, this study proposed the use of visible/near-infrared (VIS/NIR) spectroscopy, specifically at wavelengths 698 nm, 970 nm, 1200 nm, 1450 nm, and 1950 nm to predict the skin hydration and TEWL. In this study, the prediction models were built using simple linear regression, multiple regression analysis, and feed-forward backpropagation neural network (FFBPNN) with gradient descent algorithm. The FFBPNN consists of an input layer, three hidden layers with five hidden nodes per each hidden layer, and an output layer. Twenty four participants were recruited for this study; their skin hydration, TEWL, and absorbance spectra were collected from the participants' palms and arms. The participants were comprised of various skin colours, ethnics, and genders, with age ranges from 22 to 32 years old. The participants were selected at random with the main criteria that the healthy subjects were without any skin diseases. The experiment was carried out in an air-conditioned room under controlled temperature. The temperature of the air conditioner setting was set at 24 °C and the settings remain constant throughout the data collection. The data were then randomly sorted into three datasets; the first set was used for the training of FFBPNN, while the second and third dataset were used for the model selection and validation test

respectively. The same datasets were also used to create and validate the regression models. The R^2 for the FFBPNN models ranged from 79.1% to 89.1% and 41.0% to 68.0% for skin hydration and TEWL respectively. While the R^2 for multiple regression models are 24.8% and 19.8% for skin hydration and TEWL respectively. Besides, this study also found that the skin hydration and TEWL have positive correlation of R^2 = 47.51%. This result shared a similarity with other authors that found the positive correlation tends to occur in high availability of sunlight, hot and humid environment. In conclusion, FFBPNN successfully predict skin hydration and TEWL using the VIS/NIR spectroscopy and outperforms the classical multiple regression models. This study also demonstrates the combined potential of spectroscopy and neural network for quantitative analysis application.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Skin is the biggest and one of the most important organs that function to protect us from most of the environmental hazards. Well hydrated skin is more resilient against skin diseases and environmental hazards. Besides, a well hydrated skin also looks more healthy, smooth, and elastic from the cosmetic aspect. On the contrary, overhydrated and dehydrated skin cause unpleasant sensation and skin diseases, which compromise the protection provided by the skin. Skin hydration and transepidermal water loss (TEWL) measurement often used by researchers and dermatologists to evaluate the degree of skin damage and the effectiveness of a skin treatment. The most common technique for skin hydration and TEWL measurement by measuring the electrical properties of the skin such as impedance, capacitance, conductance, and resistance of the skin. However, this technique is not perfect and do contains some weaknesses. Thus, researcher are looking for an alternative way to interpret skin hydration and TEWL for the past few decades.

One of the promising techniques is spectroscopy, that is the study of the interaction of light and matter. Recently, spectroscopy is widely used in many areas such as environmental study, pollution monitoring, food quality control, and health care. With the advancement of computing power in recent years, machine learning is becoming more feasible and integrated into many areas. One of the most widely used algorithm is the feed-forward backpropagation neural network (FFBPNN).

In this study, skin hydration and TEWL is measured using DermaLab USB Probe, and skin's absorbance spectrum is collected using ASD FS3 Spectrometer. The relationship between skin hydration, TEWL, and the selected water absorbance wavelengths is studied using statistical software. Finally, the prediction models for skin hydration and TEWL are constructed based on simple regression analysis, multiple regression analysis and FFBPNN.

1.2 Problems Statement

In most cases, skin hydration and TEWL reveal the skin condition, thus, they are considered as important parameters during skin health diagnostic. Skin hydration is the water holding capacity of the skin, while TEWL is the rate of water evaporated from the skin. Skin hydration and TEWL can be used to determine the severity of skin diseases and to monitor the improvement or deterioration of diseased skin as well as the effectiveness of treatment (Augustin *et al.*, 2019; Montero-Vilchez *et al.*, 2021). Furthermore, these parameters can also be used to indicate skin damage or occupational skin diseases at early stage (Kasemsarn *et al.*, 2016; Korinth *et al.*, 2005). Skin hydration is commonly measured based on electrical method, which is by measuring the electrical properties of skin. The electrical properties of skin, such as resistance, capacitance, conductance, and impedance can be used to indicate the level of skin hydration, while the TEWL measurement is based on evaporimeter that measures the water vapor gradient (Grove *et al.*, 1999; Martinsen & Grimnes, 1998). However, these methods can be affected by several factors such as the corrosion and oxidation of electrodes, ambient temperature, humidity, and the underlying ions concentration under the skin (Barel & Clarys, 1995; Barel & Clarys, 2014; Krishnan

et al., 2017). Thus, there are many ongoing research that proposed different alternative methods for skin hydration and TEWL measurement in recent decades.

For instance, raman spectroscopy was used by Biniek *et al.* to investigate the biomechanical and structure of drying skin in year 2018. A low cost battery-free skin hydration sensor that used near field communication technology was developed by Madhvapathy *et al.* in year 2020. Two multi-wavelengths optical sensors were developed by Mamouei *et al.* to measure the water content underneath porcine (pig) skin in year 2020. Lastly, one of the latest research in year 2021, Yang *et al.* used diffused reflectance spectroscopy (wavelengths 945 nm – 1000 nm and 1230 – 1380 nm) and artificial neural network to investigate the water bonding status of normal and psoriatic skin.

Visible/near-infrared (VIS/NIR) spectroscopy method has been proposed for skin health assessment and water content assessment because the water molecule has an unique spectra that can be used for detection, identification, and measurement. There are studies that use visible/near-infrared/short-wave infrared (VIS/NIR/SWIR) spectroscopy to study the water content inside the leaf (Raj *et al.*, 2021), the quality of fruits and vegetables (Beghi *et al.*, 2018; Blakey *et al.*, 2009; Magwaza *et al.*, 2012), fish fillet (ElMasry & Wold, 2008) and non-biological samples (Knadel *et al.*, 2020; Krauklis *et al.*, 2018).

Recently, the advancement of computer sciences has enabled machine learning to be more feasible and applicable nowadays. This allows researchers to have more options and versatility for data processing and analysis. Neural network, which is part

of the machine learning, has the flexibility and ability of handling complex data. Currently, there are studies that used neural network to analyze spectra. Wang X *et al.* used backpropagation neural network (BPNN) and VIS/NIR diffused reflectance spectroscopy to predict the concentration of heavy metals in the soil (Wang *et al.*, 2020). Similarly, Wang Q *et al.* used BPNN and NIR spectroscopy to predict the concentration of nitrogen in the soil (Wang *et al.*, 2021). Besides, spectroscopy and neural network have also been used to predict the concentration of carbon in soil (Justin George *et al.*, 2020), anomaly in milk (Vasafi *et al.*, 2021), and protein in winter wheat (Soltani & Noguchi, 2018).

Therefore, in this study, VIS/NIR spectroscopy and FFBPNN were used to create the prediction models for the measurement of skin hydration and TEWL. VIS/NIR spectroscopy was chosen because of its non-invasive nature and its ability to reveal the amount of water content underneath the skin. FFBPNN is part of the machine learning and it is useful in analysing complex data such as the reflectrance and absorbance spectrum of the skin.

1.3 Scope of Study

This study investigated the relationship of skin hydration and TEWL, and alternative measurement method for skin hydration and TEWL using VIS/NIR spectroscopy. Twenty-four healthy participants were randomly selected and comprised of different genders, ethnics, and skin colours with age ranged from 22 to 32 years old. The data was collected from palms and arms of the volunteers with total of four targeted areas for each participant. This study did not investigate the chromophores that contributes to the skin colour because the peak absorbance wavelengths for

chromophores are outside of the selected wavelengths in this study. The skin hydration and TEWL data were collected using DermaLab USB Probe, while the ASD FS3 Spectrometer was used to perform the diffused reflectance spectroscopy. The data collection was done in an air-conditioned environment where the temperature was controlled by the air-conditioner. The relationship between skin temperature and water content underneath the skin is not considered in the investigation of this study. This is because the any changes in the water content will be equally detected by both DermaLab USB Probe and the spectrometer, thus skin temperature should not have any effect on the relationship between spectra, skin hydration, and TEWL. The relationship between skin hydration and TEWL was studied using statistical software Minitab and compared to previous studies by other authors.

Diffused reflectance spectroscopy was chosen because most body parts are opaque and inapplicable for transmittance spectroscopy. The reflectance spectra are then directly converted into absorbance spectra using ViewSpec Pro 6. The prediction models in this study are based on simple regression analysis, multiple regression analysis, and feed-forward backpropagation neural network. The regression analysis is done using Minitab Statistical Software, while both the neural network and its training algorithm is written in C# using Microsoft Visual Studio 2019. The training algorithm for the neural network are batch gradient descent and stochastic gradient descent. Finally, the performance of the prediction models is compared with one another.

1.4 Research Objectives

The objectives of this study are as follow:

1. To investigate the relationship between skin hydration and TEWL using statistical method with the comparison to previous studies.
2. To construct a skin hydration and TEWL prediction models based on visible/near infrared (VIS/NIR) spectroscopy, regression analysis and FFBPNN.
3. To compare and investigate the performance between regression models and FFBPNN models.

1.5 Novelty of This Study

The main objective of this study is to find out the feasibility of using both VIS/NIR spectroscopy and machine learning in predicting the skin hydration and TEWL of human skin. Some studies proposed VIS/NIR spectroscopy, whereas most of them used the classical statistical method such as partial least square, principal component analysis, and linear discriminant analysis (McIntosh *et al.*, 2002; Qassem & Kyriacou, 2019; Shin *et al.*, 2018).

Therefore, this study demonstrates the combined potential of VIS/NIR spectroscopy and FFBPNN for the measurement of skin hydration and TEWL. The strength and limitation of this combination is also discussed in this study. There are many past studies that investigated the relationship of skin hydration and TEWL under different environments and conditions that have shown contradictory results with one another. Thus, this study also investigates the relationship between skin hydration and TEWL and its similarity from other studies.

1.6 Structure of the Thesis

This thesis consists of five chapters. First chapter provides an overview of the study, problems statements, scope of study, research objectives and novelty of the study. Chapter 2 covers the literature reviews which includes previous works by other authors and basic information for skin anatomy, skin hydration, TEWL, spectroscopy, and neural network. Chapter 3 presents the detailed information about the participants, instruments, methods, and backpropagation algorithm that were used in this study. Chapter 4 presents the results and discussion in this study as well as the limitation of this study. Chapter 5 summarizes the results as well as provides some suggestions for future study. Finally, the link to the source code and human ethics consent form can be found in the Appendices.

CHAPTER 2
THEORY AND LITERATURE REVIEW

2.1 Introduction

This chapter introduces some basic theories and information about skin anatomy, skin hydration, TEWL, common technique in skin hydration and TEWL measurement, and the recent research regarding the skin hydration and TEWL assessment. In addition, the basic of absorbance spectroscopy and feedforward backpropagation neural network will be included in this chapter as well.

2.2 Skin

Skin covers our entire body with its sensory receptor and act as a protective barrier to the outside world. It gathers sensory information, produce vitamin D, waste elimination, protects us from pathogens invasion, physical hazard, chemical hazard, and ultraviolet (UV) radiation, regulates body temperature, and regulates water loss. Skin has a surface area of about 1.8 m^2 for an adult, thus it is the largest organ in the human body. The skin has two regions, epidermis and dermis, underneath these regions is the hypodermis, which loosely attached to any underlying structure such as muscle or bone (Longenbaker, 2011; Mac-Mary *et al.*, 2006).

2.2.1 Skin Anatomy

Figure 2.1 shows the structure of human skin. The hypodermis comprises of adipose (fat) tissue that helps insulate the body from cold, and provides protective padding against external assaults. Hypodermis gives the body a rounded appearance and excessive development of these adipose tissue would result in obesity.

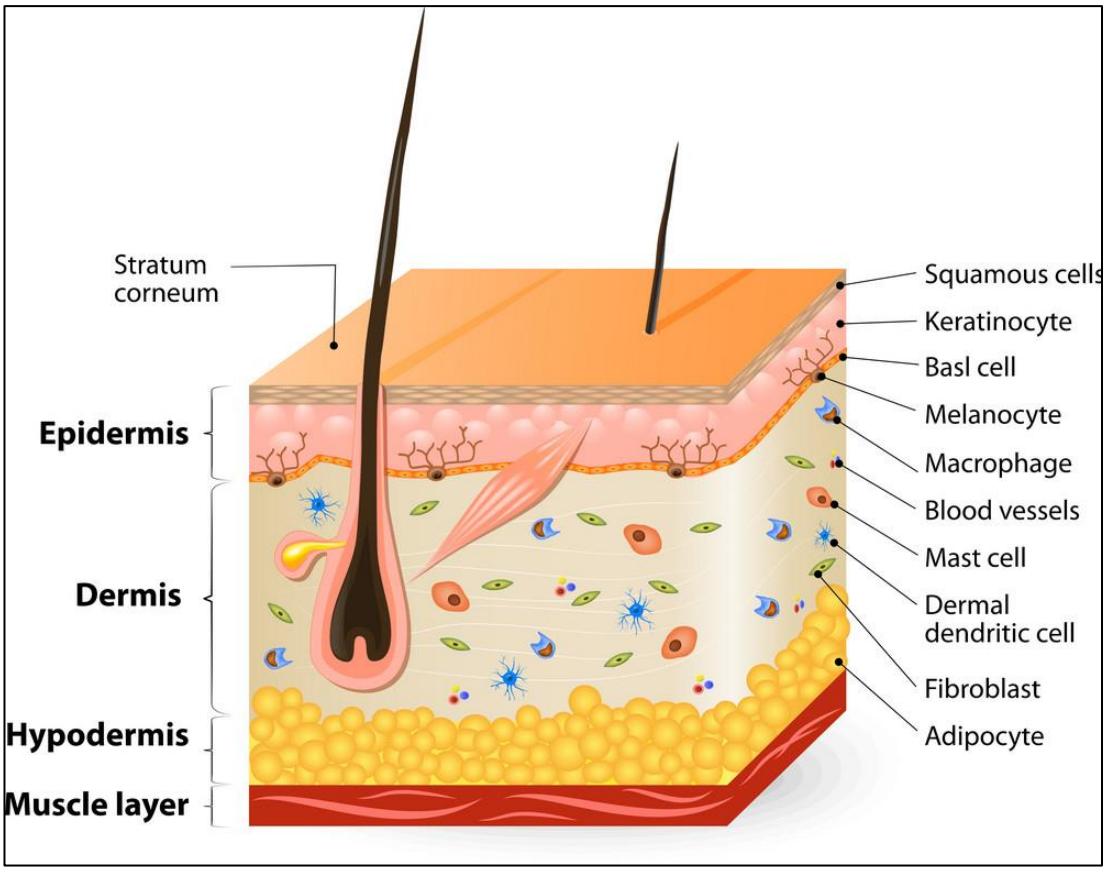

Figure 2.1 Structure of human skin (Designua, 2021)

The dermis is a dense irregular connective tissue that contains collagen and elastic fibres. These collagen fibres are flexible, allowing the skin to stretch during the movement of muscles and joints, as well as preventing the skin from tearing. Dermis also contains blood vessels and numerous sensory nerve fibres.

In the deeper layer of epidermis, it has basal cells, Langerhans cells, and melanocytes. Basal cells constantly divide and produce new cells to replace the old cells. Langerhans cells are macrophages that stimulate the immune system when they encounter foreign substances. Melanocytes are specialized cells that are responsible for the melanin production and skin colour. As the production of new cells pushes the old cells toward the surface of the skin, they become flat and hard, forming the tough,

uppermost layer of the epidermis. This hardening process is known as keratinization, the production of a fibrous and waterproof keratin protein (Longenbaker, 2011).

The topmost layer of the skin is known as stratum corneum, which consists of more than 15 layers of flattened corneocytes. It does not have any cellular structure or nucleus. It is formed when the skin cells matured and moved to the top layer of the skin, and it plays a vital role as a protective barrier against the external environment (Murphrey *et al.*, 2018). Besides that, it can receive water from dermis and from the environment to keep the skin moist (Mac-Mary *et al.*, 2006).

2.2.2 Skin Hydration and TEWL

Water is an essential element for any living organism on Earth. For a human, water is preserved through water permeability barrier and natural moisturizing factor. The permeability barrier comprises multiple lipid bilayers which prevents water loss and penetration of harmful chemical from the environment (Landmann, 1988). Natural moisturizing factor is a mixture of various amino acids and salts, it helps in absorbing atmospheric water to increase the skin hydration and it can be found on the stratum corneum. However, it can be easily washed out when exposed to soap. (Rawlings *et al.*, 1994).

The amount of water that binds to the skin is known as skin hydration and it keeps the skin soft, hydrated, and healthy. This water holding capacity is regulated by the skin to keep it moist but not over-hydrated. A compromised skin will not be able to regulate the water content as it intended for and thus causing the skin to be too dry or occlusive. Factors that can compromise the water holding capacity are infection by

pathogens, injuries due to harmful chemicals, physical injuries, inappropriate lifestyle, hereditary factors, and aging (Dąbrowska *et al.*, 2018).

In addition, TEWL is part of the regulation process by the skin to diffuse excessive water out of the skin. Like the water holding capability, this process can be influenced by harmful chemicals, physical injuries, infection, humidity, skin temperature, skin hydration, sweating, aging and the integrity of stratum corneum (Tupker *et al.*, 1989). Individuals that have abnormal level of TEWL might experience dehydration, infection, and physical hazard. In addition, TEWL often relates to skin barrier function. Thus, some skin diseases are treated with Fluocinonide 0.1% cream to improve its barrier function, which is measured by TEWL (MOH-Malaysia, 2018).

There are studies that investigate the relationship of skin hydration and TEWL, with contradicting results. Study on patients with skin disease and the application of moisturizing agent often shows a negative correlation between skin hydration and TEWL (Hagströmer *et al.*, 2006; Lee *et al.*, 2012; Manosroi *et al.*, 2011; Rim *et al.*, 2005). This phenomenon shows that skin with higher water evaporation rate often leads to dry skin, while the treatment using moisturizing agent can increase the skin hydration and reduce the TEWL.

However, there are studies that show a positive correlation between the skin hydration and TEWL (Black *et al.*, 2000; Conti *et al.*, 1995; Song *et al.*, 2015). In addition, there are few research claiming that the correlation between skin hydration and TEWL is poor (Cravello & Ferri, 2008; Q. Zhang *et al.*, 2018). A study found that the atopic (red and itchy) skin has an increased in the TEWL but there is no correlation

to the skin hydration (Loden *et al.*, 1992). Based on the European Group on Efficacy Measurement and Evaluation of Cosmetics and other Products (EEMCO) guidelines, the correlation between skin hydration and TEWL can be linear, positive, negative, or scattered, depending on the experiment, the condition, and the participant that are being measured (Berardesca *et al.*, 2018).

As shown by other studies that are mentioned above, it is possible that the correlation depends on the environmental parameters such as the climate or seasons, ambient temperature, and humidity. Hot climate, availability of sunlight, and healthy participants mostly yields positive correlation. In addition, this correlation also depends on the objective of the study; for instances, studies that focused on the performance of moisturizing agent, lotion, and treatment for dry skin often found that the correlation is negative. Lastly, the poor correlation could be due to the inappropriate experiment setup or factors that are yet to be found.

2.3 VIS / NIR Spectroscopy

VIS / NIR spectroscopy is the study of the interaction between matter and light in the visible and near infrared region. Light consists of a bunch of photons with discrete energy, frequency or wavelength, and spectroscope allows us to visually observe the intensity of each wavelength inside the light. The word “spectroscope” is derived from the Latin word “spectrum” meaning image and Greek word “skopein” meaning view. In modern day, spectroscopes are upgraded with various technologies, and they are usually known as spectroradiometer, spectrophotometer, or spectrometer. Spectrometer allow us to study and investigate the interaction of light and matter in details (Gauglitz & Moore, 2014; Hellman, 1968).

Spectrometer usually consists of diffraction grating, photodetector, filters, and waveguide, to collect and measure the intensity of light over a wide range of the electromagnetic spectrum. Light is first guided by a waveguide, then split and dispersed by diffraction grating into different direction according to their wavelengths. Light with different wavelengths then propagate and eventually reached at the corresponding photodetector where its intensity is recorded. The photodetector usually consists of an array of photodiodes covered with a filter to eliminate second and higher order light. The recorded data is a plot of intensity against wavelengths, and it is known as the spectrum (ASD, 2010). The spectrum can then be converted into reflectance, transmittance or absorbance depending on the purpose of experiment.

The interaction between light and matter is unique for different chemical compound, thus quantitative and qualitative analysis are possible by studying the spectrum. Spectroscopy is a sensitive tool that works on gas, liquid, and solid samples. It is virtually used in every field of sciences, such as remote sensing in geosciences, food quality control in food sciences, physics, and biomedical sciences.

The spectroscopy technique used in this study is diffuse reflectance spectroscopy. Diffuse reflectance spectroscopy is commonly used in VIS/NIR regions to obtain molecular spectroscopic information. It served as a logical alternative for the sample that are not applicable to transmittance spectroscopy. Such samples are mostly opaque or “infinitely thick” where light is unable to penetrate to the other end of the sample. This type of spectroscopy shows the result of combination of reflection and scattering of the incident light (Blitz, 1998). Ideally, it is best to use a reflectance spectrum without scattering effect to study the nature of the sample. However, it can be

extremely difficult to remove or separate scattering effect from reflectance spectra. Hence, scattering effect is not considered in the data processing of this study. Like most studies that used diffuse reflectance spectroscopy, the reflectance spectra in this study are converted into absorbance spectra to reduce its nonlinearities (Hunt Jr, 2021; Leone *et al.*, 2019). Figure 2.2 shows the spectra of water and human skin, the spectrum of water was obtained from Wikimedia, while the spectrum of human skin was obtained from the experiment in this study. Both spectra shared some absorption peaks that are used in this study, particularly at wavelengths 970 nm, 1200 nm, 1450 nm, and 1950 nm.

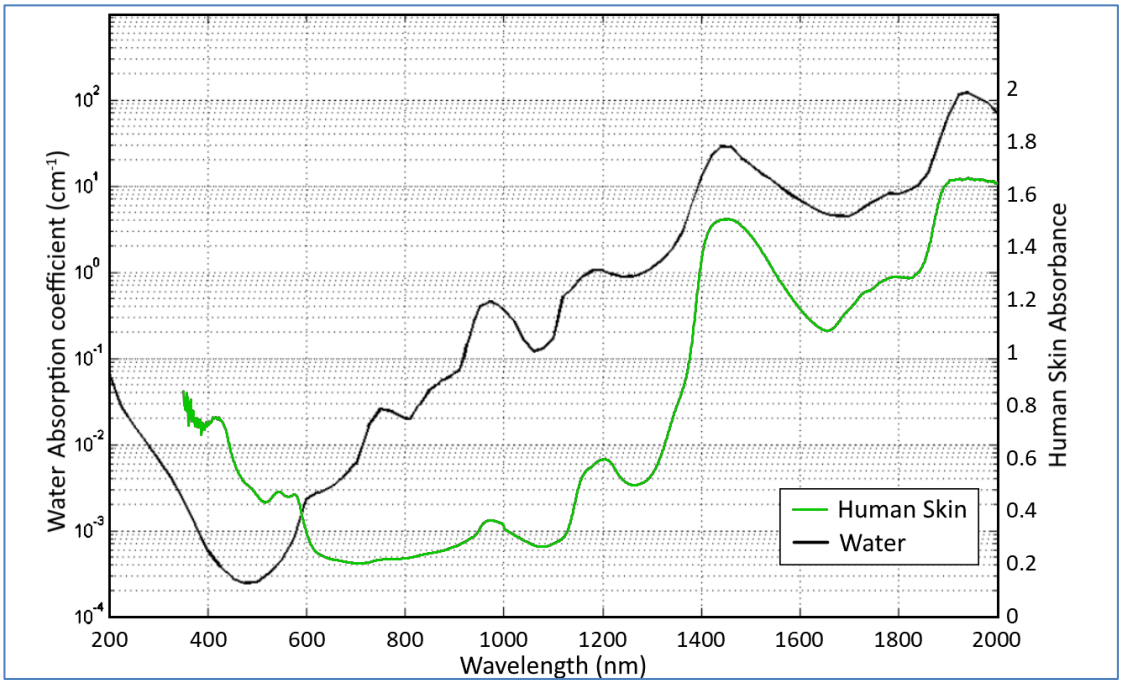

Figure 2.2 The spectra of water and human skin (Zhun310, 2009)

2.4 Light Absorption of Water Molecules

When light strikes a matter, it results in either transmission, reflection, absorption, or scattering. In this study, the light absorption by water molecules is used to predict the skin hydration and TEWL. Resonance radiation also known as

absorption, occurs when the frequency of light matches the vibration mode of the molecules. During the absorption phenomena, water molecules experience either O−H stretching, bending, or both as shown in Figure 2.3. The combination of these vibrations gives a complex spectrum, and the intensity of these vibrations increases when the absorption occurs (Carleer *et al.*, 1999). The absorption phenomena can then be observed on the absorbance spectra from a spectrometer.

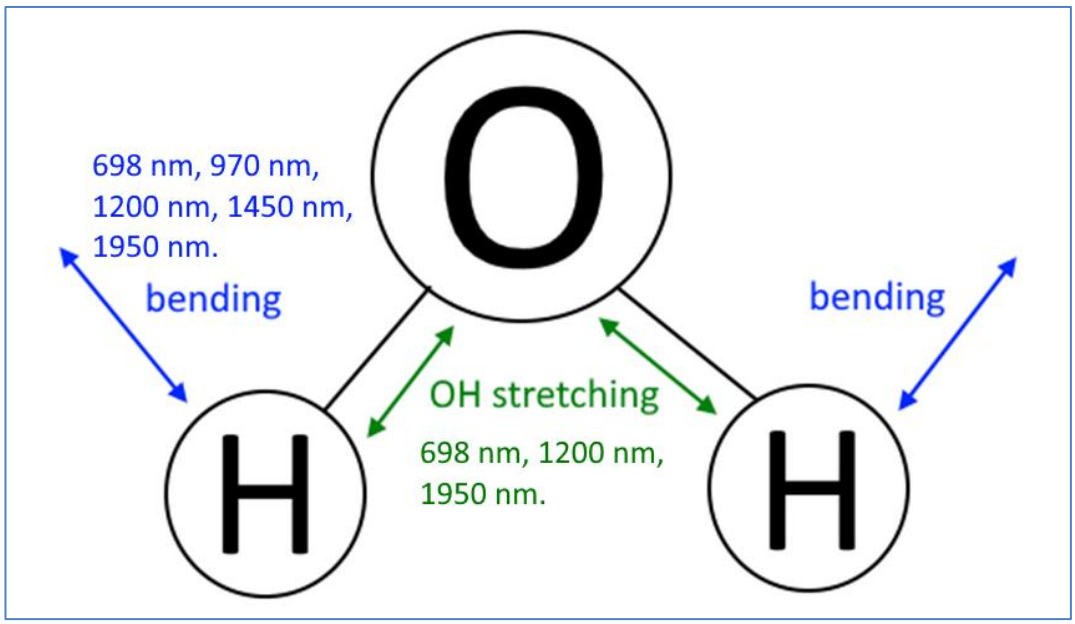

Figure 2.3 H_2O molecules when light absorption occurs

The spectrum of water is incredibly useful in many research areas, such as remote sensing, astronomy, quantitative and qualitative assessment for both organic and inorganic samples. Water can be detected using microwave, infrared, and near infrared. In this study, 698 nm, 970 nm, 1200 nm, 1450 nm, and 1950 nm are selected for water content assessment because there are studies that utilize these wavelengths with promising results (Deng *et al.*, 2012; Mamouei *et al.*, 2021; Pope & Fry, 1997). The absorption at 698 nm, 1200 nm, and 1950 nm is caused by both the O−H stretching

and bending resonance, while the absorption at 970 nm and 1450 nm is caused by O−H stretching only (Chaplin, 2017; Flaud *et al.*, 1997; Hein *et al.*, 2011; Matcher *et al.*, 1994). Besides, there are other water absorption wavelength that are not used in this study such as 760 nm and 975 nm (Raj *et al.*, 2021).

When wavelengths of light longer than 450 nm are absorbed by the water molecules, energy is transferred to one or more vibration modes of the O – H bond. The primary vibration modes for water O – H bond are at the wavelengths 2865 nm (stretch), 3049 nm (stretch), 6079 nm (bend), while other wavelengths are seen as overtone and harmonics of these vibration modes. The assignments of modes were determined experimentally by other authors who investigated the spectra of the mixture of water molecules, H_2O (water), D_2O (heavy water), and HDO. The liquid state ratio of $H_2O + D_2O \rightleftharpoons 2HDO$ gives hints and clues regarding the assignment of modes to its corresponding frequency in their studies. The combination of modes was chosen by considering the best fit of the observed band positions for both H_2O and D_2O, while the mode assignments were entirely based on the similarity of H_2O and D_2O spectra. In addition, the author also compared the ratios of resonant frequencies with the ratio of the mixture to validate their results. (Bayly *et al.*, 1963; Braun & Smirnov, 1993; Pegau *et al.*, 1997).

2.5 Regression Analysis

Regression analysis is a statistical technique for finding the relationship between variables, and it is also widely used to make future-oriented prediction model for the dependent variable. It is also widely used in many studies and experiments for its simplicity.

Regression that uses only one independent variable is known as univariate regression analysis or simple regression analysis. This type of regression is usually expressed in $y = mx + C$, where y is the dependent variable, m is the coefficient, x is the independent variable, and C is the constant. However, the accuracy using this method can be limited when the dependent variables can be affected by other independent variables. To overcome this issue, more than one independent variable needs to be taken into consideration, and the regression using multiple independent variables is known as multivariate regression analysis or multiple regression analysis.

Multivariate regression analysis is based on the assumptions that the independent variables are normally distributed, linearity, do not have extreme values, and no multicollinearity in the data. Multivariate regression analysis usually formulated as $y = C + m_1x_1 + m_2x_2 + m_3x_3 + \cdots + m_nx_n$, where n is the number of independent variables. When the correct combination of independent variables is used, multiple regression model can have a better prediction accuracy than simple regression model (Tabachnick *et al.*, 2007; Uyanık & Güler, 2013). Until today, multiple regression analysis is still widely used in spectroscopy for quantitative analysis (Basak *et al.*, 2020; Naveen *et al.*, 2020; Wang *et al.*, 2017).

2.6 Feed-Forward Backpropagation Neural Network

Machine learning is rarely seen or mentioned in the consumer market until the advancement of computing power in recent years. Modern processor allows machine learning to be more feasible. Since then, it is getting attention from various science field, and the research and the implementation of artificial intelligence (AI), machine learning, and neural network growth dramatically. Because of this, nowadays there are

many electronic devices that advertised with “Powered by AI Technology” or something similar.

FFBPNN is an algorithm that improve automatically by “learning” from the data that is given to it, and it is seen as part of the machine learning and artificial intelligent. The idea of neural network is first proposed by neurophysiologist Warren McCulloch and mathematician Walter Pitts in year 1943. The concept is based on how our body processes information and making decision using the interconnected neurons in our body, and they proposed using logical calculus to create a workable procedure that mimics the nervous net activity (McCulloch & Pitts, 1943; Mishra & Srivastava, 2014).

There are several types of neural network for different purposes, such as convolutional neural network (CNN) for image processing and computer vision, recurrent neural network (RNN) for text processing and sentiment analysis, and many more (Afana *et al.*, 2018). In this study, FFBPNN is used to estimate the skin hydration and TEWL from the skin absorbance spectrum. Feed-forward neural network (FFNN) is the most basic type of neural net that passes the information in only one direction from input to output. It is useful in various applications such as function approximation, forecasting, and pattern recognition. It needs to be trained properly before it can be used for prediction. In this study, the FFNN is trained using backpropagation learning algorithm, and FFNN that uses backpropagation algorithm is known as the FFBPNN.

Feed forward neural network consists of multiple layers, each layer contains at least one node, and each node represents a number. Figure 2.4 shows the example of a

three layers neural network. From Figure 2.4, the first layer is the input layer which consists of one input node, the middle layer is known as the hidden layer which consists of two hidden nodes, and the final layer is the output layer which consists of one output node. In the feed forward process, weighted sum of previous nodes is first calculated, then an activation function is applied on the sum, and finally the value is stored in the node. The feed forward process will continue from layer to layer until it reaches the final layer of the FFBPNN.

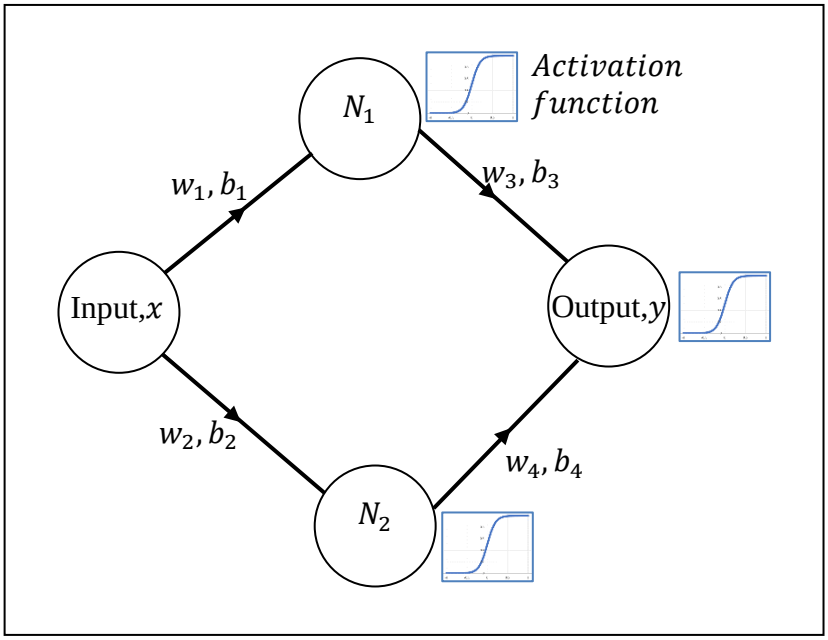

Figure 2.4 A Feed Forward Neural Network (FFNN)

To update and correct the weights and biases in a FFBPNN, a backpropagation algorithm is used. Various kinds of backpropagation algorithm have been introduced to optimise the backpropagation process, such as gradient descent and Levenberg–Marquardt algorithm (Biehl & Schwarze, 1995; Moré, 1978; Ruder, 2016). In this study, batch gradient descent and stochastic gradient descent algorithms were used to train the FFBPNN.

Initially, the weights and biases for each connection between nodes are populated with random numbers and it is meant to be fine-tuned in each iteration of the backpropagation process. There are four main steps in fine turning the weights and biases in FFBPNN using gradient descent algorithm. Firstly, determine the loss function; Secondly, determine the derivative of the loss function; Thirdly, determine the step size based on the derivative of the loss function, learning rate, and momentum. Lastly, update the weights and biases according to the step size. These four steps will repeat until the network can make good predictions.

The loss function is shown in Equation 2.1, and it is also known as the sum of squared residual (*SSR*). *SSR* represents the errors between the predicted values and the observed values. When *SSR* is plotted against multiple possible weights and biases, a curved (parabola) graph can be seen, Figure 2.5 shows the general idea of *SSR* vs weights plot, and the minimum point on this curve has the lowest error and minimum gradient.

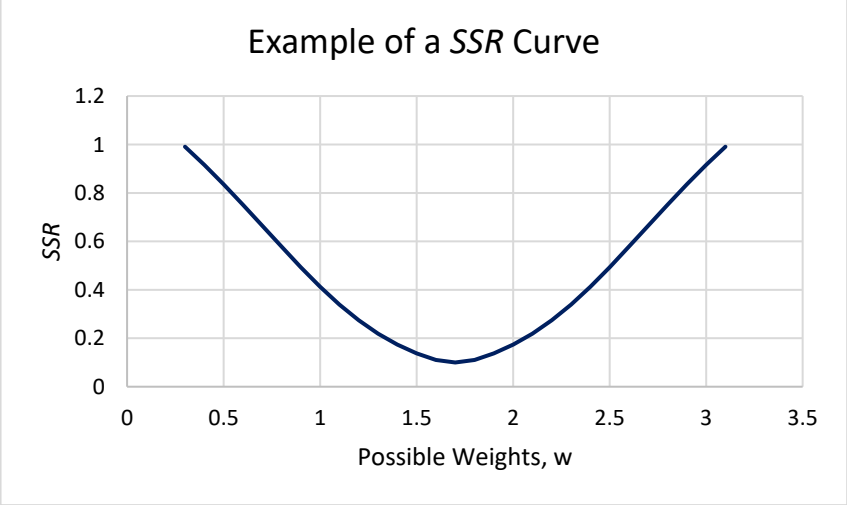

Figure 2.5 *SSR* vs Possible Weights

The gradient can be determined with the first derivative order, $\frac{d(SSR)}{d\,w}$ and $\frac{d(SSR)}{d\,b}$, and it is used to track the gradient at any point on the *SSR* curve for the weights and biases in each layer. After that, the step size can be determined using Equation 2.2 and Equation 2.3, where the learning rate is a constant real number that is smaller than 1. With each step, the algorithm is tracking the gradient and aims to descent along the *SSR* curve, thus this algorithm is known as the gradient descent. Finally, the new weights and biases are updated using Equation 2.4 and Equation 2.5, and it is now ready for the next feed forward process (Bottou, 2012; Deisenroth *et al.*, 2020; Smola & Vishwanathan, 2008).

$$SSR = \sum (y_{Observed} - y_{Predicted})^2 \tag{2.1}$$

$$\alpha_w = \Delta_w \times \eta \tag{2.2}$$

$$\alpha_b = \Delta_b \times \eta \tag{2.3}$$

$$w_{New} = w_{old} - \alpha_w \tag{2.4}$$

$$b_{New} = b_{old} - \alpha_b \tag{2.5}$$

Where,

SSR	=	Sum of squared residuals
y	=	Output
Δ	=	$\frac{d\,(SSR)}{d\,W}$
Δ	=	$\frac{d\,(SSR)}{d\,W}$
η	=	Learning rate
α	=	Step size
w	=	Weight

b = Bias

In batch gradient descent, all data is considered in each epoch to calculate its loss, derivative, and step size, while in stochastic gradient descent, it will randomly use a single data point in each epoch for the calculation of loss, derivative, and step size. While both shared the similar set of equations, stochastic gradient descent is theoretically less computationally intensive and able to achieve optimal solution faster compared to the batch gradient descent but with a tendency of having more noise. Training a neural network can be time consuming, depending on the size of data and the specification of computer. A simple logic gate operation can be easily trained within seconds, but a large batch of data can take hours or days to complete. In this study, both batch gradient descent and stochastic gradient descent were used as the method to train the FFBPNN in this study.

2.7 Skin Hydration and TEWL Assessment

There are various types of skin hydration checker available in the market, from consumer grade to industrial grade. The consumer grade skin hydration checker can be easily obtained from local pharmacy store at reasonable price, while the industrial grade devices are far more expensive and hardly accessible by the consumer. On the other hand, most of the TEWL probes are quite expensive and intended for industrial and professional purposes only. These devices are based on electrical method. The skin hydration is estimated based on the measurement of resistance, capacitance, conductance, and impedance of the skin, while the TEWL is an evaporimeter that measures the water vapor gradient.

The research regarding skin moisture and its electrical resistance and capacitance can be as early as year 1962 (RUTENFRANZ *et al.*, 1962), and the device was developed in year 1977 using a four-point probe to study the resistivity of the skin to estimate the water content of stratum corneum (Campbell *et al.*, 1977). Due to the complex electrical properties of skin, particularly its physical and chemical state, various kinds of probes were developed and proposed to study the skin (Edwards, 1988; Kohli *et al.*, 1985; Yamamoto & Yamamoto, 1981). The one that resembles the design of modern skin hydration probe was introduced in 1986 by Yamamoto. The probe consists of two dry electrodes plate that is made of flatten copper to estimate skin hydration by measuring the skin impedance (Yamamoto & Yamamoto, 1986).

There are several electrical methods that proposed by different authors to estimate the skin hydration. Bio-impedance has been studied for many years before it is used to estimate the skin hydration. The concept is to apply a sinusoidal voltage, $\Delta E \sin \omega t$ through a pair of electrodes. When the electrodes touched the skin surface, a sinusoidal current, $i \sin(\omega t + \theta)$ will flow with a phase angle, θ, and ω equals $2\pi f$, where f is the sinusoidal frequency in Hz. The phase angle represents the phase difference between the applied sinusoidal voltage and the resultant sinusoidal current. Magnitude of the impedance, Z can be measured by using $|Z| = \frac{E(t)}{i(t)}$ (Archer *et al.*, 1990; Yamamoto & Yamamoto, 1987). The second method by applying an alternating electric current that operates at a specific frequency to the skin and measures the conductance of the skin. The frequency usually ranged from 100 Hz to 10 kHz depending on the device and the person who designed the device. Due to the "skin effect", a large electrical current is formed on the skin surface (Martinsen & Grimnes, 1998). Since the hydrated skin is more conductive than dry skin, the measured

conductance can be used to represent the skin hydration, where hydrated skin has higher conductance. This method is widely adopted in the industrial grade skin hydration probe today, it is also the method that used by the skin hydration probe in this study (Constantin *et al.*, 2014; O'goshi & Serup, 2007; Sim *et al.*, 2019). Third method is widely used by the cheaper skin hydration probe that can be found in local pharmacy shop. These probes are usually powered by battery, comes with two electrodes, and measured in arbitrary unit that ranged from 0 to 100 (ScalarCorporation, 2006). Unfortunately, the exact working mechanism of it is often not mentioned in the manual and could not be found on the internet.

The instrument for TEWL assessment is an evaporimeter that measures the evaporation rate of water or the rate of the water loss from skin surface. Inside the evaporimeter are two vertically stacked sensors coated with an organic polymer that are dielectrically sensitive to humidity, and each sensor is coupled with a temperature dependent resistor. Both sensors are enclosed in a cylindrically shaped plastic capsule and open at both ends. In order the measure the TEWL, the probe is held against the skin with the sensors parallel to the skin and perpendicular to the vapor pressure gradient for roughly 15 to 30 seconds (Grove *et al.*, 1999; Roskos & Guy, 1989; Scott *et al.*, 1982). Figure 2.6 shows the images of probes that are mentioned above. Figure 2.6 (i) is the design of impedance probe that proposed by Yamamoto *et al.* (1986) , Figure 2.6 (ii) is the industrial grade conductance probe from Dermalab, Figure 2.6 (iii) is the consumer grade skin hydration probe from Scalar Corporation (ScalarCorporation, 2006) and Figure 2.6 (iv) is the TEWL probe from Dermalab.