國立臺灣大學理學院地理環境資源學系暨研究所

碩士論文

Department of Geography

College of Science

National Taiwan University

Master thesis

應用類神經網路推估
低緯度濕地二氧化碳通量變化

Application of Artificial Neural Network Model on Estimation of Carbon Dioxide Flux in Low Latitude Wetland Ecosystem

呂姿儀

Tzu-Yi Lu

指導教授：莊振義　博士

Advisor: Jehn-Yih Juang, Ph. D.

中華民國 106 年 6 月

June, 2017

doi:10.6342/NTU201703598

國立臺灣大學碩（博）士學位論文
口試委員會審定書

應用類神經網路推估低緯度濕地二氧化碳通量變化
Application of Artificial Neural Networks Model on Estimation of Carbon Dioxide Flux in Low Latitude Wetland Ecosystem

本論文係呂姿儀君（R04228007）在國立臺灣大學地理環境資源學系、所完成之碩（博）士學位論文，於民國 106 年 6 月 16 日承下列考試委員審查通過及口試及格，特此證明。

口試委員：

（簽名）

（指導教授）

doi:10.6342/NTU201703598

Acknowledgement

這份論文的完成感謝一路上許多人實質或精神層面的幫助，但憑我一己之力是無法順利完成的。首先，特別感謝我的指導教授莊振義老師，在大三時提供我進研究室實習的機會，在研究室學習的四年中，增進了我對環境科學、儀器操作、待人處事等相關的知識，並擴大了我對於學術研究的視野。感謝老師的栽培與教導，使我在這段期間收穫豐碩。同時，感謝四位給予珍貴建議的口試委員，張斐章老師、黃倬英老師、林楨家老師、溫在弘老師，讓這份論文能夠更加嚴謹。

時光匆匆，轉眼之間碩士班兩年的時光就這樣度過了。這段時間，說長不長，說短不短，但卻是充滿各種幸福與成長的歲月。在小時候的人生規劃中，未曾想過會有碩士班的這兩年的機緣，但感謝有這兩年的時光，看到、聽到、遭遇到，也學到了更多，讓我成為如今的面貌，並提醒著自己要時時努力向前。雖然這本論文中仍有許多美中不足之處，但這兩年來的大大小小、點點滴滴，都化作一字一句累積在此。

感謝研究室的各位夥伴。宇晨學長、釗榮學長、松靜學長，在我甫進入研究室學習時，不厭其煩的指導我、各種罩我、包容我各種雷，在研究室當小學妹的時光是非常快樂的回憶；淇雅、祥恒在各種時候的神救援，總是受到你們許多照顧；思方、子瑄、逸倫各種野外互相照應。

感謝親愛的呱呱，住在一起的一年，絕對是最快樂、最珍貴、最溫暖的一年。感謝你總是在我論文崩潰而慌張的每個當下，耐心聆聽我的苦惱，溫柔的給我鼓勵，理性的為我分析方向。因為有你，我才能總是從不知所措的慌亂中回到正軌，真的非常非常感謝，接下來換你加油了，希望我也能為你做到如你為我所做的這種程度，然後也祝你接下來也能夠一切順利。

感謝親愛的鴻錡，這兩年中，之所以能夠堅持下來，努力完成論文的最大功臣，莫過於你。在每一個想放棄的瞬間，淚流不止的當下，對於茫茫前途感到徬徨的時刻，都是你的手牽著我繼續走下去。不只是這兩年，在一起至今，你也總是很辛苦，忍受著我的壞脾氣，導正我朝向理性成熟的方向發展，在每個懶散的時候鞭策我、督促我向前，總是有條不紊地和我討論著論文每個階段的問題，給我許多中肯的建議，沒有你的話，我大概沒辦法這麼快畢業吧哈哈，真的無比偉大！最後，感謝爸媽家人的支持與照顧，讓我可以無後顧之憂的完成大學及研究所的學業，順利畢業。

呂姿儀 謹誌於

國立臺灣大學地理環境資源學系

中華民國一零六年八月

doi:10.6342/NTU201703598

摘要

目前濕地研究多集中於高緯度的泥炭地(peatlands)，低緯度濕地的研究相對較少，然而低緯度濕地面積占全球濕地面積之 70%，若以泥炭地之研究結果進行全球濕地碳通量的推估，可能會與實際情形產生誤差。在碳收支的研究中，渦度相關法(eddy-covariance method)被認為是最直接而準確地量測生態系統通量的方式，但受限於此法對量測環境的要求，故無通量塔處之碳通量變化趨勢須以推估的方式進行。在眾多推估方式中，類神經網路(artificial neural network)在各種不同生態系統的推估研究中，已證實具有良好的模擬能力，能夠精確掌握氣象資料與碳通量資料之間的關聯性。

本研究使用類神經網路作為推估工具，採用關渡塔一(GDP-T1)、關渡塔二(GDP-T2)、佛羅里達站(US-Esm)三個低緯度濕地測站資料(包括氣象及二氧化碳通量)作為輸入及輸出參數，配合倒傳遞演算法(back propagation algorithm)進行模式建立與推估。模型訓練完成後，計算相關係數(Correlation coefficient)、均方根誤差(RMSE)、平均絕對百分比誤差(MAPE)等指標，以討論推估結果與觀測結果之間的誤差情形，以及模型推估二氧化碳通量變化模式之能力。GDP-T1 及 GDP-T2 站最佳結果出現在日間模型(R=0.89 及 R=0.87)，US-Esm 站點最佳結果出現在夜間模型(R=0.62)。跨站模擬最佳情況可高達 R=0.73，此結果顯示未來以類神經網路應用於大範圍推估為可行的。

本研究所建立之模型可應用於濕地生態系統進行二氧化碳通量數據的推估，提高全球濕地碳收支推估精確程度，並進一步研究無通量塔處之二氧化碳通量變化特徵，分析相似生態系統二氧化碳通量的變異模式，如季節或年際變異等，克服通量塔設置之空間限制及節省儀器架設成本。在面對氣候變遷時，此模型的推估能力也可增加對於未知風險的了解，作為評估未來變化趨勢的依據之一。

關鍵字：渦度相關法、倒傳遞演算法、二氧化碳通量、氣象因子

doi:10.6342/NTU201703598

Abstract

In wetland studies, few attentions have been given to low-latitude wetland ecosystem presently, but it accounts for about 70% of the global wetland area. Therefore, it's very important to consider the contribution of this significant portion on global carbon (C) budget. In the past decades, eddy-covariance method has been widely applied in many C budget studies at the ecosystem scale, but there are still several limitations affecting the performance of EC methods. In order to overcome the abovementioned limitations, many linear or non-linear statistical techniques are applied to fill the measurement gap. Among various methods, the Artificial Neural Network (ANN) method is considered to be an excellent means to identify the complex non-linear relationship between the CO_2 flux and meteorological variables.

In this study, a back-propagation ANN model was applied to quantify CO_2 flux at three low-latitude wetland sites (Guandu Nature Park Tower One (GDP-T1), Guandu Nature Park Tower Two (GDP-T1), and Florida Everglades short hydroperiod marsh (US-Esm)) in East Asia and the US. Meteorological variables were used as the input parameters to train the ANN to predict the CO_2 exchange. The best results of the GDP-T1 (R=0.89) and GDP-T2 (R=0.87) occurred in the simulation of the daytime (DT) model, and that of the US-Esm (R=0.62) in the nighttime (NT) models. The cross-site simulation was feasible, the best result could up to 0.73 in terms of R. This model provided a quick, efficient, and highly accurate estimation, and could be conducted to estimate the dynamics of CO_2 flux where there is no direct in-situ flux measurement. The simulation capability is helpful to characterize the spatial/temporal variations in low-latitude wetland ecosystems, and improve the quantification of global C budget.

Keyword: eddy-covariance method, back propagation algorithm, carbon dioxide flux, meteorological factors

doi:10.6342/NTU201703598

Contents

Acknowledgement i

摘要 ii

Abstract iii

Contents iv

List of figures vi

List of tables vii

1. Introdution 1

1.1. Wetland ecosystems 1

1.2. Eddy covariance method 3

1.3. Estimation of C flux 5

1.4. The application of ANN models on C budget 6

1.5. The relationship of environmental factors and NEE 11

1.6. Objectives 12

2. Methods 14

2.1. Site descriptions 14

2.1.1. Guandu Nature Park Tower One (GDP-T1) 14

2.1.2. Guandu Nature Park Tower Two (GDP-T2) 15

2.1.3. Florida Everglades short hydroperiod marsh (US-Esm) 16

2.1.4. Brief summary 16

2.2. Study design 18

2.3. Artificial Neural Network (ANN) 19

2.4. Input and output variables 23

2.5. Data pre-processing 24

doi:10.6342/NTU201703598

2.6. Model comparison indices 26
3. Results 29
3.1. The model performance in each site 29
3.2. The data distribution at each site 31
3.2.1. All data 34
3.2.2. Daytime data 35
3.2.3. Nighttime data 36
3.3. The capability of cross-site prediction 40
3.4. The comparison of data distribution between different site 41
4. Discussion 44
4.1. The data period in each site 44
4.2. The advantage of the meteorological variables as the input data 44
4.3. The comparison of MAPE in different site 45
4.4. The results of data division 45
5. Conclusions and Future Works 47
Reference 49
Appendix I: The maximum (max) and the minimum (min) of input parameters used in the normalization process at each site. 55
Appendix II: The RMSE curve of three groups in the training phase. 56
Appendix III: The simulation results in each site. 61
Appendix IV: The cross-site simulation results. 72
Appendix V: The code used in the training phase 81
Appendix VI: The code used in the simulation phase. 85

doi:10.6342/NTU201703598

List of figures

Fig. 1. The location of the study sites. 14

Fig. 2. The location of the towers.....18

Fig. 3. The structure of ANN.....21

Fig. 4. The operation process in each neuron.....21

Fig. 5. The Log-Sigmoid transfer function used in the ANN models. 22

Fig. 6. The simulation result of the AL group of GDP-T1 from selected time period. 30

Fig. 7. The simulation result of the NT group of GDP-T2 from selected time period 30

Fig. 8. The data distributions of AL group at each site 36

Fig. 9. The data distribution of DT group of each site 38

Fig. 10. The data distribution of NT group of each site 39

Fig. 11. Ensemble averages of hourly CO_2 flux of the three sites.....42

doi:10.6342/NTU201703598

List of tables

Table 1 Relative references which used the ANN model 10

Table 2 The details of the study sites 18

Table 3 The numbers of neuron in the hidden layer of each model 23

Table 4 The numbers of effective datasets in each site 26

Table 5 The numbers of the training data and the simulation data in each site 26

Table 6 The criteria of MAPE 28

Table 7 The model performance of each group in the simulation phase 31

Table 8 Comparison of the model performance in daytime and nighttime interval of each group in the simulation phase 31

Table 9 The result of the *R* with Fc in each site. 34

Table 10 The simulation results of GDP-T1 and US-Esm in the GDP-T2 model .. 41

Table 11 Comparison of the model performance in daytime and nighttime interval of each group in the simulation phase 41

Table 12 The *R* of each site 43

doi:10.6342/NTU201703598

1. Introdution

1.1. Wetland ecosystems

Wetland ecosystem is one of the most important terrestrial ecosystems, but it has been neglected in the past due to its small proportion of the terrestrial ecosystem area. The wetland ecosystem is the natural optimum environment to store soil organic carbon (SOC) and sequestrate the carbon from the atmosphere due to the anoxic wet condition, and it is estimated that about 20 to 30 % of the global soil pool of 2500 Pg of carbon (C) is stored in the wetlands (Lal, 2008; Mitsch et al., 2013). It represents a large component of the SOC pool for long term storage. Despite the fact that the total coverage of the wetland ecosystem only account for 6-8% of the terrestrial area (Mitsch and Gosselink, 2007; Mitsch et al., 2013), the wetland ecosystem is one of the largest biological C pool and stored more C in the soil than other biomes which occupy larger area (Bernal and Mitsch, 2008). Moreover, the wetland ecosystem is also regarded as the net C sink which has the high capacity for C sequestration with an average of 118 g-C m^{-2} yr^{-1}, and one of the most productive ecosystems around the world, which accounts for about 6.3% of the terrestrial net primary production (NPP) (Neue et al., 1997; Mitsch et al., 2013; Fennessy, 2014). For these reasons, the wetland ecosystem plays an important role in the global C cycle.

Most wetland studies focused on the northern peatlands, and few have investigated the low latitude wetland ecosystems (Bernal and Mitsch, 2012). It is estimated that 455 Pg of C was stored in the northern peatlands (Gorham, 1995). In the context of global warming, the permafrost degradation in peatlands will have a great influence on the global C dynamics. The themes of the northern peatlands studies included the influence on the global C budget, the estimation of net C storage, the temporal and spatial variation, and

doi:10.6342/NTU201703598

the analysis of control factors of C flux (Armentano and Menges, 1986; Bubier et al., 1998; Waddington and Roulet, 2000). Although these results were conducted to understand wetland ecosystems more, they only represented a portion of the wetland ecosystem. However, current studies have noticed that the C sequestration of ecosystems varies over different environmental conditions (e.g. vegetation, hydrogeomorphology, climate, and latitude). Bernal and Mitsch (2012) estimated the soil C pool and the sediment accretion rate to compare six temperate wetland communities in Ohio to explain that wetland ecosystems have different C sequestration capabilities because of the climate conditions, vegetation communities, and hydrogeomorphic characteristics. In this study, the depressional wetland sites had higher organic content in soil C and lower bulk density than the riverine wetland sites. Additionally, the six sites had different C sequestration rates, and the rates of depressional wetland sites were higher than the riverine wetland sites in general. The highest C sequestration rate was found in the *Quercus palustris* forested wetland community (473 g C m^{-2} yr^{-1}) in the depressional wetland group. It is noted that the capabilities of C sequestration were quite different even though the kinds of sites were in the similar hydrogeomorphic conditions, respectively.

Mitsch et al. (2013) discovered that the capability of C sequestration of wetland ecosystems under several climate conditions varied quite differently, and most of the C retention occur in low-latitude wetlands. Seven sites, including three temperate wetlands and four tropical wetlands, were chosen to estimate the C budget, and the estimation results of the C budget of these sites were also compared to those of the fourteen wetlands investigated in the previous studies. There were totally six tropical wetlands, seven temperate wetlands, and eight boreal wetlands described in this study. The results revealed that the average sequestration rate of carbon dioxide (CO_2) was 214±66 g C m^{-2} yr^{-1} for the tropical/subtropical wetlands, 320±51 g C m^{-2} yr^{-1} for the temperate

doi:10.6342/NTU201703598

wetlands, and 49±18 g C m^{-2} yr^{-1} for the boreal wetlands. However, if taking the total area of wetlands into consideration, the total amount of C retention was 5.6×10^{14}, 1.6×10^{14}, and 1.1×10^{14} g C yr^{-1} for the tropical/subtropical, temperate, and boreal wetlands, respectively. As a result, when it comes to global wetland C budget, the variations of the net ecosystem exchange (NEE) of wetland ecosystems at different latitudes should not be lumped together.

Due to influences of the year-round growing seasons, the characteristics of NEE in the low latitude wetlands differ from the northern peatlands and other temperate wetlands (Schedlbauer et al., 2010). For this reason, understanding the C exchange dynamics and their control factors is important to the estimation of global wetland C budget.

1.2. Eddy covariance method

The eddy-covariance (EC) method is a meteorological technique, which directly measures the fluxes of scalars or energy components across the interface between the atmosphere and the plant canopies. In this study, the negative values of CO_2 flux represented the C uptake, and on the contrary, the positive values represented the C release. Continuously observed EC data represents point measurement with a footprint ranging from meters to kilometers depending on the measurement height, surface roughness, and atmospheric stability. This method is quite accurate when the atmospheric condition is steady, the underlying plant is homogeneous, and the instruments are situated on a flat terrain.

The weakness of the EC method is that it is critical to the environment as mentioned above, and the instruments are very expensive. In addition to the uneven distribution of the present sites, sometimes it is difficult to find adequate locations to establish the flux

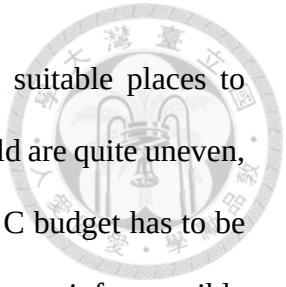
doi:10.6342/NTU201703598

towers. The EC method is strict with the background of sites. The EC towers need to be located on relatively flat terrains with vegetation extending horizontally about 100 times the sampling height (Baldocchi et al., 2001). The measurements can only be accurate when the atmospheric conditions remain steady. Although the EC instruments can measure accurately, they have to be maintained and calibrated frequently. That is to say, the EC data is helpful for understanding the exchange of mass and energy and interactions between the atmosphere and ecosystems, but when it comes to the application of the EC instruments, many conditions should be taken into consideration including site selection, instrument placement, and instrument calibration and post-processing (Baldocchi et al., 2001).

Although the EC system is able to measure the high-accurate data, some conditions may lead to the measurement errors. First, the measurement may be inaccurate in the unstable environment; second, radiative cooling leads to stable conditions that suppress turbulent mixing, so it is hard to recognize the variation of turbulent for instruments; third, the instruments are not calibrated correctly resulting in the measurement errors (Ooba et al., 2006; Dragomir et al., 2012).

The FLUXNET was established to gather the long-term observations from regional micrometeorological tower sites together, and the data has been applied to quantify the spatial/temporal variations of ecosystems and to understand the underlying mechanisms responsible for observed fluxes and C pools of the global C cycle. The Fluxnet lists 650 tower sites which are in operation, but most sites are distributed over Europe, North America, and Northeast Asia. Moreover, among all of the sites, wetland sites are quite sparse, so the understanding of wetland ecosystems is relatively insufficient. Further, the majority are located in middle or high latitude areas.

doi:10.6342/NTU201703598



Based on the reasons mentioned above, it is not easy to find suitable places to establish EC towers, and the distribution of EC towers around the world are quite uneven, and most towers are clustered in developed countries. Therefore, the C budget has to be estimated in the place where it cannot be directly measured. In order to infer possible temporal and spatial patterns of C flux at the place without *in-situ* measurements, some current studies attempt to predict the variation of C flux by extrapolating existing observations to large scale with the help of modeling.

1.3. Estimation of C flux

Data from the EC system are usually sampled with high frequency and half-hourly averaged for continuous measurement all around the year. Nevertheless, the yearly average data coverage is merely 65% because of the system failures or data rejection (Falge et al., 2001). In order to fill the missing data for complete dataset, many investigators have developed and implemented some site-specific techniques. There are some common approaches such as interpolation, look-up tables, mean diurnal variation method, the Monte Carlo method, linear regression, nonlinear regression, and the artificial neural network (ANN) (Falge et al., 2001 ; Hollinger et al., 2004; Knorr and Kattge, 2005; Ooba et al., 2006).

The non-linear regression method is the most conventional way to estimate the C flux, which uses the photosynthetic photon flux density (PPFD), the air temperature, and some empirical parameters to calculate the C flux based on the characteristics of the bioreaction. It is reasonable and easy to calculate. However, it may lead to over-simplification of ecosystem and neglect some information in the environmental condition (Ooba et al., 2006). Other estimation methods, such as look-up tables and the mean

doi:10.6342/NTU201703598

diurnal variation method was considered to be insufficiently accurate (Carrara et al ., 2003 ; Mogami et al ., 2003).

Among many estimation methods, the ANN technique is a good way because of its fast and accurate calculation capability. This method has been applied for estimating NEE within the past two decades (Van Wijk and Bouten, 1999; Papale and Valentini, 2003; Melesse and Hanley, 2005). However, compared to other methods, the ANN technique has the better performance for C flux prediction than other methods (Van Wijk et al., 2002; Ooba et al., 2006). The ANN technique is unnecessary for ANN users to have special background knowledge of the field to which ANN technique is applied. In addition, this method is more sensitive to the relationship between the meteorological inputs and C flux than other methods (Melesse and Hanley, 2005).

In the ANN models, the flexible structure can identify the non-linear relationship between the input data and the output data. Owing to its excellent capability of calculation, the ANN model can deal with numerous datasets at the same time and calculate the result quickly (Melesse and Hanley, 2005). If the input data is sufficient, the ANN model can establish the relationship between the input and output accurately. The ANN technique is suitable for some kinds of problems, which are too complicated to define the background conditions, which need to calculate the results quickly, and which have to establish the non-linear relationship between the input and the output (Chang and Chang, 2015).

1.4. The application of ANN models on C budget

Some studies have improved that the relationship between environmental factors and C flux can be established by ANN models (Van Wijk and Bouten, 1999; Van Wijk, Bouten, and Verstraten, 2002; Melesse and Hanley, 2005; Leuning et al., 2005; Ooba et al., 2006).

doi:10.6342/NTU201703598

Most of the current studies focused on the forest and grassland ecosystems, and the discussion on the wetland ecosystem is limited, especially in the CO_2 flux. Ooba et al. (2006) used 15-day data and six input parameters to establish the relationship between the environmental factors and CO_2 flux with a back-propagation neural network to compare three different kinds of ecosystems. The R^2 values of the data simulation of forest, grassland, and wheat field models were 0.86, 0.75, and 0.94, respectively. It showed that the ANN models had a great ability to predict the CO_2 flux. Moreover, this study also concluded that the different ecosystems need to choose different input parameters.

For the time being, few study uses the ANN model to qualify the CO_2 flux of wetland, but some wetland studies applied the ANN models to CH_4 flux. Morin et al. (2014) use the ANN model to determine the environmental drivers to the CH_4 flux. The data was divided into four groups, including summer day, summer night, winter day, and winter night, to discuss the seasonal and diurnal impacts. The model input parameter included air temperature, soil temperature, net short-wave and long-wave radiations, atmospheric pressure, the half-hourly change in atmospheric pressure, relative humidity, turbulent velocity, wind speed, and fluxes of sensible heat, CO_2, and water vapor. An important concept proposed in this study was that the correlation of the input and output did not represent the direct causal relationships but might use to explain the shared mechanisms between methane and other environmental variables or fluxes. The results revealed that the variables related to volatilization of water in the wetland are highly correlated with methane emissions, and the fluxes of latent heat and NEE were benefited for the models of CH_4.

In addition to the studies which used the ANN models to establish the relationship between the environmental factors and flux data in the specific sites, some studies used

doi:10.6342/NTU201703598

the ANN models to predict large scale variation of flux. Papale and Valentini (2003) used data of 16 forest ecosystems to build an ANN model in order to estimate the total NEE of the whole forest ecosystem at Europe scale. This study used the data observed in 1997, and the input parameters included the Normalized Difference Vegetation Index (NDVI), land cover characteristics, dew point temperature, maximum air temperature, minimum air temperature, mean air temperature, and the season. The estimation of overall European C uptake was about 0.47 GtC yr^{-1}, and was within the reasonable range reported by the estimation of previous studies.

Moffat et al. (2007) compared 15 techniques for estimating the missing flux data and concluded that ANN performed well among all the methods, and reproduced the half-hourly flux data better than nonlinear regressions. In this study, three kinds of ANN was selected to discuss, and all of them in four scenarios generated low root mean square error (RMSE) and high the coefficient of determination (R^2). The former is from 0.9 to 3.0 g C m^{-2} d^{-1}; the latter is between 0.36 and 0.92.

The input and output parameters of relative references were shown in **Table 1**. At present, most studies use three kinds of data as the input parameter (Papale and Valentini, 2003; Melesse and Hanley, 2005; Ooba et al., 2006; Morin et al., 2014). The first kind is the micrometeorological data, including the latent heat flux, sensible heat flux, and soil heat flux. Second, meteorological variables are also regarded as applicable to ANN; for instance, air temperature, soil temperature, net radiation, photosynthetic photon flux density (PPFD), vapor pressure deficit (VPD), precipitation, and horizontal wind speed. The last kind is attribute data such as the leaf area index (LAI), NDVI, land cover characteristics, and time series.

The CO_2 flux of ecosystems will influence the concentration of CO_2 in the

doi:10.6342/NTU201703598

atmosphere and lead to corresponding feedbacks in the biosphere and the atmosphere. Accordingly, the simulation results of ANN will be useful for understanding the spatial and temporal variations of CO_2 flux and the mechanism of C exchange.

doi:10.6342/NTU201703598

Table 1 Relative references which used the ANN model

Reference	Ecosystem	Data division	Data period	Input parameter	Output parameter
Van Wijk and Bouten (1999)	6 European forests	N/A	1996~1997 (DOY=150~250)	global radiation, temperature,and vapour pressure deficit.	NEE
Van Wijk, Bouten, and Verstraten (2002)	forested area in the Netherlands	N/A	1995~1996	global radiation, air temperature, wind speed, vapour pressure deficit, and precipitation.	NEE
Papale and Valentini (2003)	16 European forest	N/A	1997.1~ 1997.12	air temperature, dew point temperature, the maximum normalized difference vegetation index, the land cover type, and the season	NEE
Melesse and Hanley (2005)	Forest, grassland, and cropland in the US.	N/A	15 days	net radiation, latent heat, sensible heat, soil heat flux, air temperature, and soil temperature.	NEE
Ooba et al. (2006)	Japanese larch plantation in northern Japan.	N/A	2002.05~ 2002.09	Day of the year, time, photosynthetic photon flux density, vapor pressure deficit, net radiation, air temperature, wind speed, and latent heat flux.	NEE
Moffat et al. (2007)	6 different European forest	daytime /nighttime	2000~ 2002	Latent energy, global radiation, photosynthetic photon flux density, air temperature, soil temperature and water contents, relative humidity, precipitation, friction velocity, leaf area index.	NEE
Morin et al. (2014)	wetland in the US	N/A	2011.05~ 2013.12	air temperature, soil temperature, net short-wave and long-wave radiations, atmospheric pressure, the half-hourly change in atmospheric pressure, relative humidity, turbulent velocity, wind speed, and fluxes of sensible heat, C dioxide, and water vapor.	Methane flux

doi:10.6342/NTU201703598

1.5. The relationship of environmental factors and NEE

It has been proved that the variation of C flux is relative to the climate drivers in short scale (Stoy et al., 2009), but the characteristic of the exchange of C varies as a result of environmental factors including not only atmospheric conditions but also soil characteristics and plant properties, which affect both the quantity and quality of the organic matter (Brix, Sorrell, and Lorenzen, 2001; Inglett et al., 2012). Consequently, the C sequestration of ecosystems under different circumstances should be discussed individually. According to the Ramsar Convention, the total area of the low latitude wetland ecosystems accounts for about 70% of the global wetland area and distributes over Latin America and the Caribbean, Africa, and Asia area (Ramsar, 2015). In other words, taking all kinds of wetland ecosystems into consideration together may lead to inaccurate estimation of the global wetland C budget.

There have been some discussions on the relationship between the environmental factors and the CO_2 flux. The wetland studies of the variation of the net exchange of CO_2 indicated that the NEE is mainly influenced by the temperature, radiation, growing season, water, and LAI dynamics (Bubier et al., 1998; Bonneville et al., 2008; Han et al., 2015; Zhang et al., 2016). Zhang et al. (2016) concluded that the air temperature, shallow depths soil temperature, and the photosynthetic active radiation (PAR) are high-related to the CO_2 flux, respectively. By means of the proportion of variance (R^2) explained by a linear fit, the air temperature and the shallow depths soil temperature (0, 2, and 4 cm) showed the strong relationship to the CO_2 flux, with the R^2 were higher than 0.7. In the correlation results of PAR and CO_2 flux, the PAR was negative-related to the CO_2 flux; moreover, the correlation coefficient (R) of the growing season was much higher than that of the non-growing season. The flooding period of wetland ecosystems had impacts on the pattern of the NEE exchange, which led to the smaller

doi:10.6342/NTU201703598

range of the variation of the diurnal NEE. The effective photosynthetic leaf area may be reduced when the vegetation was submerged, so the exchange of CO_2 may be influenced (Han et al., 2015). In addition, the LAI is dependent on the variation of the temperature and precipitation. According to the study conducted by Bonneville et al. (2008), the LAI also has the strong relationship to the NEE because the uptake of the CO_2 has relations with the green photosynthesizing portion of plants.

1.6. Objectives

In this study, the estimation was focused on the CO_2 flux between the wetland ecosystem and atmosphere. An ANN model with the back-propagation algorithm was applied to quantify the CO_2 flux at three low-latitude wetland sites in East Asia and North America. Because ANN is more sensitive to the relationship between meteorological factors and C flux than other methods, this study will use the meteorological variables as the input parameters instead of the flux data and attribute data. Since the weather stations are more widespread than EC towers, this model is expected to apply to the places where is no tower established.

Meteorological variables were used as the input parameters to train the ANN to predict the CO_2 exchange. This model provides a quick, efficient, and highly accurate estimation, and can be conducted to estimate the dynamics of CO_2 flux where there is no direct in-situ flux measurement. The simulation capability is helpful to characterize the spatial/temporal variations in low-latitude wetland ecosystems and improve the quantification of low latitude wetland C budget.

The following is the purposes of this study. First of all, ANN models will be constructed to characterize the relationship between the meteorological variables and

doi:10.6342/NTU201703598

CO_2 flux in order to ensure this method is applicable in the wetland ecosystem. The data at each site will be used to train the ANN models and evaluate the model performance. The second objective is to verify the cross-site capability of ANN. The *in-situ* observations of low-latitude wetlands are limited, so this study attempts to apply the proposed ANN models to estimate the variations of CO_2 flux. The ANN models trained by a common reed wetland ecosystem will be used as reference models, and data from the other two wetlands will be estimated according to the reference models.

doi:10.6342/NTU201703598

2. Methods

2.1. Site descriptions

Three wetland ecosystem sites from two different countries were selected for this study. One site is situated in the southeast part of the Everglades National Park, Florida, United States; the others are located in the western part of the Guandu Nature Park, Taipei, Taiwan. Although the two areas are far from each other, they are similar in climate, vegetation, latitude, and elevation conditions. Data from these sites were used to train the ANN model to predict CO_2 flux.

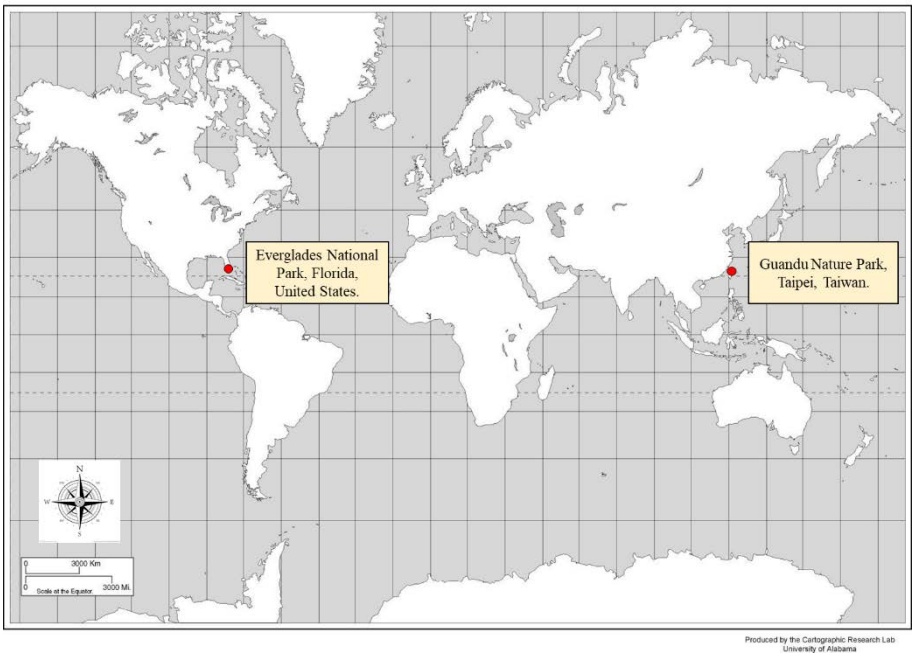


Fig. 1. The location of the study sites.

(Map source: http://alabamamaps.ua.edu/contemporarymaps/world/world/)

2.1.1. Guandu Nature Park Tower One (GDP-T1)

The Guandu tower one (25°7’N, 121° 28’E, hereafter GDP-T1) is registered on the Asiaflux network, and located in the western part of the Guandu Nature Park, which

doi:10.6342/NTU201703598

is in the northwest part of Taipei City. This area is part of the Guandu Nature Reserve, so it will not be interrupt by human activities. This site is at the junction of Tamsui River and Jilong River, and merely 10 km away from the estuary of Tamsui River; therefore, the characteristics of hydrological properties, energy properties, and biogeochemical cycles in this area will be influenced by the tidal and stream flow.

The GDP-T1 site is a grass marsh ecosystem which is dominated by para grass (*Brachiaria mutica*), with the canopy height about 1.2 m. Because this area is along with Tamsui River, the soil type is alluvium clay. In this site, the dry-wet seasonal variation is very mild. The climatic type is humid subtropical climate (Cfa). The annual mean temperature is 23.0 ℃, and the mean annual precipitation is about 2405.1 mm (Lee et al., 2015). The East Asian monsoon will influence the climate over different seasons, and typhoons have significant impact on this site during the summer especially.

2.1.2. Guandu Nature Park Tower Two (GDP-T2)

The second EC tower (hereafter GDP-T2) is also established in the Guandu Nature Park, and near the GDP-T1. The most environment condition is almost the same, including the topography, the soil type, and the climatic condition. Nevertheless, the predominant species here is common reed (*Phragmites australis*), and its canopy height is approximately 3.0 m. The two towers have different characteristics of the NEE and ecosystem respiration, and the difference shows that the two predominant species have different sensibilities to the environmental fluctuations. (Lee et al., 2015).

doi:10.6342/NTU201703598

2.1.3. Florida Everglades short hydroperiod marsh (US-Esm)

This station (25°26’16.43’’N, 80°35’40.56’’W) is registered in the Amerflux network, named Everglades short hydroperiod marsh (US-Esm) site, and co-located with the site managed by Florida Coastal Everglades Long Term Ecological Research (FCE LTER) program. The US-Esm site is located at the Everglades National Park, which is in the southern part of Florida.

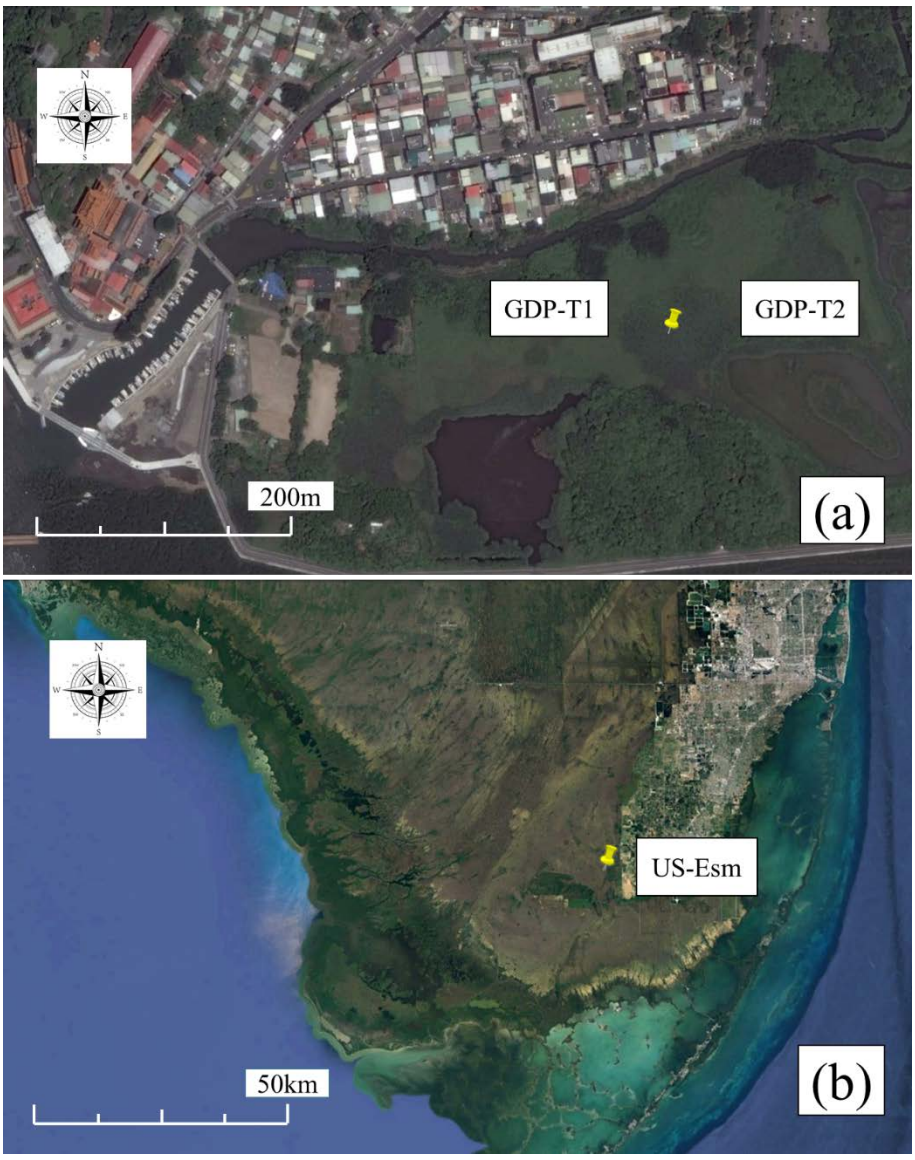
This site is a freshwater marsh which is seansonally inundated for about four to five months per year, and dominated by sawgrass (*Cladium jamaicense*) and muhly grass (*Muhlenbergia capillaris*) with a mean canopy height of 0.73 m (Schedlbauer et al., 2012). The major soil type is shallow murl soil (about 0.14 m), overlying limestone bedrock (Schedlbauer et al., 2010). The climatic seasonality here is moderate (Schedlbauer et al., 2010), and the climatic type is humid subtropical climate (Cwa) in the Köpen’s climate classification. The annual mean temperature is 23.8 ℃, with an average maximum of 29.4 ℃ and an average minimum 18.2 ℃ (NCDC, 2017). The annual precipitation is 1346 mm (NCDC, 2017). In this area, approximately 60% of rainfall occurs during the wet season (June to September), and about 25% is in the dry season (November to April) (Duever et al., 1994).

2.1.4. Brief summary

According to **Table 2**, the three sites had some similarities and dissimilarities. There were two reasons why these three sites were chosen in this study. First, the GDP-T1 and GDP-T2 site were established in the same place, so the conditions of environmental factors should be identical. If the datasets from GDP-T1 site can be simulated accurately by the GDP-T2 model, it is more likely that the GDP-T2 model

doi:10.6342/NTU201703598

can apply to other ecosystems where the environmental conditions were similar. Secondly, the environmental conditions in US-Esm site were also similar in terms of latitude and Köpen's climate classification. However, it is important to note that the difference of mean annual precipitation between the GDP-T2 and US-Esm was significant, thus the comparison between the two places might be attributable to the difference in the precipitation in terms of the climatic condition. Besides, the datasets from the US-Esm site were observed in 2010. It might be helpful for examining the capability of the cross-site data simulation.



doi:10.6342/NTU201703598

Fig. 2. The location of the towers.

(a) The Guandu Nature Park tower one (GDP-T1) and tower two (GDP-T2). (b) The Florida Everglades short hydroperiod marsh (US-Esm)

Table 2 The details of the study sites

Site name	GDP-T1	GDP-T2	US-Esm
Location	(25°7'N, 121° 28'E)		(25°26'16.43''N, 80°35'40.56''W)
Dominant Species	para grass	common reed	sawgrass and muhly grass
Canopy height	1.2 m	3.0 m	0.73 m
Soil type	alluvium clay		shallow murl soil
Climatic type	humid subtropical climate (Cfa)		humid subtropical climate (Cwa)
Annual mean temperature	23.0 ℃		23.8 ℃
Mean annual precipitation	2405 mm		1346 mm
Data period	2014		2010

2.2. Study design

There were four stages in this study. In the first place, three scenarios were set to discuss whether the data division is helpful for the model training. The three scenarios were including the all data group, the daytime data group, and the nighttime data group. According to the theory of EC method, the nighttime CO_2 flux is hard to measure because the friction velocity is too weak. The unapparent relationship between the meteorological variables and CO_2 flux may lead to the worse model performance, so the datasets will be grouped into daytime data and nighttime data to examine the results.

In this stage, the observed data of each group from each study site was randomly divided into to two parts. The first part was used to train the ANN models, and the other part was for the simulation phase and calculating the model performances. The model

doi:10.6342/NTU201703598

performances were evaluated by the three statistical indices mentioned in Section 2.7. Next, this study used the *R* and the statistical test to analyze the relationship between the meteorological variables and the CO_2 flux. Through these results, the relationships between the meteorological variables and the CO_2 flux could be more obvious.

After the training and simulation phase, the models from the GDP-T2 were chosen as the main models to calculate the cross-site prediction. Because the dominant species in the GDP-T2 area was the common reed, which is the most widespread and representative species in the fresh wetland ecosystem (Clevering and Lissner, 1999; Park, Hong, and Kim, 2013; Vymazal, 2013; Lee et al., 2015), it is considered that the models trained by the GDP-T2 data are applicable to a great deal of areas where the environmental conditions are similar. Besides, if the capability of the cross-site prediction was proved, the models can be applied to more wetland ecosystems.

The last phase was to compare the relationships between the main model and the other two sites, respectively. Same as the second phase, the ratios between the meteorological variables and the CO_2 flux were testified by the *R* and the statistical test to discuss the differences in the two conditions (GDP-T1 vs. GDP-T2 and US-Esm vs. GDP-T2).

2.3. Artificial Neural Network (ANN)

The ANN technique, which imitates the concept of working network of biological brain, is a kind of machine learning. It calculates a great deal of data through an empirical non-linear regression model. ANN extract the relationship between the input and the output variables from observed datasets. The structure of ANN consists of three parts including neurons, weights, and layers. Neurons are the constitutive unit to calculate. In the computing step, a neuron will receive a set of input (x) and weight (w).

doi:10.6342/NTU201703598

First, this set will be summarized by the summation function; then, the result (z) will be transformed by the activation function into the output of this neuron (**Fig. 3**). Neurons may be connected to many other neurons by weights, which can be enforcing or inhibitory on the activation state of the connected neurons. According to the characteristic of calculation, neurons will form different layers. Typically, ANN are composed of multiple layers including an input layer, an output layer, and hidden layer(s).

Because the meteorological variables and the C flux are continuous variable data, an ANN model with the back propagation algorithm was built and used to estimate in this study. The back propagation algorithm is a kind of supervised learning to train the model. The training process included forward pass and backward pass. In the forward pass phase, the input data was imported into the model in order to calculate a simulation value as the output of the model this time. Therefore, this model started the backward pass phase. This value was compared to the real output data. The difference between the real output and simulated output was called error. Then, the ANN model adjusted the weights according to the error of this time to minimize the error of next time. Through the two process, the training process stopped when the error was smaller than the threshold which the investigator set before the model was carried out.

doi:10.6342/NTU201703598

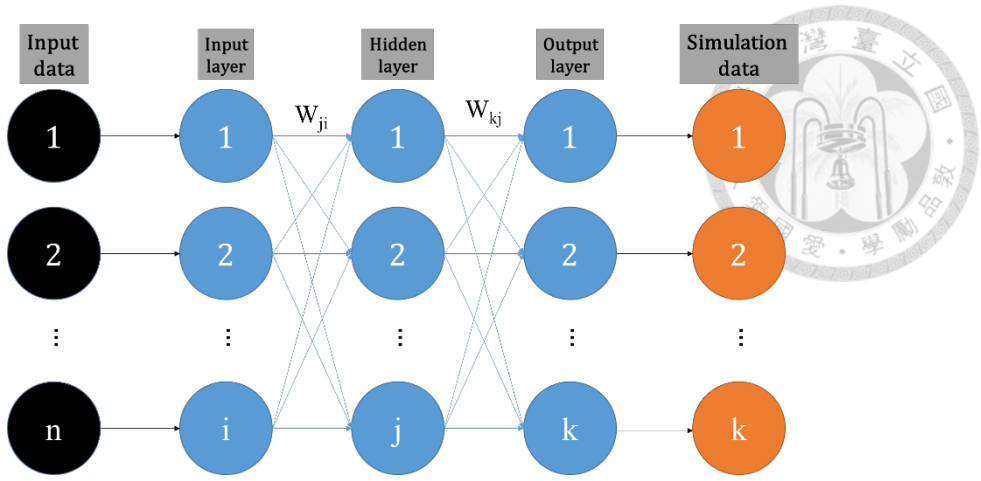


Fig. 3. The structure of ANN.

The black and orange neurons represent the input data and simulation data, respectively. The blue neurons are the computing units of this model. In the input layer, the number of neurons depends on the number of the input data variables; in the same way, the variable number of simulation data will decide the neuron number of the output layer. It is assumed that there are i, j, and k neurons in the input, hidden, and output layer, respectively. Weights link neurons between different layers.

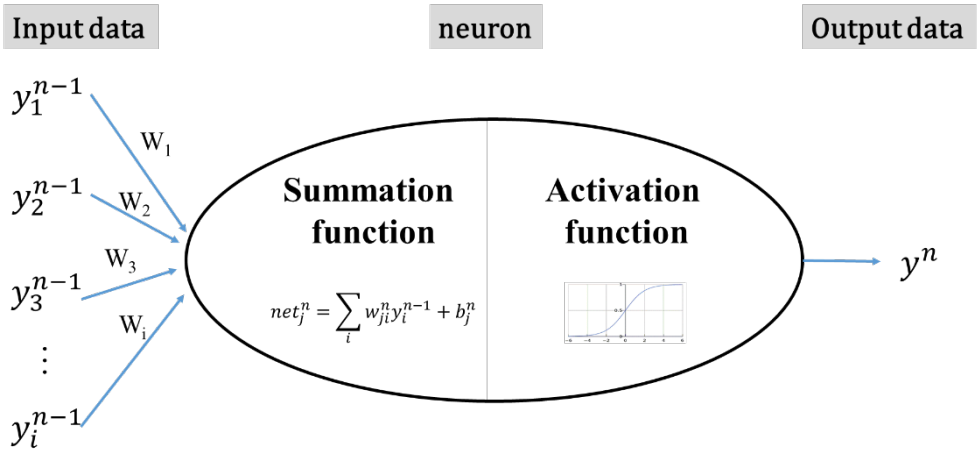


Fig. 4. The operation process in each neuron.

In the computing step, y_i^{n-1} represents the input data from the previous layer, and w_i represents the corresponding weight. This set of data will be calculated by the summation function and the activation function chronologically. Finally, y^n is the result of this neuron, and it will be sent to the next layer.

doi:10.6342/NTU201703598

There was only one hidden layer in all the models in this study. Too many hidden layers make the model too complicated so that the model is possible to be overfitting to some specific data (Rowbottom, Webb, and Oldham, 1999). The activation function is Log-Sigmoid transfer function (Eq. (1)) and is described using **Fig. 5.**

$$f(x) = \frac{1}{1+e^{-x}} \quad \text{.....(1)}$$

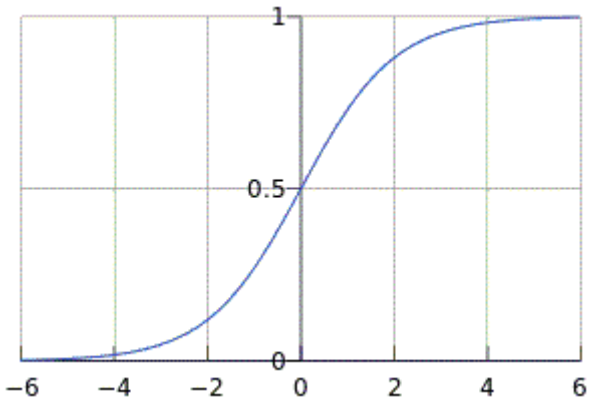


Fig. 5. The Log-Sigmoid transfer function used in the ANN models.

The number of neuron in the hidden layer was decided based on the constructive algorithm, which Kwok and Yeung (1997) proposed. The results of each model were in **Table 3,** and the RMSE curve figures were in the **Appendix I** for the purpose of preventing the influence of the random initial weights, every parameter setting was run in ten times to analyze.

doi:10.6342/NTU201703598

Table 3 The numbers of neuron in the hidden layer of each model

Group / Site name	All data	Daytime data	Nighttime data
GDP-T1	7	5	7
GDP-T2	5	4	3
US-Esm	7	5	6

2.4. Input and output variables

Ooba et al. (2006) demonstrated the ANN model with the variables which were consist of energy balance terms, and ended up the best performance among all the input type. In this study, the input parameters excluded from the flux data such as latent heat flux (W m^{-2}) and photosynthetic photon flux density (PPFD, μmol m^{-2} s^{-1}). Some studies revealed that air temperature, water and the PPFD were the important variables which will influence the NEE of wetland ecosystems (Jimenez et al., 2012; Lee et al., 2015). However, the latent heat flux and PPFD will not be measured in the general weather stations, so the VPD and net radiation, rather than the latent heat flux and PPFD, were used to provide the relevant information in this study. The stationarity of the background environment will also influence the exchange of materials between the atmosphere and the ecosystem (Baldocchi, 2001), therefore the wind data was taken into consideration. Similarly, the horizontal wind speed data was used in this study because of the data availability in the general weather station.

In brief, the chosen four variables were air temperature (hereafter Tair, ℃), net radiation (hereafter Rn, W m^{-2}), vapor pressure deficit (hereafter VPD, kPa), and horizontal wind speed (hereafter wind, m s^{-1}) as the input. The CO_2 flux (hereafter Fc, μmol m^{-2} s^{-1}) was the output.

In this study, half-hourly data from January to December 2014 in GDP-T1 and

doi:10.6342/NTU201703598

GDP-T2 was used. The data from US-Esm was from the January to December 2010. All of the variables from the EC towers were averaged every 30 minute, and 17520 data sets in total from each tower.

2.5. Data pre-processing

In order to choose the data set which had an obvious relationship between the meteorological variables and CO_2 flux, the data sets were selected based on two principles. The first and most important thing is that there was no missing data in the selected data set. Second, the variables should distribute over a reasonable range, and the procedure followed the standard protocol used in other flux studies (Baldocchi et al. 2001).

Three scenarios were designed to analyze the differences between these groups. The first condition was taking all the data into consideration, called the all data group (hereafter AL group). The model trained by this group might be more comprehensive on account of the diversification of data, or might underperform because the relationship is equivocal. The others were the daytime data group (hereafter DT group) and the nighttime data group (hereafter NT group) owing to the different properties of meteorological variables and CO_2 flux. The atmospheric condition is unstable because the radiative heating in the daytime, it is helpful for the development of turbulent. In this situation, the mass is easy to exchange in vertical direction, and the variation of CO_2 flux is easier to be measured. On contrary, the atmospheric condition is more stable in the nighttime. In this study, the division of data into daytime data and night data relied on the value of incoming shortwave radiation.

After the process, the data sets were normalized to build up the ANN models. The

doi:10.6342/NTU201703598

variations of climate conditions and CO_2 flux are dynamic, so the minimum and maximum of data from specific temporal intervals may not be representative enough. Furthermore, in order to make the relationship between the meteorological variables and CO_2 flux more obvious, the normalized values of each variable ranged from 0.15 to 0.85. The extrema were calculated by the maximum and minimum of datasets which had been pre-processed. Thus, extremums of each group in the normalization phase were shown in **Appendix I**.

After the data pre-processing, the numbers of the effective data was shown in **Table 4**. The the effective data was divided into training data and simulation data randomly, and the numbers of training and simulation data were shown in **Table 5**. In the selected data, depending on the total amount of each scenario in each site, about 300-800 datasets were used to train the model and the rest was for the simulation.

doi:10.6342/NTU201703598

Table 4 The numbers of effective datasets in each site

Site name / group	GDP-T1	GDP-T2	US-Esm
AL	4,982	3,232	8,228
DT	3,014	1,791	5,180
NT	1,968	1,441	3,048

Table 5 The numbers of the training data and the simulation data in each site

	GDP-T1		GDP-T2		US-Esm	
Number of data	training	simulation	training	simulation	training	simulation
AL	982	4,000	732	2,500	728	7,500
DT	514	2,500	791	1,000	680	4,500
NT	968	1,000	941	500	548	2,500

2.6. Model comparison indices

The statistical indices were used to evaluate the performance of the models in this study. The performance of the model depended on the comparison of the observed and the predicted CO_2 flux data. The indices were including the R, the RMSE, and the mean absolute percentage error (MAPE). Let n be the sample number. The $y_{o,i}$ and $y_{p,i}$ represented the i^{th} observed CO_2 flux data, and the predicted CO_2 flux data, respectively.

- Correlation coefficient (R)

$$R = \frac{\sum_i^n (y_{p,i} - \overline{y_{p,i}})(y_{o,i} - \overline{y_{o,i}})}{\sqrt{\sum_i^n (y_{p,i} - \overline{y_{p,i}})^2 \sum_i^n (y_{o,i} - \overline{y_{o,i}})^2}} \quad \text{.....(2)}$$

which the $\overline{y_{p,i}}$ and $\overline{y_{o,i}}$ represented the average of $y_{p,i}$ and $y_{o,i}$. The value of this index ranged from -1 to 1. This index indicates the correlation condition between

doi:10.6342/NTU201703598

the observed data and the predicted data. In this study, the value of *R* should range from 0 to 1. When the value is smaller than 0, it represents the prediction is unreasonable. Then, the correlation analysis and the significant test of each group was calculated and shown in **Table 4**. The significance level of correlation analysis is 0.05 ($\alpha = 0.05$), and the hypothesis of significant test is as following:

$$\begin{cases} H_0: \text{the variables } \mathbf{isn't} \text{ related to each other significantly.} \\ H_1: \quad \text{the variables } \mathbf{is} \text{ related to each other significantly.} \end{cases}$$

- Root Mean Square Error (RMSE)

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_{o,i}-y_{p,i})^2}{n}} \quad \text{.....(3)}$$

The smaller RMSE is, the smaller error between the observed data and predicted data is. This index has the same unit with the original data, so it doesn't distribute over a specific range. RMSE is more sensitive to the outliers than other indices; therefore, the MAPE was used to prevent the influence of the scale of original data.

- Mean absolute percentage error (MAPE)

$$\text{MAPE} = \frac{\sum_{i=1}^{n} 100\times|(y_{o,i}-y_{p,i})/y_{o,i}|}{n} \quad \text{.....(4)}$$

MAPE represents the error as a percentage of the observed data. This index is not influenced by the scale of the observed and predicted data, and is good for the comparison of different models. Based on the theory of Lewis (1982), when the MAPE is small, the model performs well. The criteria of the MAPE value and the corresponding interpretation is in **Table 6**.

doi:10.6342/NTU201703598

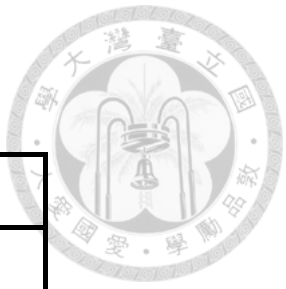


Table 6 The criteria of MAPE

Criterion interval	**Forecasts capability**
<10%	highly accurate
10% ~ 20%	good
20% ~ 50%	reasonable
>50%	inaccurate

doi:10.6342/NTU201703598

3. Results

3.1. The model performance in each site

The model performances in the simulation phase are shown in **Table 6**, and all figures of the simulation result are in **Appendix III**. The best result occurred in the AL group of the GDP-T1 site. However, if the AL group were divided into the daytime part and the nighttime part to calculate the statistical indices, the results changed consequently (shown in **Table 7**).

Because the AL data included daytime and nighttime data, and were normalized with the same extremums, so the characteristics of daytime data and nighttime were concentrated on some part of the range between [0, 1] after the normalization. If the relationship between meteorological variables and CO_2 flux of daytime/nighttime data was quite apparent, it might influence the overall results of the AL group. In other words, the model performance of simulation result could not be determined by the results calculated by the whole data of the AL group.

Although the difference was quite small, the best result occurred in the daytime group of the GDP-T1 in terms of *R*. If the results of the undivided group (*i.e.* AL group) were partitioned into daytime part and nighttime part, the RMSE values were higher than the divided groups (*i.e.* DT / NT group) and the *R* values became lower.

In the **Fig. 6**, the black line represented the observed data, and the red line was the results of the model simulation. The simulation of the ANN model not only depicted the trend of the variation of the CO_2 flux, but also predicted the values accurately. Although some peaks were not as fitting as other data, the major portion was highly accurate. In other words, if the environmental conditions are similar to the historical conditions, this model can provide the considerably reliable estimation results.

doi:10.6342/NTU201703598

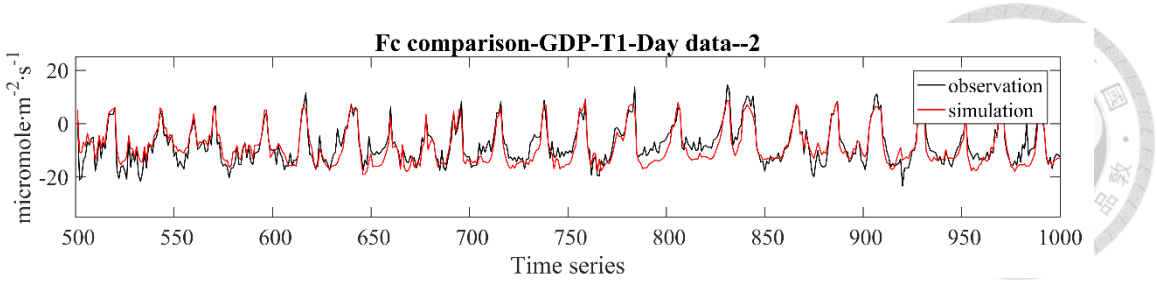


Fig. 6. The simulation result of the AL group of GDP-T1 from selected time period.

The total result in the three site proved that the meteorological variables have the capability to predict the variation of CO_2 flux. In these sites, the accuracy of the AL groups was the best among the three scenarios. The diversification of data may be helpful for the model to be more comprehensive, and the same condition can be observed in **Table 3**. The undivided groups had to use more calculation resource to train the model, and the numbers of neuron in the hidden layer were more than other groups.

Notably, all the MAPE values in each site were less than 20% in **Table 7**, which represented that the forecasts capabilities of the models were good, but the performances of the three nighttime groups were poor in terms of R, especially the result of the GDP-T2 site. In the **Fig. 7**, though the simulation data estimated the rough variation of CO_2 flux more and less, it didn't display the fluctuation of the CO_2 flux. In addition, most results were underestimated. The other figures of simulation in each site were in **Appendix IV**.

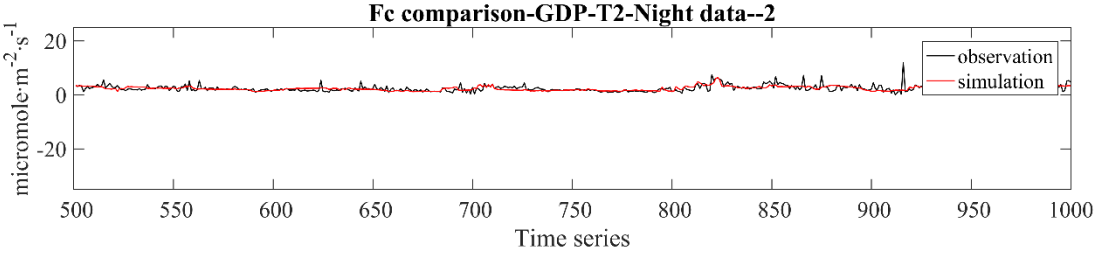


Fig. 7. The simulation result of the NT group of GDP-T2 from selected time period

doi:10.6342/NTU201703598

Table 7 The model performance of each group in the simulation phase

site	group	MAPE (%)	RMSE (μmol · m^{-2} ·s^{-1})	R (unitless)
GDP-T1	AL	7.8%	3.1369	0.94*
	DT	10.9%	3.6240	0.89*
	NT	18.2%	2.0354	0.65*
GDP-T2	AL	6.4%	2.9628	0.91*
	DT	9.1%	3.1602	0.87*
	NT	8.6%	1.5823	0.59*
US-Esm	AL	5.5%	0.9283	0.81*
	DT	7.2%	1.0117	0.60*
	NT	9.2%	0.7227	0.62*

* represents that the result reaches a statistically significant level (*p*-value <0.05).

Table 8 Comparison of the model performance in daytime and nighttime interval of each group in the simulation phase

Time period	Data group	GDP-T1		GDP-T2		US-Esm	
		RMSE	R	RMSE	R	RMSE	R
Daytime	AL_daytime	3.6019	0.89*	3.3856	0.84*	1.0168	0.60*
	DT	3.6240	0.89*	3.1602	0.87*	1.0117	0.60*
Nighttime	AL_nighttime	2.2394	0.53*	2.3345	0.28*	0.7571	0.55*
	NT	2.0354	0.65*	1.5823	0.59*	0.7227	0.62*

* represents that the result reaches a statistically significant level (*p*-value <0.05).

(Unit of RMSE is μmol · m^{-2} · s^{-1}, and *R* is unitless)

3.2. The data distribution at each site

Using the single variable to build the ANN model and calculate the *R* between the observed Fc and simulation Fc and the results were shown in **Table 9**. Namely, all results were significantly related to the CO_2 flux except the wind of daytime data at the GDP-T1 and US-Esm site.

The trends in the order of the magnitude of *R* in the three groups were similar in

doi:10.6342/NTU201703598

all sites. In the AL group situation, the Rn was the predominant variable to the simulation of Fc among the four inputs at all the sites. The results of *R* revealed that the results of Rn was highly-correlated to the Fc, and all the values were less than -0.7 in these sites. The second and the third variable at all the sites were quite approximate. At GDP-T2 and US-Esm, the second and the third variable were VPD and Tair, respectively. On contrast, the *R* of Tair was higher than that of VPD at GDP-T1. The absolute values of wind were the smallest at all the sites similarly. It showed that the wind data was modestly correlated to the Fc data. In terms of the data of AL group, the capability of predicting Fc data in each result were more or less look-alike.

Among the DT groups, the sequential orders of *R* were almost the identical with the results of AL group, but there was a little bit different in the second and the third variable. The second variable was VPD at GDP-T1 site and GDP-T2 site, but the result of the US-Esm site was opposed to the other sites. The lowest values were the *R* of the wind at all the sites. Notably, the results of the GDP-T1 and the US-Esm site didn't reach the significant level which meant those simulation results using the daytime wind data as the input variable didn't correlate to the observed data. Finally, in the NT group, there were some differences of the orders in these sites. The orders of the GDP-T1 and the GDP-T2 were totally the same, and the values of each input variable were close to each other roughly. The values of Tair were the highest, and those of the wind data were the lowest. The characteristics of the US-Esm were entirely different. Unlike the other two data groups, the highest *R* at the US-Esm site was the Rn. However, the Tair was the lowest one at the US-Esm site.

Then, the original data distributions were used to analyze the characteristics at each site. Data were divided into five groups, and the average and mode of the Fc were calculated in each group. In **Fig. 8 -10**, the average and mode of each group were shown

doi:10.6342/NTU201703598

on the blue line with stars and red line with circles, respectively. The error bars were defined by twice standard deviation of each group, and stretch out from the average of each group. The horizontal axis represented the meteorological variables, and the vertical axis represented the CO_2 flux.

doi:10.6342/NTU201703598

Table 9 The result of the *R* with Fc in each site.

Site name	**GDP-T1**		**GDP-T2**		**US-Esm**	
AL						
input	***R***	**order**	***R***	**order**	***R***	**order**
Tair	0.5743*	2	0.4249*	3	0.3408*	3
Rn	0.9333*	1	0.9043*	1	0.7860*	1
wind	0.2590*	4	0.2605*	4	0.3175*	4
VPD	0.5723*	3	0.4873*	2	0.4567*	2
DT						
input	***R***	**order**	***R***	**order**	***R***	**order**
Tair	0.3360*	3	0.3836*	3	0.3449*	2
Rn	0.8575*	1	0.7872*	1	0.5286*	1
wind	0.0317	4	0.2102*	4	0.1258	4
VPD	0.3906*	2	0.4156*	2	0.3156*	3
NT						
input	***R***	**order**	***R***	**order**	***R***	**order**
Tair	0.5729*	1	0.5541*	1	0.2818*	4
Rn	0.3576*	3	0.3819*	3	0.5308*	1
wind	0.2352*	4	0.2759*	4	0.3142*	3
VPD	0.4701*	2	0.4748*	2	0.4349*	2

* represents that the result reaches a statistically significant level (*p*-value <0.05).

3.2.1. All data

In **Fig. 8**, the blue line with star represented the average of Fc of each group, and the red line with circle represented the mode of Fc of each group. In general, the CO_2 flux and the meteorological variables varied inversely. With the rise of the values of meteorological variables, the values of CO_2 flux decreased. The scale of vertical axis of the US-Esm site was smaller than the other sites because of the variation of Fc at the US-Esm site. In most conditions, the values of mode of the GDP-T1 and the GDP-T2 site were higher than the averages. In other words, some strong extreme values, which were smaller than the modes, influenced the averages of those groups. This condition could be found in the results of wind and VPD. On contrary, the values of mode of the

doi:10.6342/NTU201703598

US-Esm site were lower than the average, and this conditions occurred in the results of Tair and VPD obviously. The more apparent the difference between the mode and the average were, the larger the length of the error bar was. It also coincided with the impact of the extremums. That is to say, the extremums might influence the overall conditions of the AL group.

3.2.2. Daytime data

The characteristics of daytime data were more obvious than the AL group. The data distributions in the figures of the GDP-T1 and the GDP-T2 were more centered, and the lengths of error bar became a little bit smaller than those of AL group. In the results of Tair and Rn (**Fig. 9**), the relationship was clearer than the AL group. The larger the values of meteorological variables were, the smaller the averages of Fc were. Because the higher values of Tair and Rn might be helpful for the photosynthesis of vegetation, so the C uptake increased. There were similar conditions in the results of VPD. However, in the results of wind (**Fig. 9**), this relationship was still not obvious. The relationship between the wind and the Fc might not be linear, or the variation was hard to observe in these figures of wind.

doi:10.6342/NTU201703598

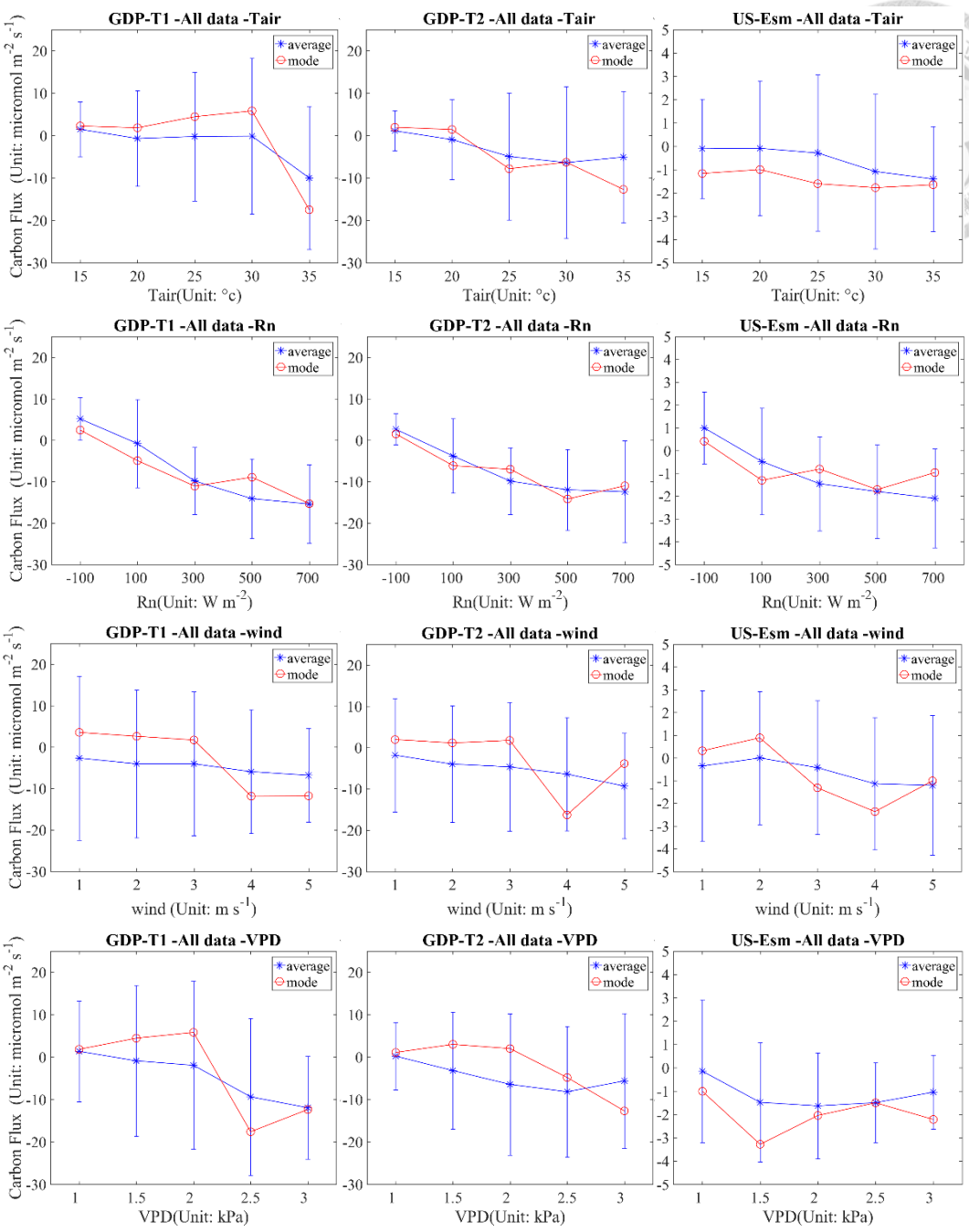


Fig. 8. The data distributions of AL group at each site

3.2.3. Nighttime data

In the nighttime group, the trends of the average were consistent with that of the modes, and the two kinds of values were quite close to each other than AL and DT data. In addition to the influence of data division which made the characteristics more

doi:10.6342/NTU201703598

apparent, it might be influenced by the original data scale of the data. Because the variation of Fc in the nighttime data was small, the influences of extrema values might also be slight. The variations of four kinds of meteorological variables in the GDP-T1 and GDP-T2 were similar. When the meteorological variables increased, the values of Fc also increased. However, the relationship in the figures of US-Esm were a little bit different from other sites. That responded to the results shown in **Table 12**. In the figures of Tair in the GDP-T1 and GDP-T2, when the values of Tair increased, those of Fc increased, too. The variation in US-Esm was nearly a horizontal line, and the right tail even went down slightly. Similar conditions could be found in the other kinds of meteorological variables.

doi:10.6342/NTU201703598

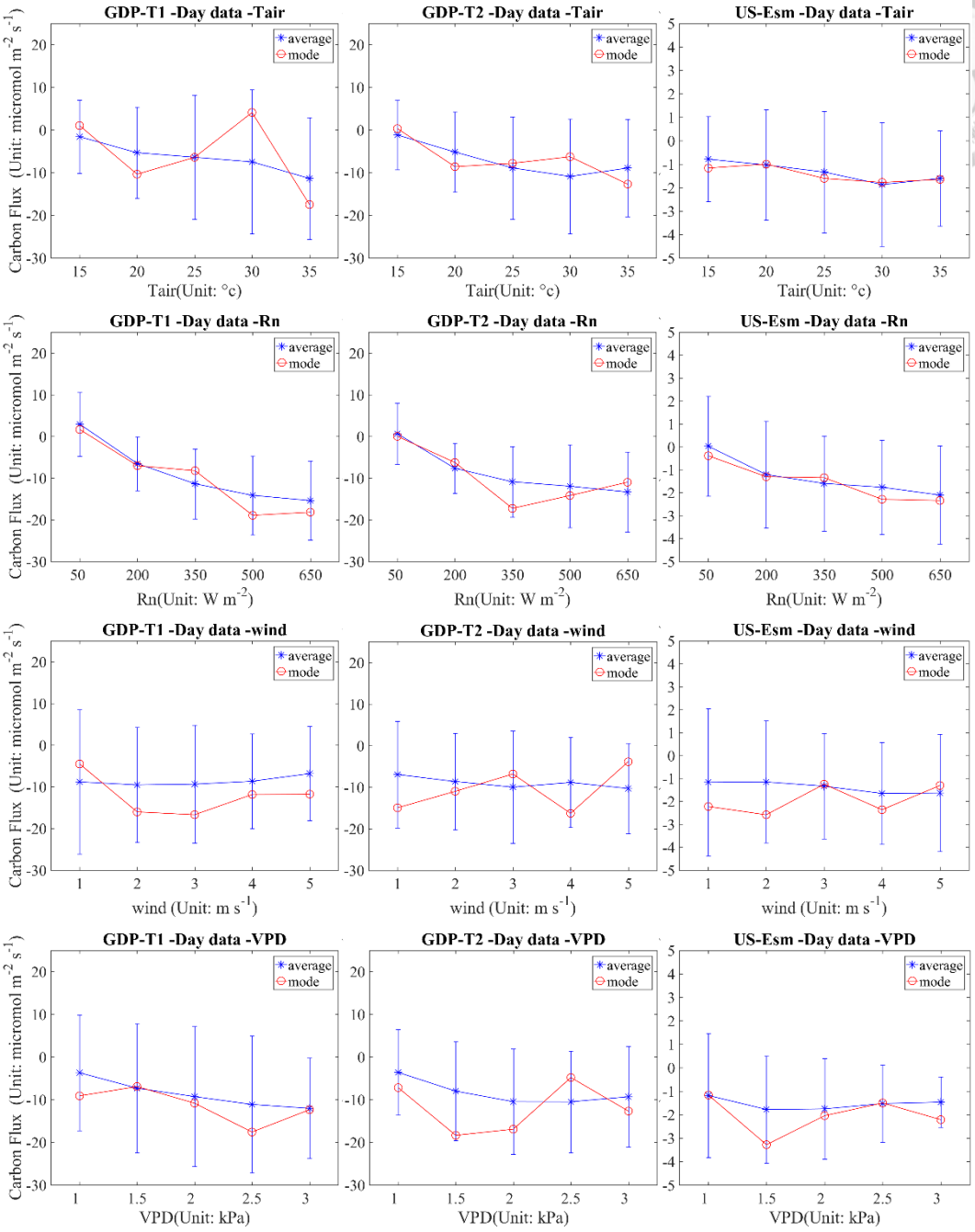


Fig. 9. The data distribution of DT group of each site

doi:10.6342/NTU201703598

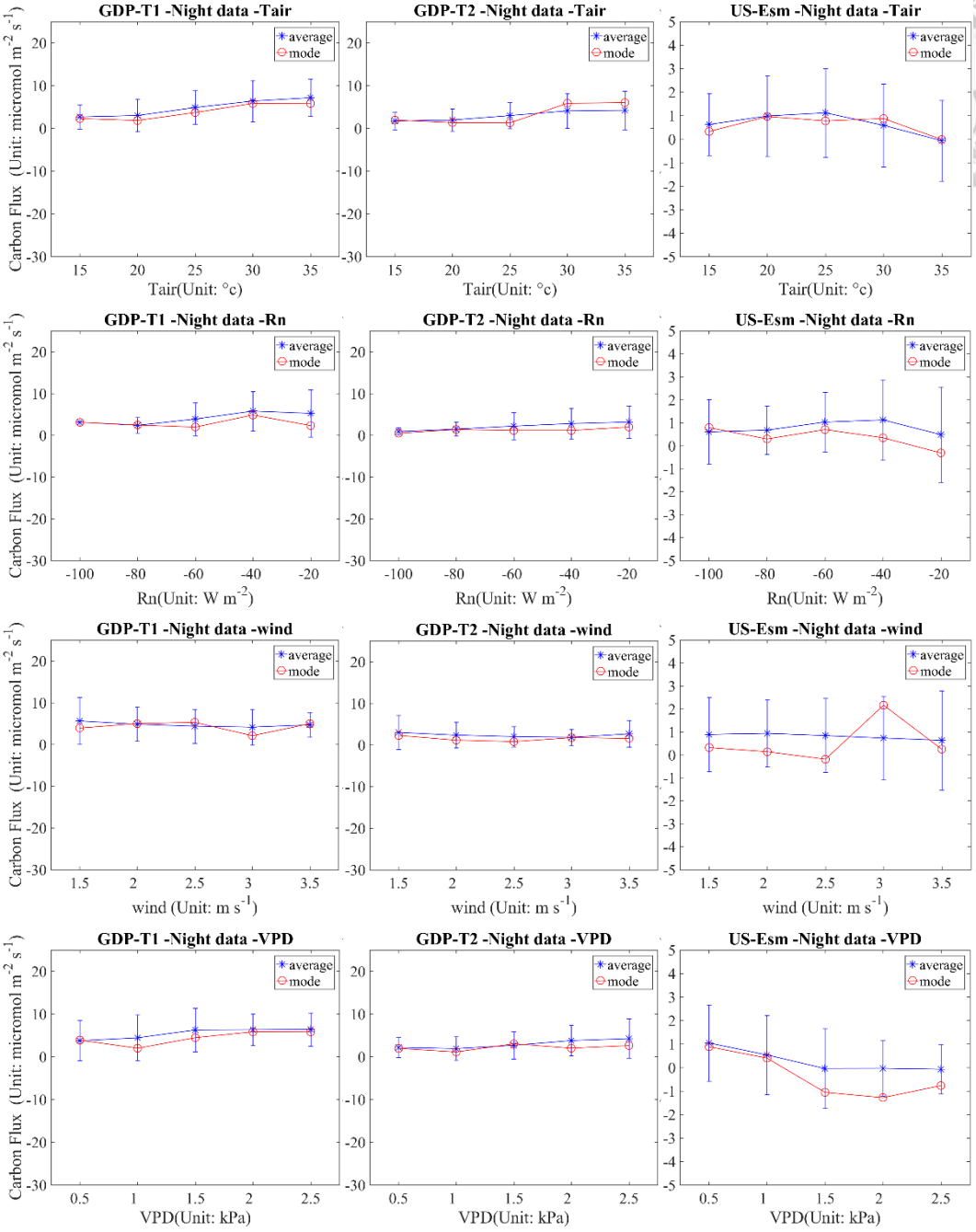


Fig. 10. The data distribution of NT group of each site

doi:10.6342/NTU201703598

3.3. The capability of cross-site prediction

The models trained by the data from the GDP-T2 site was used as the reference models to verify the capability of cross-site prediction. The model performances of the cross-site prediction in both sites were good or reasonable in terms of the MAPE, and the other indices were also good.

In the same way as the divided model performances of the simulation results at each site, the better results occurred in the DT group at GDP-T1 sites, and in general, the results of GDP-T1 were better than the US-Esm in terms of the *R* values (shown in **Table 10**). Because the GDP-T1 is quite near the GDP-T2, the climatic condition is more similar. It is reasonable that the reference models predict the data from GDP-T1 more accurately. In the results of the GDP-T1 site, the difference of *R* was significant in both daytime period and nighttime period. The *R* of DT group was almost twice of that of the daytime data of AL group, and the *R* of NT group was up to six times larger than that of the nighttime data of AL group. Moreover, it is noted that the results of *R* in the condition of nighttime data of the US-Esm site were different from other results. The simulation result of NT model was worse than that of AL model in the US-Esm data.

doi:10.6342/NTU201703598

Table 10 The simulation results of GDP-T1 and US-Esm in the GDP-T2 model

site	group	MAPE (%)	RMSE ($\mu mol \cdot m^{-2} \cdot s^{-1}$)	*R* (unitless)
GDP-T1	AL	17.8%	7.4267	0.79*
	DT	17.6%	6.1745	0.73*
	NT	33.6%	2.7298	0.55*
US-Esm	AL	14.1%	2.1281	0.72*
	DT	29.9%	3.6070	0.38*
	NT	47.3%	2.8695	0.28*

* represents that the result reaches a statistically significant level (*p*-value <0.05).

Table 11 Comparison of the model performance in daytime and nighttime interval of each group in the simulation phase

Group	Time period	GDP-T1		US-Esm	
		RMSE ($\mu mol \cdot m^{-2} \cdot s^{-1}$)	R (unitless)	RMSE ($\mu mol \cdot m^{-2} \cdot s^{-1}$)	R (unitless)
AL	Daytime	8.5967	0.47*	1.9061	0.36*
DT		6.1745	0.73*	3.6070	0.38*
AL	Nighttime	5.1425	0.09*	2.4599	0.50*
NT		2.7298	0.55*	2.8695	0.28*

* represents that the result reaches a statistically significant level (*p*-value <0.05).

3.4. The comparison of data distribution between different site

The variation of Fc in these three sites were quite different, the data used in this study was calculated the average hourly C flux (shown in **Fig. 11**). In this figure, the differences between the three sites were obvious. The black line with circle represented the C flux of GDP-T1, the red line with cross represented that of GDP-T2, and the blue line with star represented that of US-Esm. The curves of GDP-T1 and GDP-T2 were similar, but the range of the GDP-T1 was larger than that of GDP-T2. By comparison, the variation of the US-Esm was quite gentle and unapparent, and the shape of this

doi:10.6342/NTU201703598

curve was nearly flat.

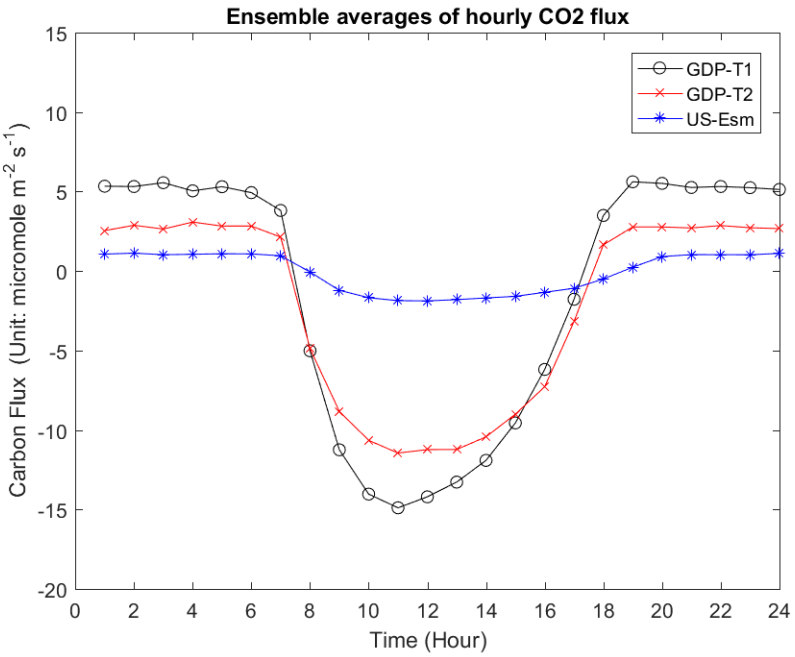


Fig. 11. Ensemble averages of hourly CO_2 flux of the three sites.

In order to compare the relationship between the meteorological variables and CO_2 flux, the results of the *R* were used to explain the characteristics of CO_2 flux variation. Based on the ratio of the meteorological variables and the CO_2 flux, the values of *R* (shown in **Table 12**) were calculated to interpret the comparisons between the three sites. According to **Table 12**, all the *R* values of GDP-T1 and GDP-T2 were higher than those of US-Esm and GDP-T2. All the results of AL group and DT data results related to the data of GDP-T2 significantly, but there were some values of US-Esm site didn't in the NT data group.

In the results of AL group, the values of Rn was the highest in both groups, and the value of wind was the lowest. All the *R* values of GDP-T1 were highly correlated to the ratio of GDP-T2, and the values were higher than 0.8. In the results of US-Esm site, only the result of Rn was higher than 0.8. The results of Tair and VPD were

doi:10.6342/NTU201703598

moderately correlated (R=0.6654 and 0.4860, respectively), and that of wind was modestly correlated (R=0.2635).

If the data was divided into daytime and nighttime group, the results of R changed more or less. The trend was similar to the AL group in the DT group, but there were some slight differences in the values. All the results of the GDP-T1 were smaller than those in the AL group, they still in the range of highly-correlated values ($R \geq 0.7$). The values of the US-Esm also reduced, and the highest value in the DT group was the ratio of Tair. In the NT group, the results of the GDP-T1 were totally opposite to the other two group. The highest value was the ratio of wind, and the lowest one was the ratio of Rn. In the results of the US-Esm in the NT group, the ratio of Tair was the highest (R=0.4059), and the others didn't related to the data of the GDP-T2 significantly. Because three of four ratios of the US-Esm site did not relate to those of the GDP-T2, the characteristics between the meteorological variables and CO_2 flux were different in the nighttime data. Therefore, it might be the reason why the result of cross-site simulation in the NT model was worse than that of the AL model.

Table 12 The R of each site

group	All data (AL)		Daytime data (DT)		Nighttime data (NT)	
Site / ratio	GDP-T1	US-Esm	GDP-T1	US-Esm	GDP-T1	US-Esm
Fc / Tair	0.9568*	0.6654*	0.9249*	0.4676*	0.3506*	0.4059*
Fc / Rn	0.9568*	0.8083*	0.9391*	0.4288*	0.2392*	-0.1662
Fc / wind	0.8599*	0.2635*	0.7873*	0.1201*	0.7573*	0.0550
Fc / VPD	0.9042*	0.4860*	0.8723*	0.2334*	0.2994*	0.0958

* represents that the result reaches a statistically significant level (p-value <0.05).

doi:10.6342/NTU201703598

4. Discussion

4.1. The data period in each site

The time period of the data in these sites was different. The data period of GDP-T1 and GDP-T2 was from 2014, and that of US-Esm was 2010. Because both the GDP-T1 site and the GDP-T2 site were in the Guandu Nature Park, the data period of these two sites should be the same in order to discuss the difference of the Fc variation between the two sites in the similar climate condition. The better condition was to choose the same period in the US-Esm site to compare, but the data in 2010 was better than another period on the basis of the data availability. The 2010 data is more complete, which provided various conditions to test the ability of the model performance. Therefore, 2010 data in the US-Esm was used in this study.

4.2. The advantage of the meteorological variables as the input data

Most of the previous studies used the micrometeorological data (e.g. the sensible heat flux, latent heat flux, soil heat flux) or the parameters which are hard to collect (e.g. LAI and NDVI). Although these variables can provide the highly relevant information about the variation of CO_2 flux, the costs of the data collection were considerable in aspects of time and money. Compared to other kinds of data, the meteorological variables are easier to be obtained from the weather stations around the world, and the resolution of the data will be higher relatively. Therefore, even the results of simulation of the models which using the meteorological variables as the input data were not very accurate, it was still a potential way in the estimation of CO_2 flux.

doi:10.6342/NTU201703598

4.3. The comparison of MAPE in different site

According to **Table 7** and **Table 10**, there were a little bit different in the variation trend. **Table 10** was taken as an example. The results of AL group showed that the simulation performance of GDP-T1 was better than US-Esm in terms of the *R* values. However, the values of MAPE and RMSE in the US-Esm simulation were smaller than those of GDP-T1. This conditions could be explained based on the definitions of MAPE and RMSE.

First, the MAPE value represented the error as a percentage of the observed data. The data from these sites was totally different. The characteristics of meteorological conditions and CO_2 dynamics in these sites were not the same. Nonetheless, the total number of the GDP-T1 site and the US-Esm site were different. That is to say, the two MAPE values should not be compared to each other.

Second, the RMSE value has the same unit with the original data, so it will be influenced by the scalar of the original data. The original Fc of the GDP-T1 was distributed from the -40 to 25 $\mu mol \cdot m^{-2} \cdot s^{-1}$, but that of the US-Esm site was between the -14 and 13 $\mu mol \cdot m^{-2} \cdot s^{-1}$. The range of the data of US-Esm was smaller, and it might affect the RMSE value.

4.4. The results of data division

According to the current studies, the discussion of the effects of data division is limited, and only Moffat et al. (2007) divided the data into daytime and nighttime data. In this study, the difference between daytime and nighttime data was taken into consideration, and the results of the divided group were compared to those of the undivided group. Based on the comparison of the two kinds of data, the importance of

doi:10.6342/NTU201703598

data division was more obvious. The results of MAPE could not be compared to the different data group, but the values of RMSE were higher, and the values of *R* were lower in the AL group. That is to say, the model performances of the undivided group were worse than the divided group.

The improvements in the model performances of the cross-site simulations were clearer than those of the simulation at each site. However, the values of *R* in the cross-site simulation were in the range between the modestly correlated and the moderately correlated except the results of DT model at the GDP-T1 site. There might be some characteristics of the data distribution should be analyzed further. Even though the capability of the cross-site simulation in the NT model was not accurate, the capability of the DT model could provide the understanding of C uptake of the low-latitude wetland ecosystem.

doi:10.6342/NTU201703598

5. Conclusions and Future Works

In this study, the meteorological variables were used as the input data to construct the ANN models in order to estimate the variation of CO_2 flux, and the results proved that it is workable. All the results of MAPE were in the range of reasonable simulation, so the ANN models can be used to quantify the CO_2 flux in the wetland ecosystems. Among all the results, the best results of the GDP-T1 and GDP-T2 site occurred in the simulation of the DT model, and their values of *R* were 0.89 and 0.87, respectively. The best results of the US-Esm site occurred in the simulation of the NT models, and the values of *R* was 0.62. Even though there were not micrometeorological data, the predicted data still fitted the variation of observed data well. The division groups (DT and NT group) provided the knowledge of different characteristics in specific temporal condition. For the most part, the models trained by the DT data is helpful to the understanding of the C sequestration.

The second part of this study is to verify the cross-site simulation capability of the ANN models. The results of GDP-T1 were better than those of the US-Esm. The best results occurred in the simulation of the GDP-T1 in the DT model, and the values of *R* was up to 0.73. Although the results of the US-Esm were not superior to those of the GDP-T1, the results of *R* were close to the moderately correlated. Therefore, it showed that the cross-site simulation is feasible, but these models should be applied to where the climatic conditions are quite similar.

In this study, the ANN models performed well in simulating the Fc data in each site which showed the potential for predicting the Fc with the meteorological variables and being applied more extensively although there is still much to be desired in the accuracy of cross-site simulation. To improve the model performances, the influence of

doi:10.6342/NTU201703598

the seasonality might be taken into consideration. Because the EC towers which observing the wetland ecosystem are sparse, these results might provide more information of the exchange of the C, and can help to estimate the variation of CO_2 flux more accurately. Thereby, the estimation results can be used for analyzing the variation and trend of the CO_2 flux, and help to explain the global missing C sink.

doi:10.6342/NTU201703598

Reference

Amundson, R., Stern, L., Baisden, T., and Wang, Y. (1998). The isotopic composition of soil and soil-respired CO 2. Geoderma, 82(1), 83-114.

Armentano, T. V., and Menges, E. S. (1986). Patterns of change in the C balance of organic soil-wetlands of the temperate zone. The Journal of Ecology, 755-774.

Baldocchi, D. (2014). Measuring fluxes of trace gases and energy between ecosystems and the atmosphere–the state and future of the eddy covariance method. Global change biology, 20(12): 3600-3609.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., Paw U, K. T., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., Wofsy, S. (2001). Fluxnet: A new tool to study the temporal and spatial variability of ecosystem-scale C dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society, 82(11): 2415-2434.

Bernal, B., and Mitsch, W. J. (2008). A comparison of soil carbon pools and profiles in wetlands in Costa Rica and Ohio. Ecological Engineering, 34(4), 311-323.

Bernal, B., and Mitsch, W. J. (2012). Comparing C sequestration in temperate freshwater wetland communities. Global Change Biology, 18(5), 1636-1647.

Bonneville, M. C., Strachan, I. B., Humphreys, E. R., and Roulet, N. T. (2008). Net ecosystem CO 2 exchange in a temperate cattail marsh in relation to biophysical properties. Agricultural and Forest Meteorology, 148(1), 69-81.

Bousquet, P., Peylin, P., Ciais, P., Le Quéré, C., Friedlingstein, P., and Tans, P. P. (2000). Regional changes in C dioxide fluxes of land and oceans since 1980. Science, 290(5495), 1342-1346.

Brix, H., Sorrell, B. K., and Lorenzen, B. (2001). Are Phragmites-dominated wetlands a net source or net sink of greenhouse gases? Aquatic Botany, 69(2): 313-324.

Bubier, J. L., Crill, P. M., Moore, T. R., Savage, K., and Varner, R. K. (1998). Seasonal patterns and controls on net ecosystem CO2 exchange in a boreal peatland complex. Global Biogeochemical Cycles, 12(4), 703-714.

Carrara, A., Kowalski, A. S., Neirynck, J., Janssens, I. A., Yuste, J. C., & Ceulemans, R. (2003). Net ecosystem CO 2 exchange of mixed forest in Belgium over 5 years. Agricultural and Forest Meteorology, 119(3), 209-227.

Chang, F. J., and Chang, L. C. (2015). Artificial neural network: Theory and application (2nd edition). Tsang Hai Publishing: Taipei.

Chmura, G. L., Anisfeld, S. C., Cahoon, D. R., Lynch, J. C. (2003). Global C sequestration in tidal, saline wetland soils. Global biogeochemical cycles, 17(4).

doi:10.6342/NTU201703598

doi:10.1029/2002GB001917.

Clevering, O. A., and Lissner, J. (1999). Taxonomy, chromosome numbers, clonal diversity and population dynamics of Phragmites australis. Aquatic Botany, 64(3), 185-208.

Dragomir, C. M., Klaassen, W., Voiculescu, M., Georgescu, L. P., van der Laan, S. (2012). Estimating annual CO2 flux for Lutjewad station using three different gap-filling techniques. The Scientific World Journal, 2012. doi:10.1100/2012/842893.

Duever, M. J., Meeder, J. F., Meeder, L. C., and McCollom, J. M. (1994). The climate of south Florida and its role in shaping the Everglades ecosystem. In: Everglades: The Ecosystem and its Restoration, ed. S. M. Davis, and J. C. Ogden, 225-248. Delray Beach: St. Lucie Press.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. L., Meyers, T., Moncrieff, J., Moors, E., Munger, J. W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S. (2001). Gap filling strategies for defensible annual sums of net ecosystem exchange. Agricultural and forest meteorology, 107(1), 43-69.

Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N.-O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T., Law, B. E., Meyers, T., Moncrieff, J. Moors, E., Munger, J. W., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., Wofsy, S. (2001). Gap filling strategies for defensible annual sums of net ecosystem exchange. Agricultural and forest meteorology, 107(1), 43-69.

Fennessy, M. S. (2014). Wetland ecosystems and global change, in Bill Freedman (ed.), Global Environmental Change, Netherlands: Springer, 255-261.

FLUXNET (2015) Historical Site Status. http://fluxnet.ornl.gov/site_status (last accessed 30 December 2015).

Gorham, E. (1995). The biogeochemistry of northern peatlands and its possible response to global warming. Biotic Feedback in the Global Climatic System, GM Woodwell and FT Mackenzie (eds.), Oxford University Press, New York, 169-187.

Gower, S. T., Kucharik, C. J., and Norman, J. M. (1999). Direct and indirect estimation of leaf area index, f APAR, and net primary production of terrestrial ecosystems. Remote sensing of environment, 70(1), 29-51.

Han, G., Chu, X., Xing, Q., Li, D., Yu, J., Luo, Y., Wang, G., Mao, P., and Rafique, R. (2015). Effects of episodic flooding on the net ecosystem CO2 exchange of a supratidal wetland in the Yellow River Delta. Journal of Geophysical Research:

doi:10.6342/NTU201703598

Biogeosciences, 120(8), 1506-1520.

Hollinger, D. Y., Aber, J., Dail, B., Davidson, E. A., Goltz, S. M., Hughes, H., Leclerc, M. Y., Lee, J. T., Richardson, A. D., Rodrigues, C., Scott, N. A., Achuatavarier, D., Walsh, J. (2004). Spatial and temporal variability in forest–atmosphere CO2 exchange. Global Change Biology, 10(10), 1689-1706.

Inglett, K. S., Inglett, P. W., Reddy, K. R., Osborne, T. Z. (2012). Temperature sensitivity of greenhouse gas production in wetland soils of different vegetation. Biogeochemistry, 108(1-3): 77-90.

Jimenez, K. L., Starr, G., Staudhammer, C. L., Schedlbauer, J. L., Loescher, H. W., Malone, S. L., and Oberbauer, S. F. (2012). Carbon dioxide exchange rates from short-and long-hydroperiod Everglades freshwater marsh. Journal of Geophysical Research: Biogeosciences, 117(G4).

Jung, M., M. Reichstein, Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., Williams, C. (2011). Global patterns of land-atmosphere fluxes of C dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations. Journal of Geophysical Research: Biogeosciences (2005–2012), 116(G3).

Kaimal, J. C. and J. J. Finnigan. 1994. Atmospheric boundary layer flows: Their structure and measurement. Oxford university press, New York.

Knorr, W., and Kattge, J. (2005). Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling. Global Change Biology, 11(8), 1333-1351.

Kwok, T. Y., and Yeung, D. Y. (1997). Constructive algorithms for structure learning in feedforward neural network for regression problems. IEEE Transactions on Neural Network, 8(3), 630-645.

Lal, R. (2008). C sequestration. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1492), 815-830.

Lee, S. C., Fan, C. J., Wu, Z. Y., and Juang, J. Y. (2015). Investigating effect of environmental controls on dynamics of CO2 budget in a subtropical estuarial marsh wetland ecosystem. Environmental Research Letters, 10(2), 025005.

Leuning, R., Cleugh, H. A., Zegelin, S. J., and Hughes, D. (2005). C and water fluxes over a temperate Eucalyptus forest and a tropical wet/dry savanna in Australia: measurements and comparison with MODIS remote sensing estimates. Agricultural and Forest Meteorology, 129(3), 151-173.

Lewis, C. D. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. London: Butterworth Scientific.

doi:10.6342/NTU201703598

Melesse, A. M. and Hanley, R. S. (2005). Artificial neural network application for multi-ecosystem C flux simulation. Ecological Modelling, 189(3): 305-314.

Mitsch, W. J., and Gosselink, J. G. 2007. Wetlands, 4th ed. John Wiley & Sons, New York

Mitsch, W. j., Bernal, B., Nahlik, A. M., Mander, Ü., Zhang, L., Anderson, C. J., Jørgensen, S. E., Brix, H. (2013). Wetlands, C, and climate change. Landscape Ecology, 28(4): 583-597.

Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G.,Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J. H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., Stauch, V. J. (2007) Comprehensive comparison of gap-filling techniques for eddy covariance net C fluxes. Agricultural and Forest Meteorology, 147(3): 209-232.

Mogami, J., Hirata, R., Hirano, T., Miura, M., Fujinuma,Y., Inukai, K. (2003). Comparison of gap-filling methods for the data sets of net ecosystem exchange and energy fluxes in a larch forest. J. Agric. Meteorol. (Hokkaido), 55: 9-16.

Morin, T. H., Bohrer, G., Frasson, R. D. M., Naor-Azreli, L., Mesi, S., Stefanik, K. C., and Schäfer, K. V. R. (2014). Environmental drivers of methane fluxes from an urban temperate wetland park. Journal of Geophysical Research: Biogeosciences, 119(11), 2188-2208.

National Climatic Data Center (NCDC). 2015. Climate at a Glance: U.S. Time Series. http://www.ncdc.noaa.gov/cag/ (last accessed 2017.02.24).

Neue, H. U., Gaunt, J. L., Wang, Z. P., Becker-Heidmann, P., and Quijano, C. (1997). C in tropical wetlands. Geoderma, 79(1), 163-185.

O'Connell, M. J. (2003). Detecting, measuring and reversing changes to wetlands. Wetland Ecology and Management, 11: 397-401.

Ooba, M., Hirano, T., Mogami, J. I., Hirata, R., Fujinuma, Y. (2006). Comparisons of gap-filling methods for C flux dataset: a combination of a genetic algorithm and an artificial neural network. Ecological modelling, 198(3): 473-486.

Papale, D., and Valentini, R. (2003). A new assessment of European forests C exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biology, 9(4), 525-535.

Papale, D., and Valentini, R. (2003). A new assessment of European forests C exchanges by eddy fluxes and artificial neural network spatialization. Global Change Biology, 9(4), 525-535.

Park, J., Hong, M. G., & Kim, J. G. (2013). Relationship between early development of plant community and environmental condition in abandoned paddy terraces at mountainous valleys in Korea. Journal of Ecology and Environment, 36(2), 131-140.

doi:10.6342/NTU201703598

Ramsar. (2015). Wetlands of International Importance (Ramsar Sites). http://www.ramsar.org/about/wetlands-of-international-importance-ramsar-sites (last accessed 30 December 2015).

Rayner, P. J., Law, R. M., Allison, C. E., Francey, R. J., Trudinger, C. M., and Pickett-Heaps, C. (2008). Interannual variability of the global C cycle (1992–2005) inferred by inversion of atmospheric CO2 and δ13CO2 measurements. Global Biogeochemical Cycles, 22(3).

Rowbottom, C. G., Webb, S., and Oldham, M. (1999). Beam-orientation customization using an artificial neural network. Physics in medicine and biology, 44(9), 2251.

Running, S. W., Baldocchi, D. D., Turner, D. P., Gower, S. T., Bakwin, P. S., and Hibbard, K. A. (1999). A global terrestrial monitoring network integrating tower fluxes, flask sampling, ecosystem modeling and EOS satellite data. Remote Sensing of Environment, 70(1), 108-127.

Schedlbauer, J. L., Oberbauer, S. F., Starr, G., and Jimenez, K. L. (2010). Seasonal differences in the CO 2 exchange of a short-hydroperiod Florida Everglades marsh. Agricultural and Forest Meteorology, 150(7), 994-1006.

Schedlbauer, J. L., Munyon, J. W., Oberbauer, S. F., Gaiser, E. E., and Starr, G. (2012). Controls on ecosystem C dioxide exchange in short-and long-hydroperiod Florida Everglades freshwater marshes. Wetlands, 32(5), 801-812.

Stoy,P. C., Richardson, A. D., Baldocchi, D. D., Katul, G. G., Stanovick, J., Mahecha, M. D., Reichstein, M., Detto,M., Law, B. E., Wohlfahrt, G., Arriga, N., Campos, J., McCaughey, J. H.,Montagnani, L.,Paw U, K. T., Sevanto, S., Williams, M. (2009). Biosphere-atmosphere exchange of CO2 in relation to climate: a cross-biome analysis across multiple time scales. Biogeosciences, 6(10): 2297-2312.

Turetsky, M. R., Wieder, R. K., Vitt, D. H., Evans, R. J., and Scott, K. D. (2007). The disappearance of relict permafrost in boreal north America: Effects on peatland carbon storage and fluxes. Global Change Biology, 13(9), 1922-1934.

Van Wijk, M. T., and Bouten, W. (1999). Water and C fluxes above European coniferous forests modelled with artificial neural network. Ecological Modelling, 120(2), 181-197.

Van Wijk, M. T., Bouten, W., Verstraten, J. M. (2002). Comparison of different modelling strategies for simulating gas exchange of a Douglas-fir forest. Ecological Modelling, 158(1): 63-81.

Vymazal, J. (2013). Emergent plants used in free water surface constructed wetlands: a review. Ecological engineering, 61, 582-592.

Waddington, J. M., and Roulet, N. T. (2000). C balance of a boreal patterned peatland. Global Change Biology, 6(1), 87-97.

Watanabe, T., Yokozawa, M., Emori, S., Takata, K., Sumida, A., and Hara, T. (2004).

doi:10.6342/NTU201703598

Developing a multilayered integrated numerical model of surface physics–growing plants interaction (MINoSGI). Global Change Biology, 10(6), 963-982.

Zhang, Q., Sun, R., Jiang, G., Xu, Z., and Liu, S. (2016). C and energy flux from a Phragmites australis wetland in Zhangye oasis-desert area, China. Agricultural and Forest Meteorology, 230, 45-57.

Zhou, L., Zhou, G., Jia, Q. (2009). Annual cycle of CO2 exchange over a reed(Phragmites australis)wetland in Northeast China. Aquatic Botany, 91(2): 91-98.

doi:10.6342/NTU201703598

Appendix I: The maximum (max) and the minimum (min) of input parameters used in the normalization process at each site.

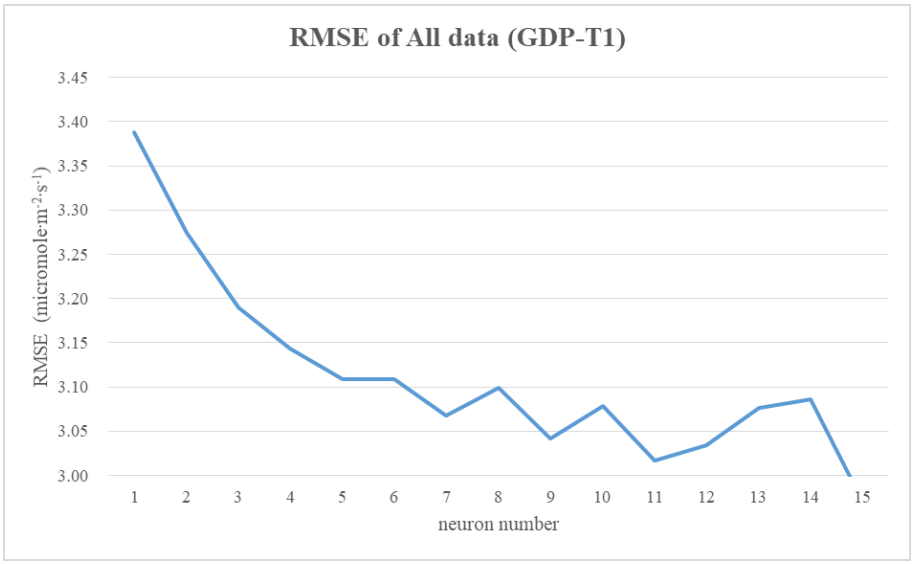
GDP-T1						
group / variable	AL group		DT group		NT group	
	max	min	max	min	max	min
Tair	44.72	2.56	44.29	5.04	41.19	3.19
Rn	1062.00	-298.00	1054.31	-254.42	37.96	-117.29
wind	10.23	-1.77	10.23	-1.74	5.36	-0.91
VPD	5.46	-0.70	5.46	-0.68	3.50	-0.36
Fc	24.57	-39.28	24.08	-39.19	18.19	-3.14
GDP-T2						
group / variable	AL group		DT group		NT group	
	max	min	max	min	max	min
Tair	50.22	4.40	50.14	4.86	43.77	5.54
Rn	2286.22	-521.92	2108.59	-468.69	2286.22	-521.92
wind	7.71	-1.25	7.70	-1.24	7.56	-1.23
VPD	8.59	-1.24	8.57	-1.14	5.57	-0.70
Fc	24.19	-37.67	22.40	-37.35	19.04	-8.49
US-Esm						
group / variable	AL group		DT group		NT group	
	max	min	max	min	max	min
Tair	44.65	-11.16	44.65	-11.16	42.45	-6.94
Rn	1257.34	-344.12	1253.22	-320.80	354.34	-184.77
wind	8.70	-1.17	8.69	-1.13	7.73	-1.00
VPD	3.67	-0.65	3.67	-0.65	3.46	-0.61
Fc	12.62	-13.50	5.64	-12.27	11.12	-4.99

doi:10.6342/NTU201703598

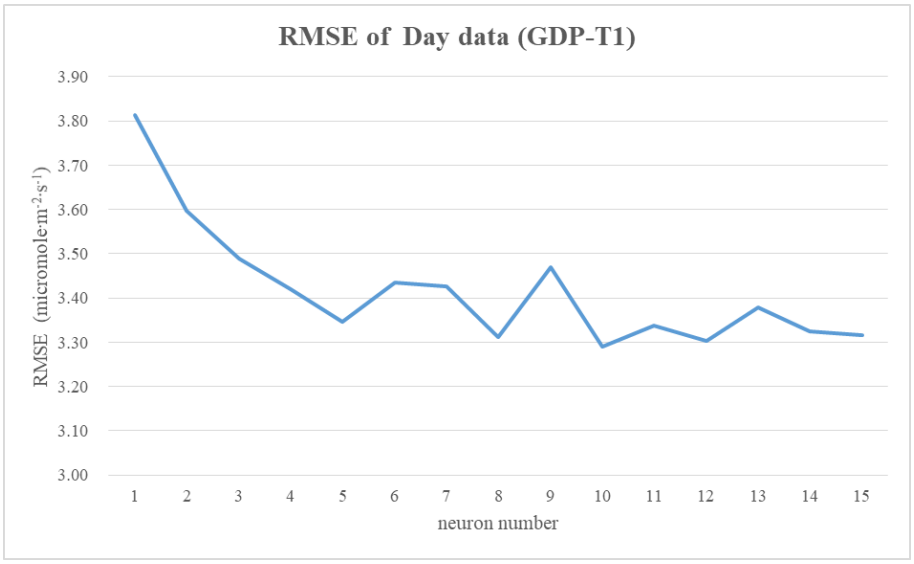
Appendix II: The RMSE curve of three groups in the training phase.

1. GDP-T1

a. All data

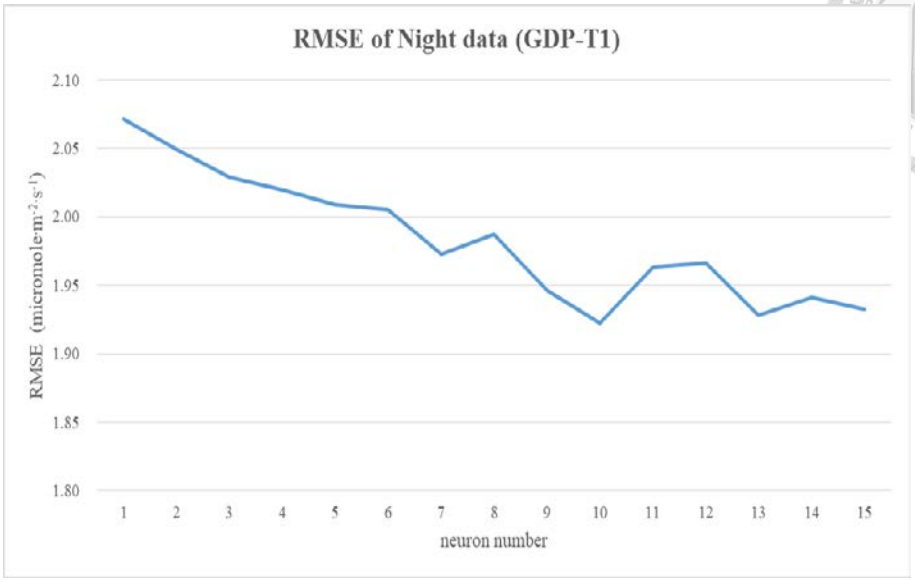


b. Daytime data



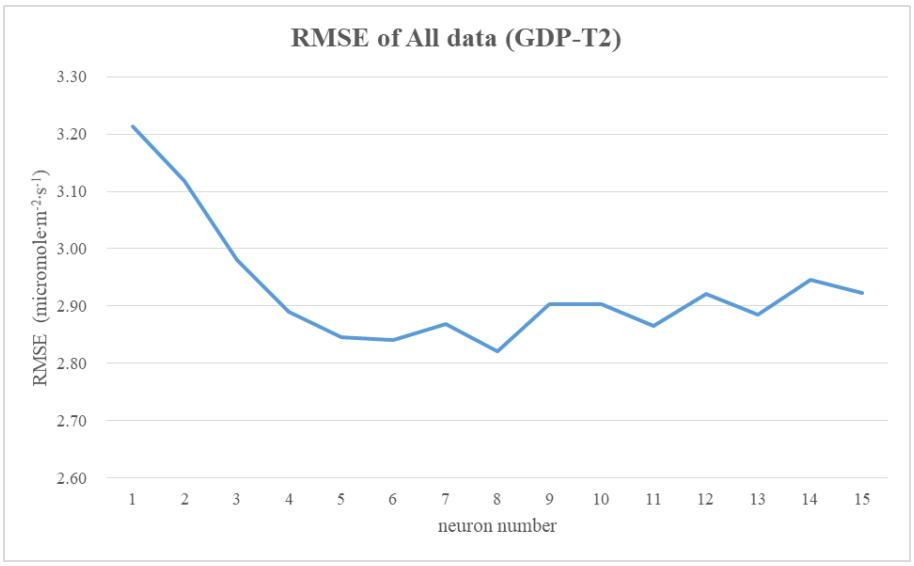
doi:10.6342/NTU201703598

c. Nighttime data



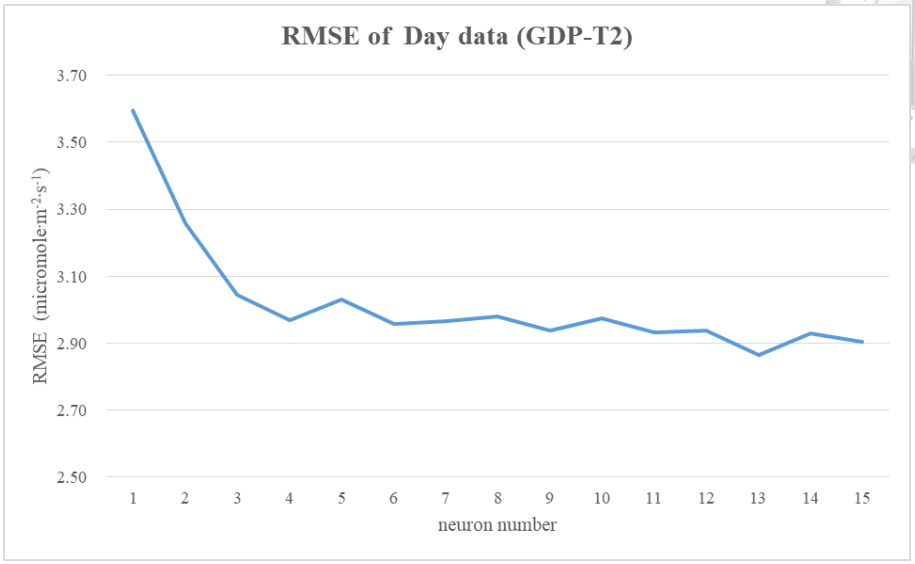
2. GDP-T2

a. All data

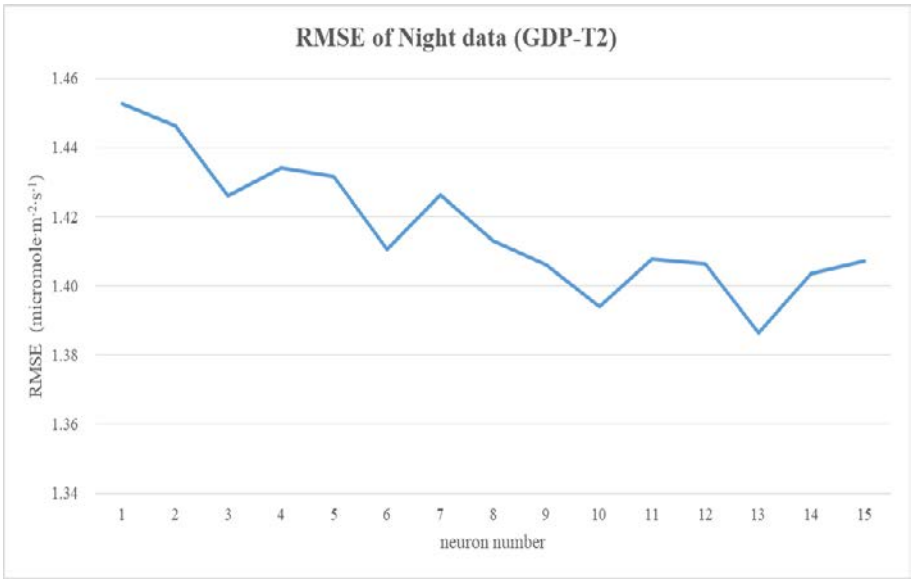


doi:10.6342/NTU201703598

b. Daytime data



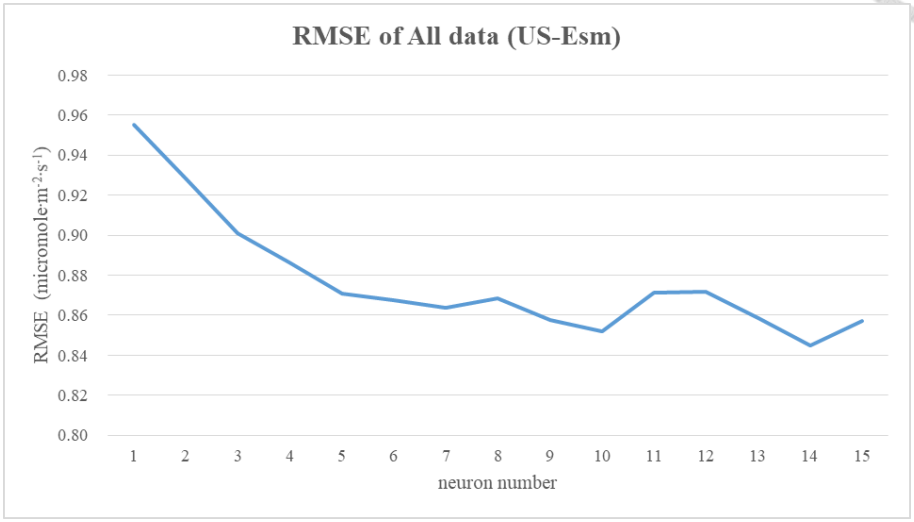
c. Nighttime data



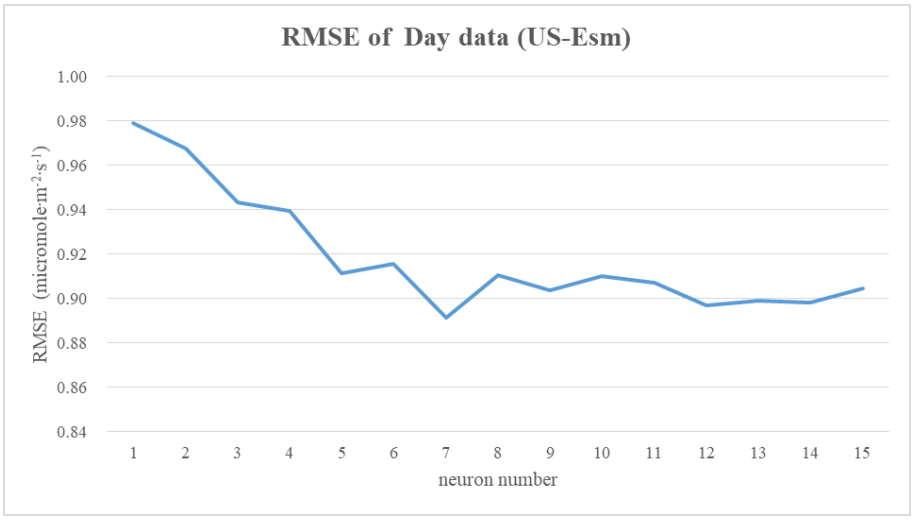
doi:10.6342/NTU201703598

3. US-Esm

a. All data

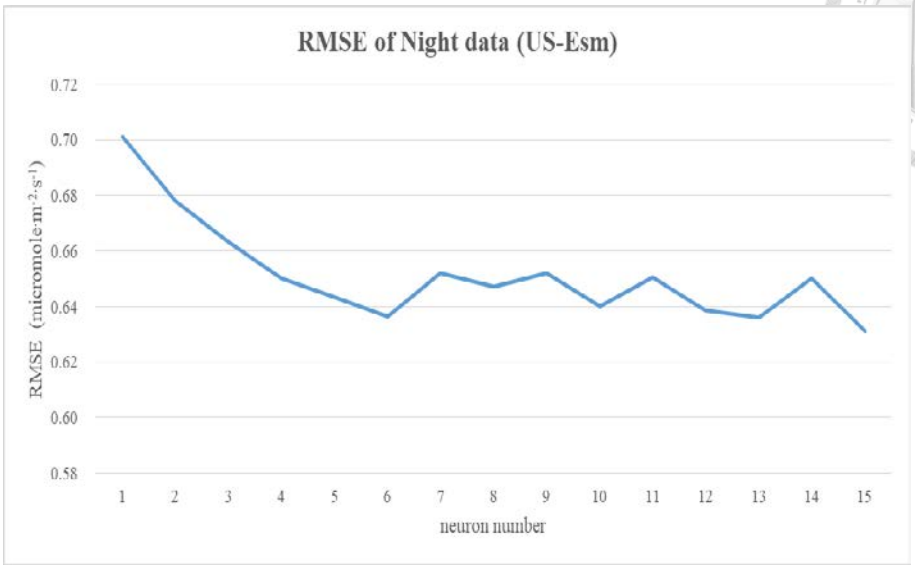


b. Daytime data



doi:10.6342/NTU201703598

c. Nighttime data



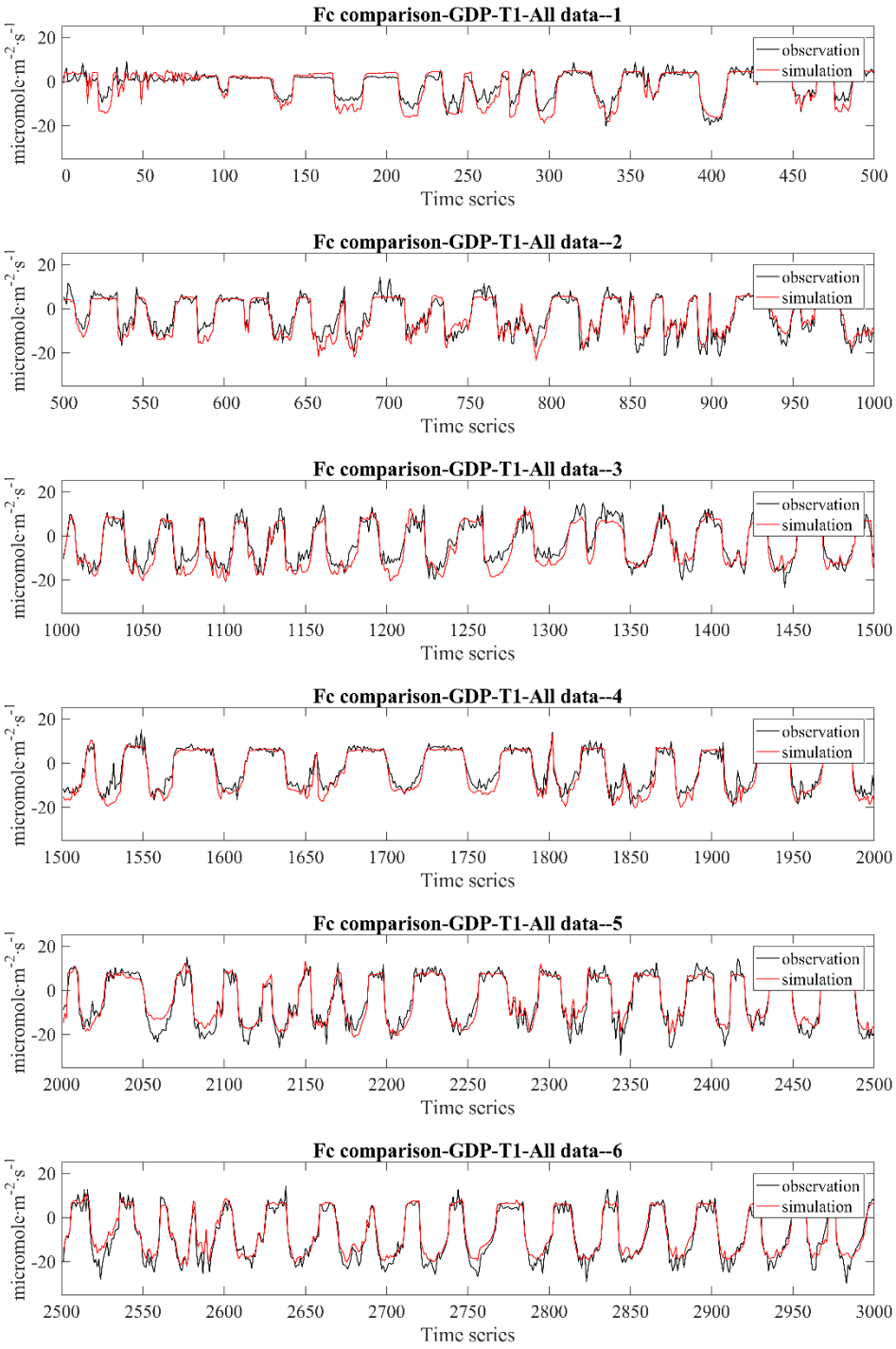
doi:10.6342/NTU201703598

Appendix III: The simulation results in each site.

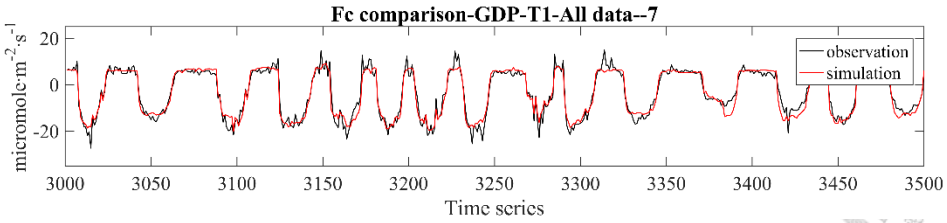
The simulation results in each site.

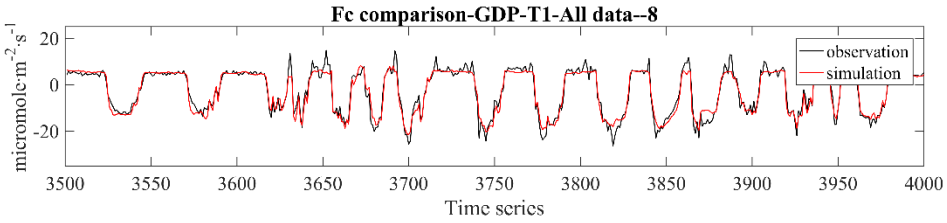
1. GDP-T1

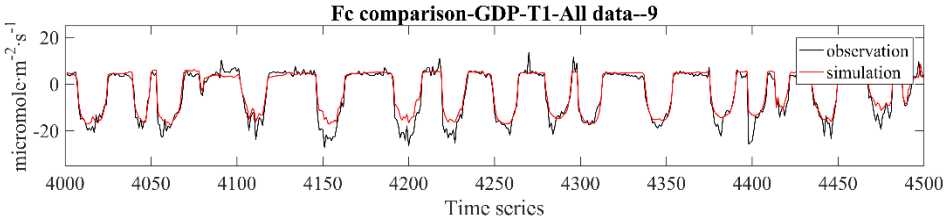
a. All data

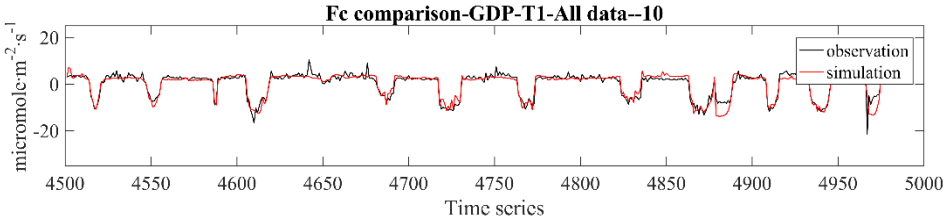


doi:10.6342/NTU201703598

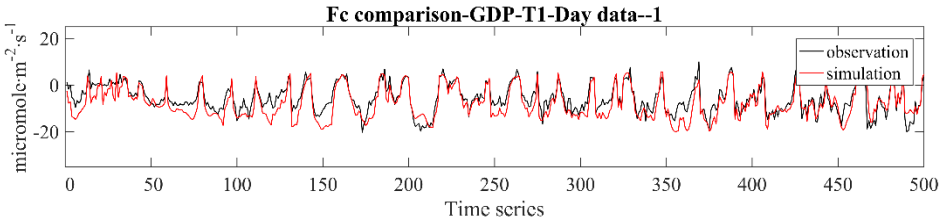


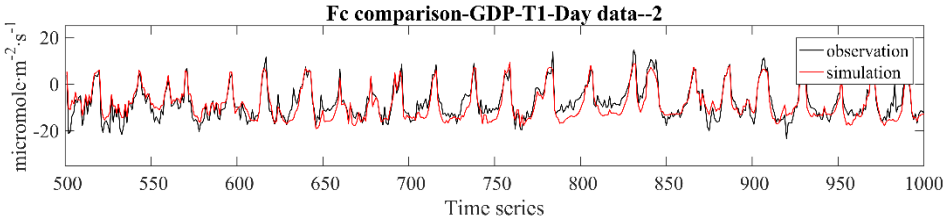


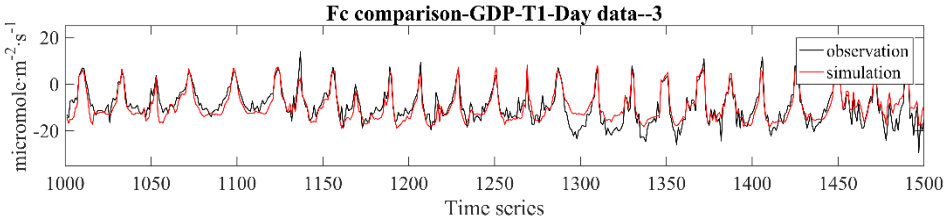




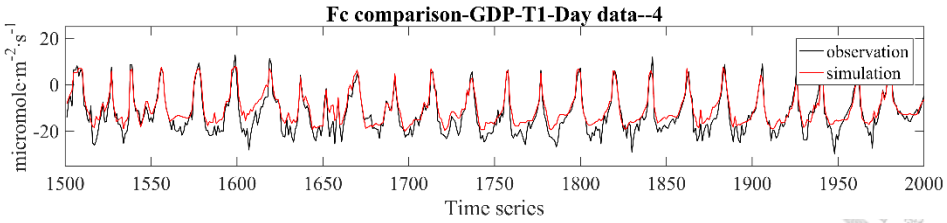
b. Daytime data

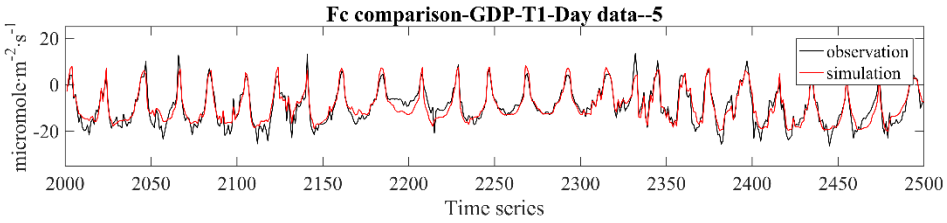


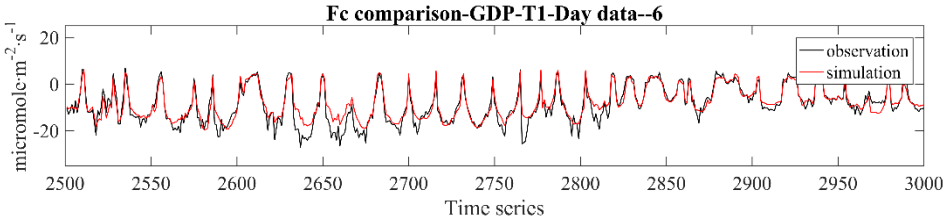


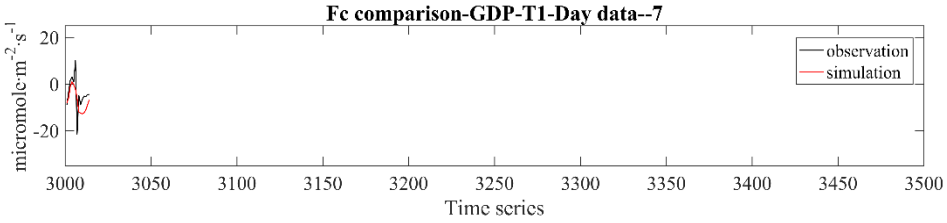


doi:10.6342/NTU201703598

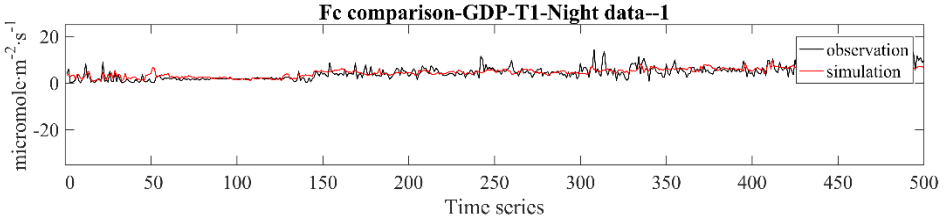


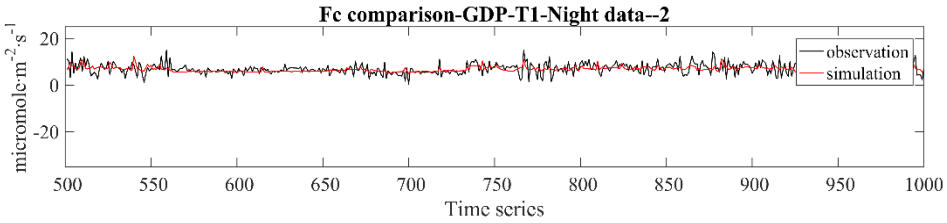


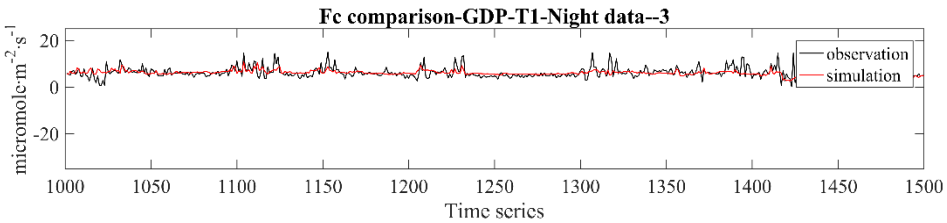




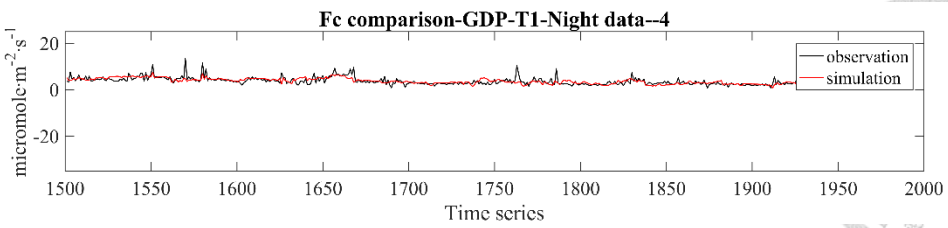
c. Nighttime data





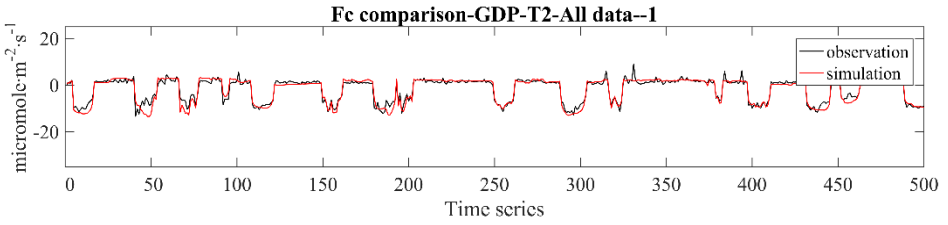


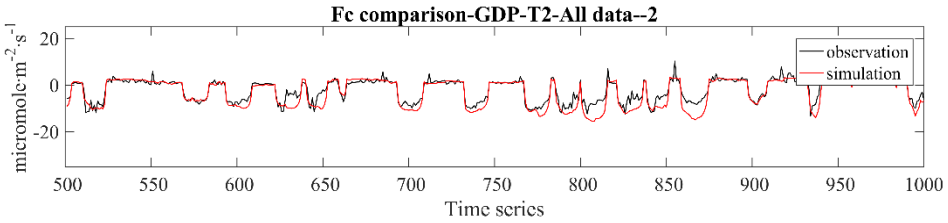
doi:10.6342/NTU201703598

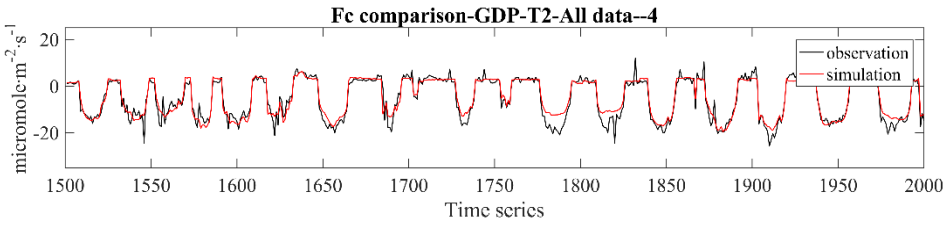


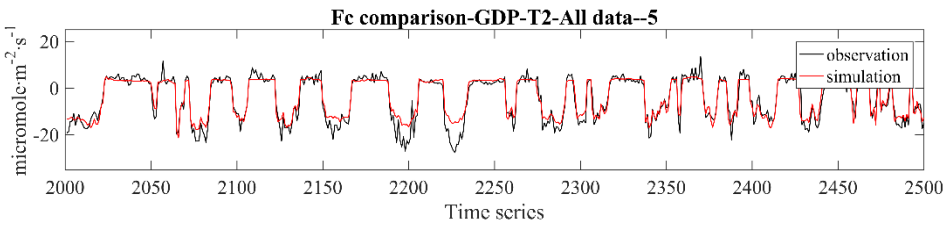
2. GDP-T2

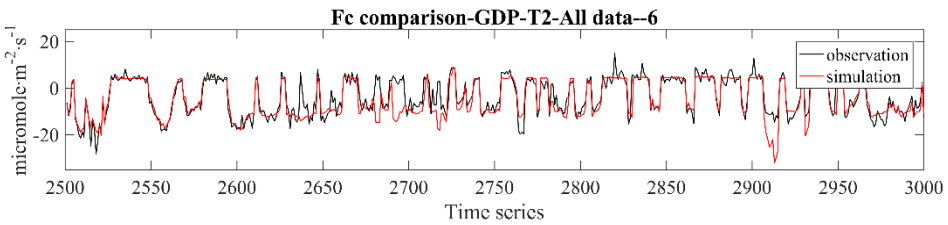
a. All data



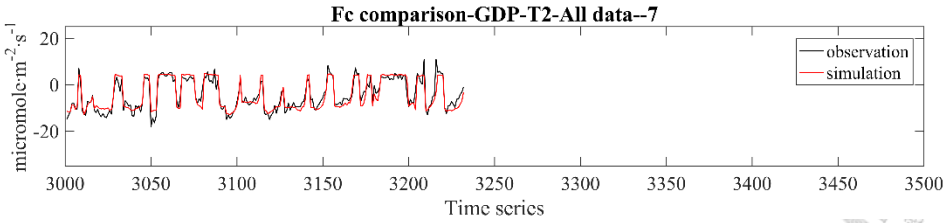




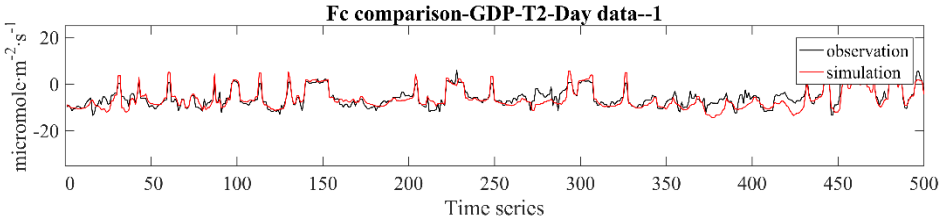


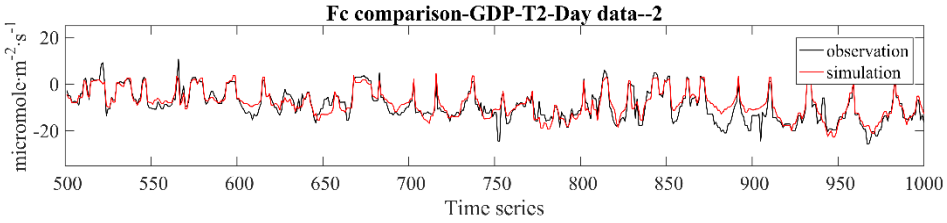


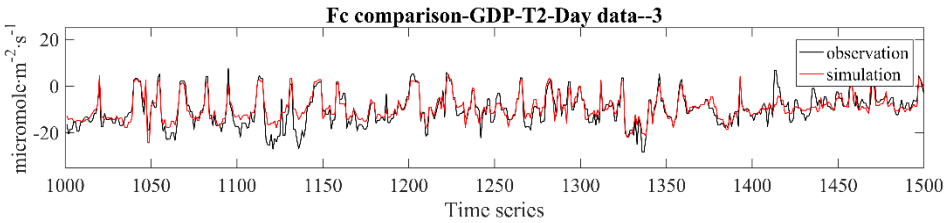
doi:10.6342/NTU201703598

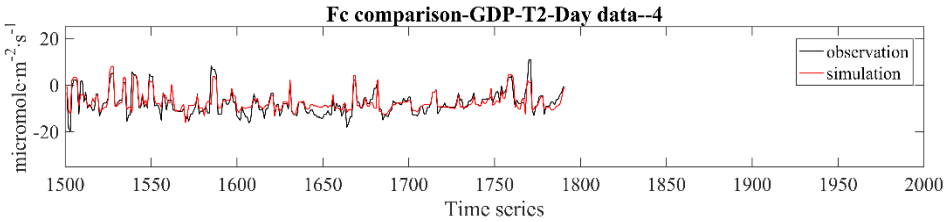


b. Daytime data

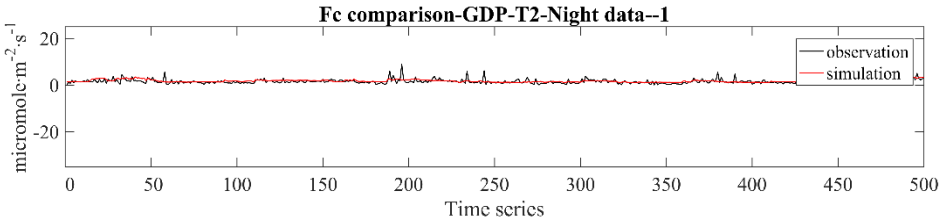




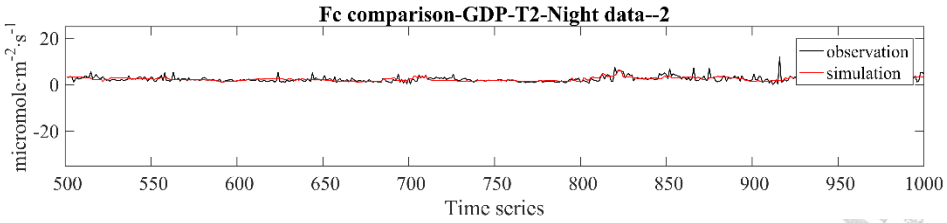


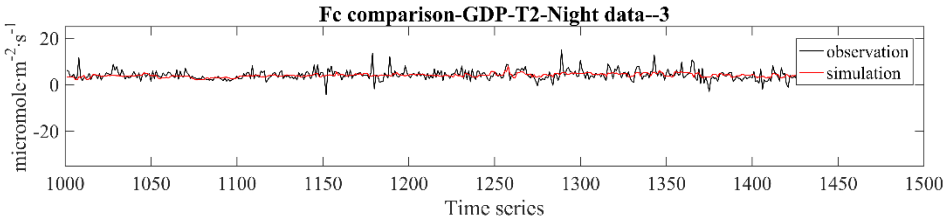


c. Nighttime data



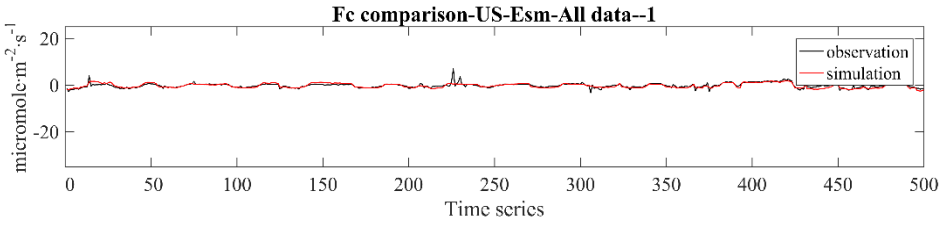
doi:10.6342/NTU201703598

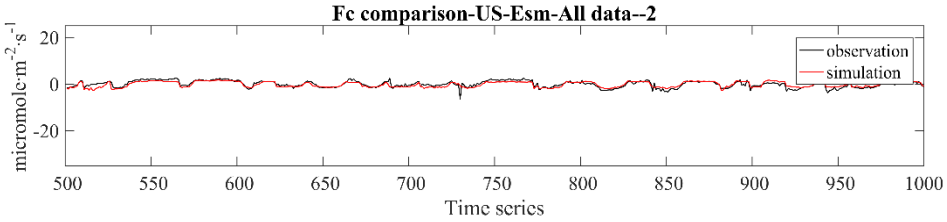


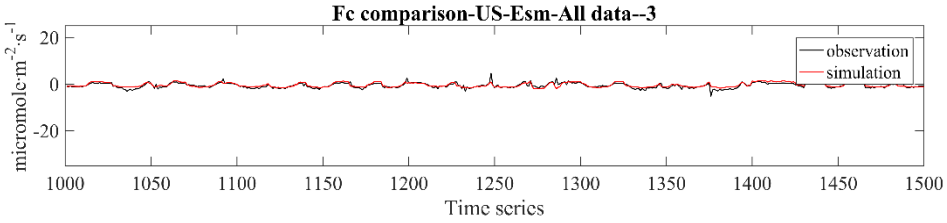


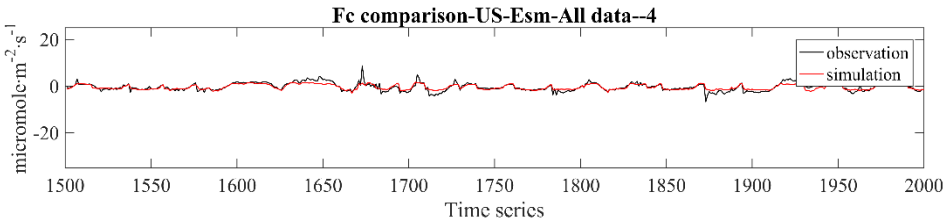
3. US-Esm

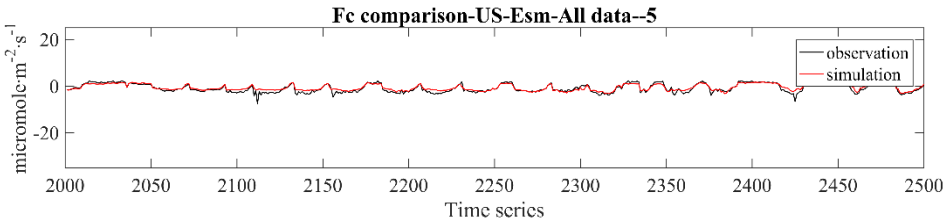
a. All data



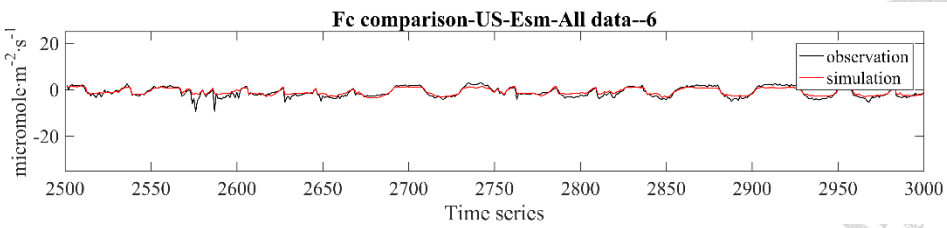


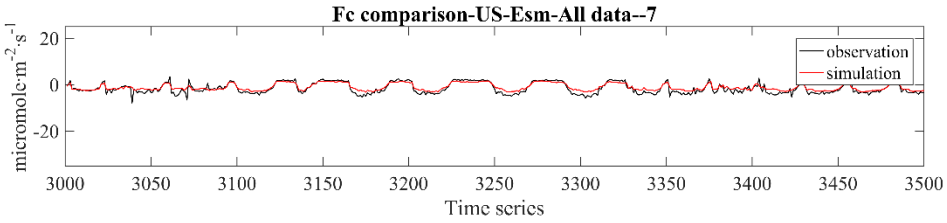


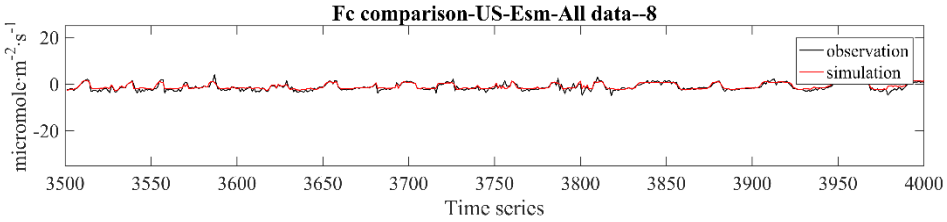


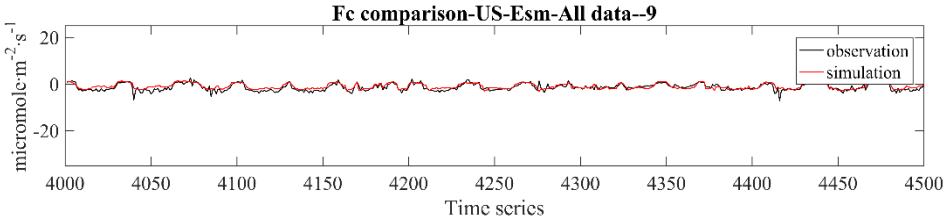


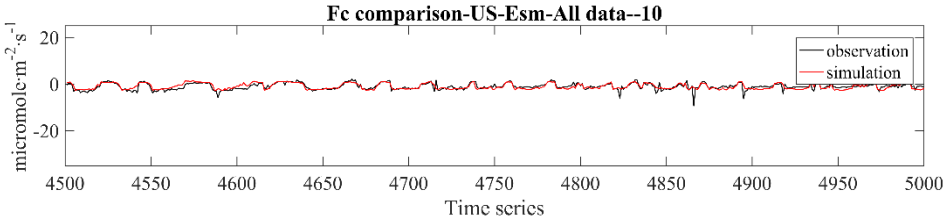
doi:10.6342/NTU201703598

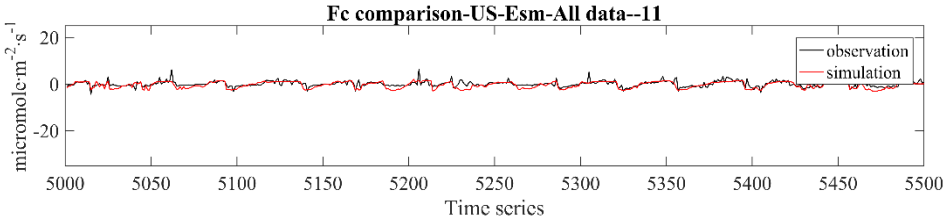


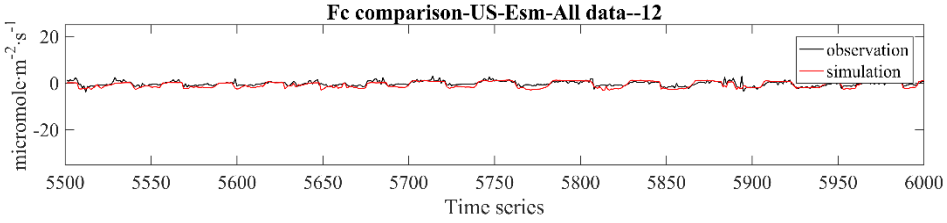




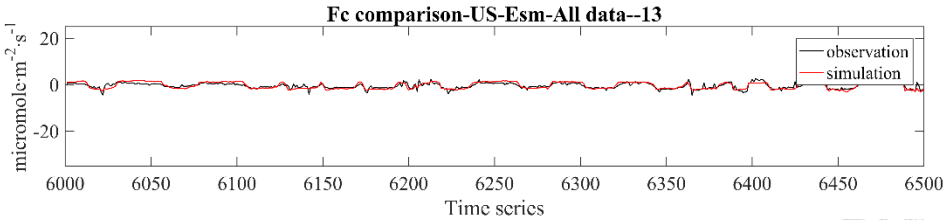


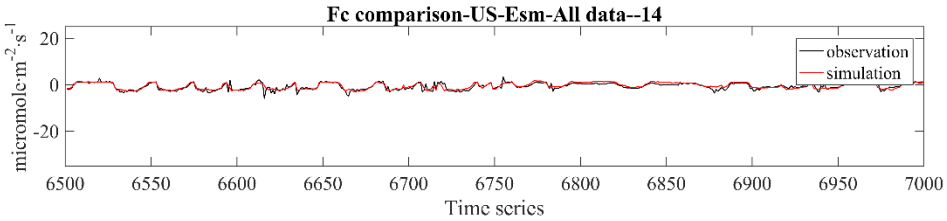


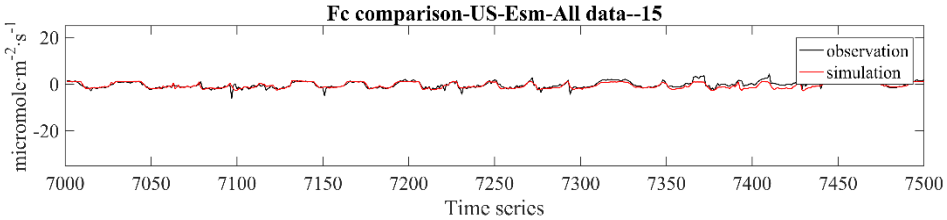


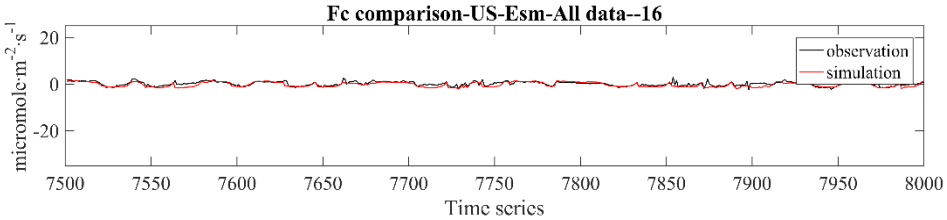


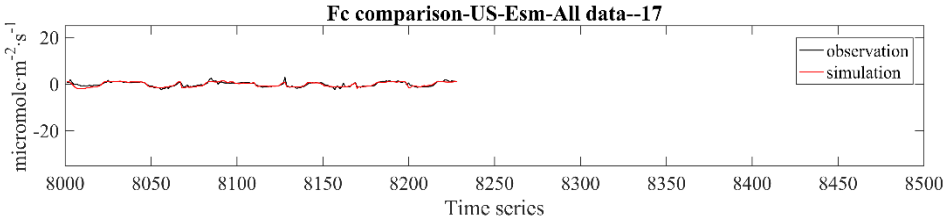
doi:10.6342/NTU201703598



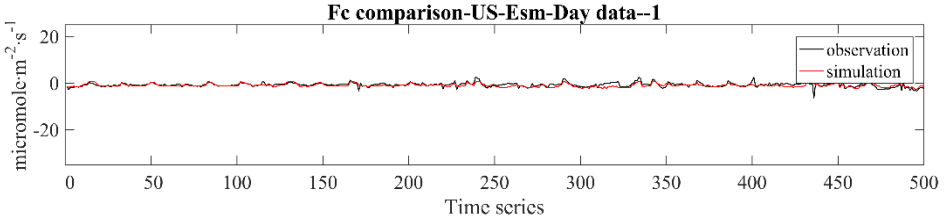


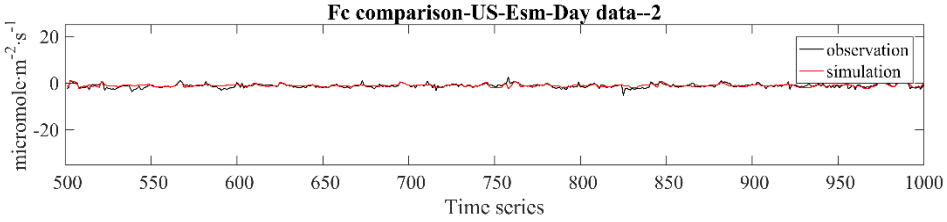




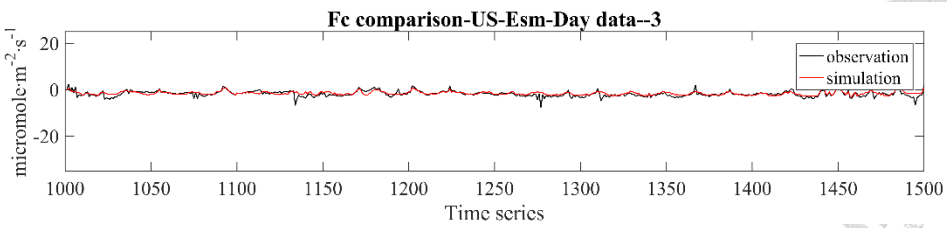


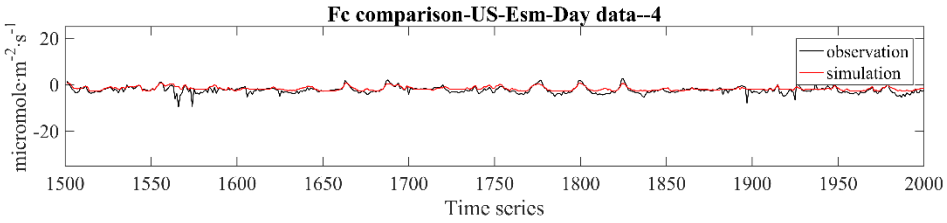
b. Daytime data

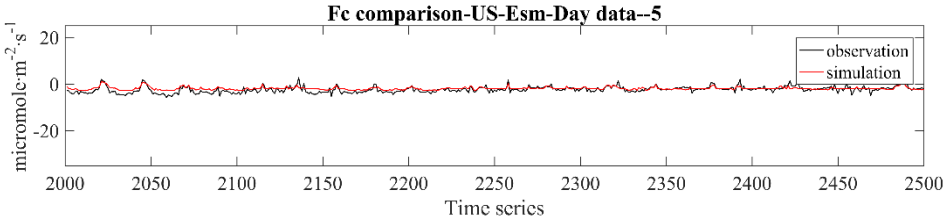


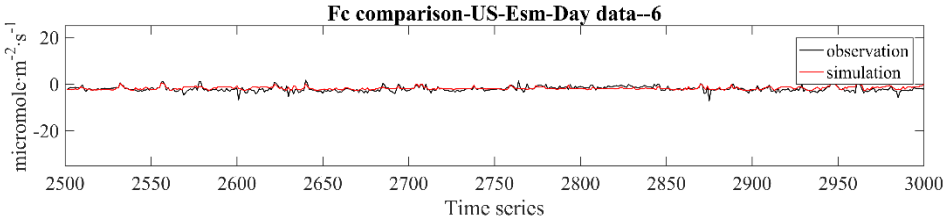


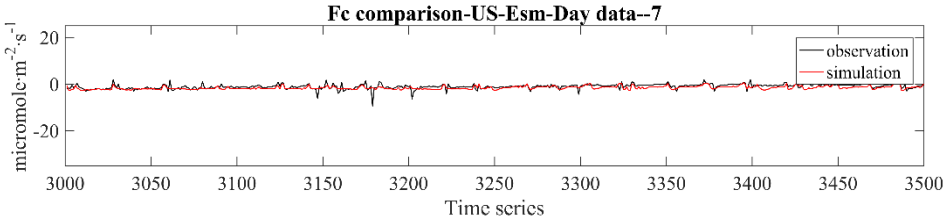
doi:10.6342/NTU201703598

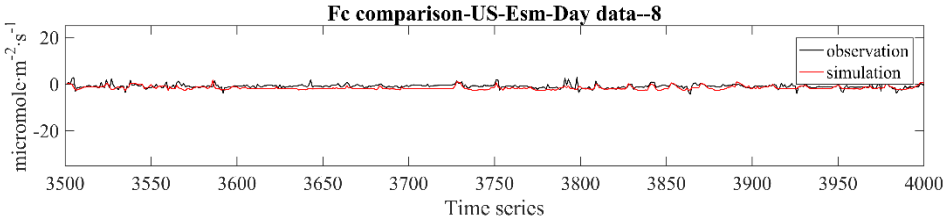


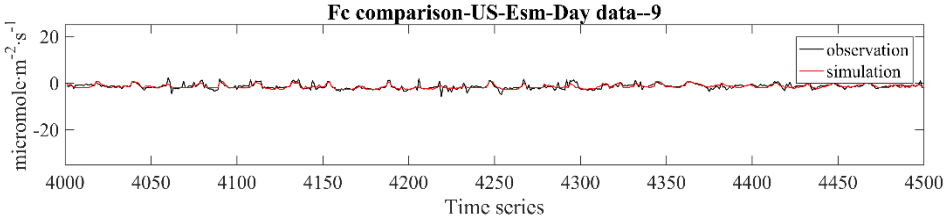




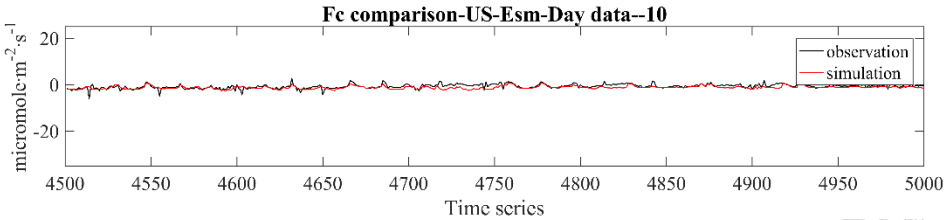


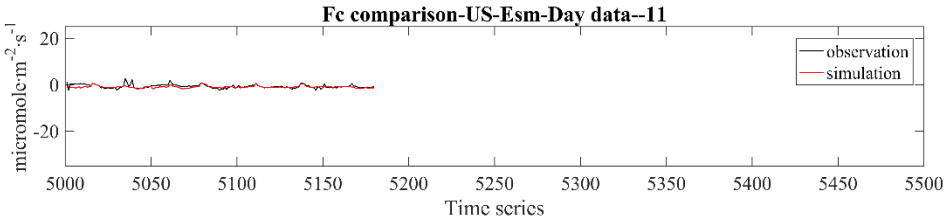




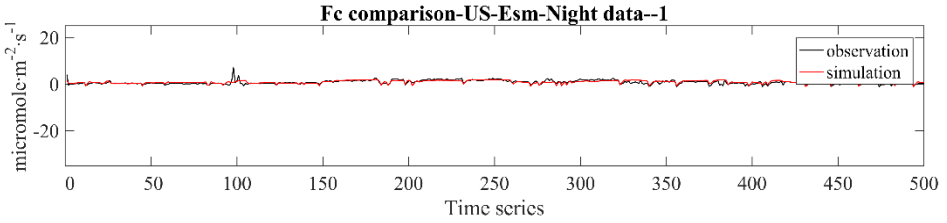


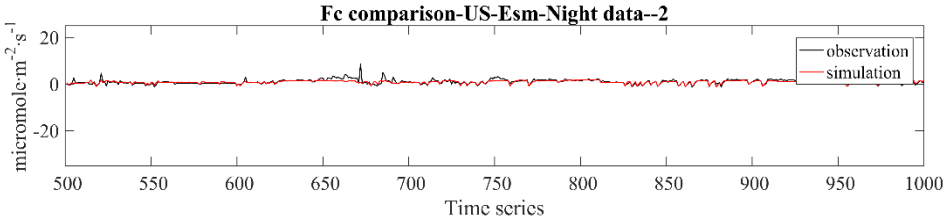
doi:10.6342/NTU201703598

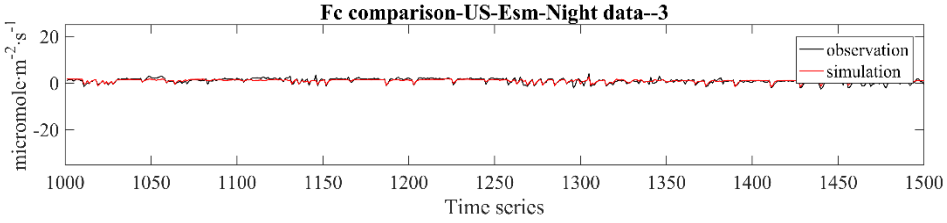


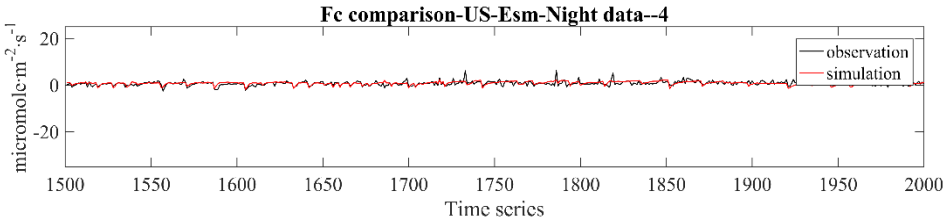


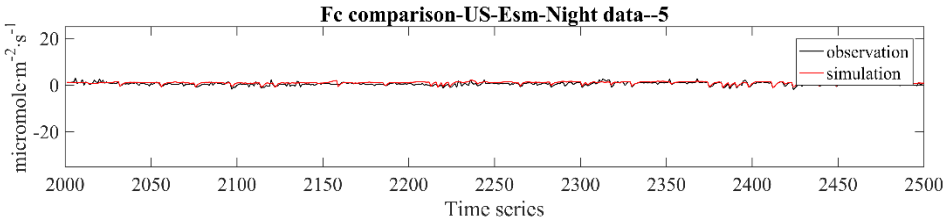
c. Nighttime data



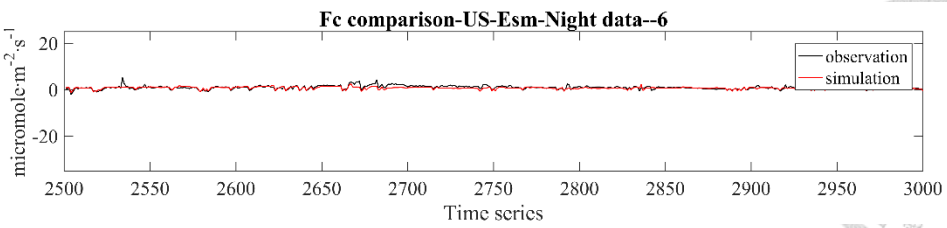


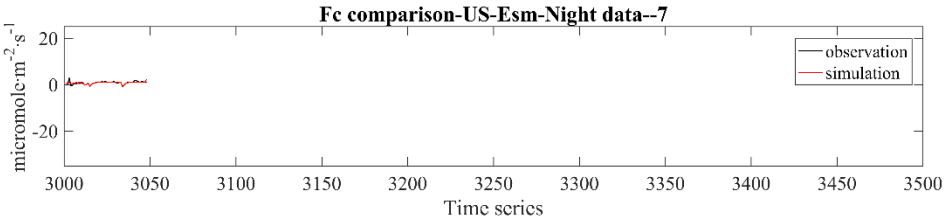






doi:10.6342/NTU201703598

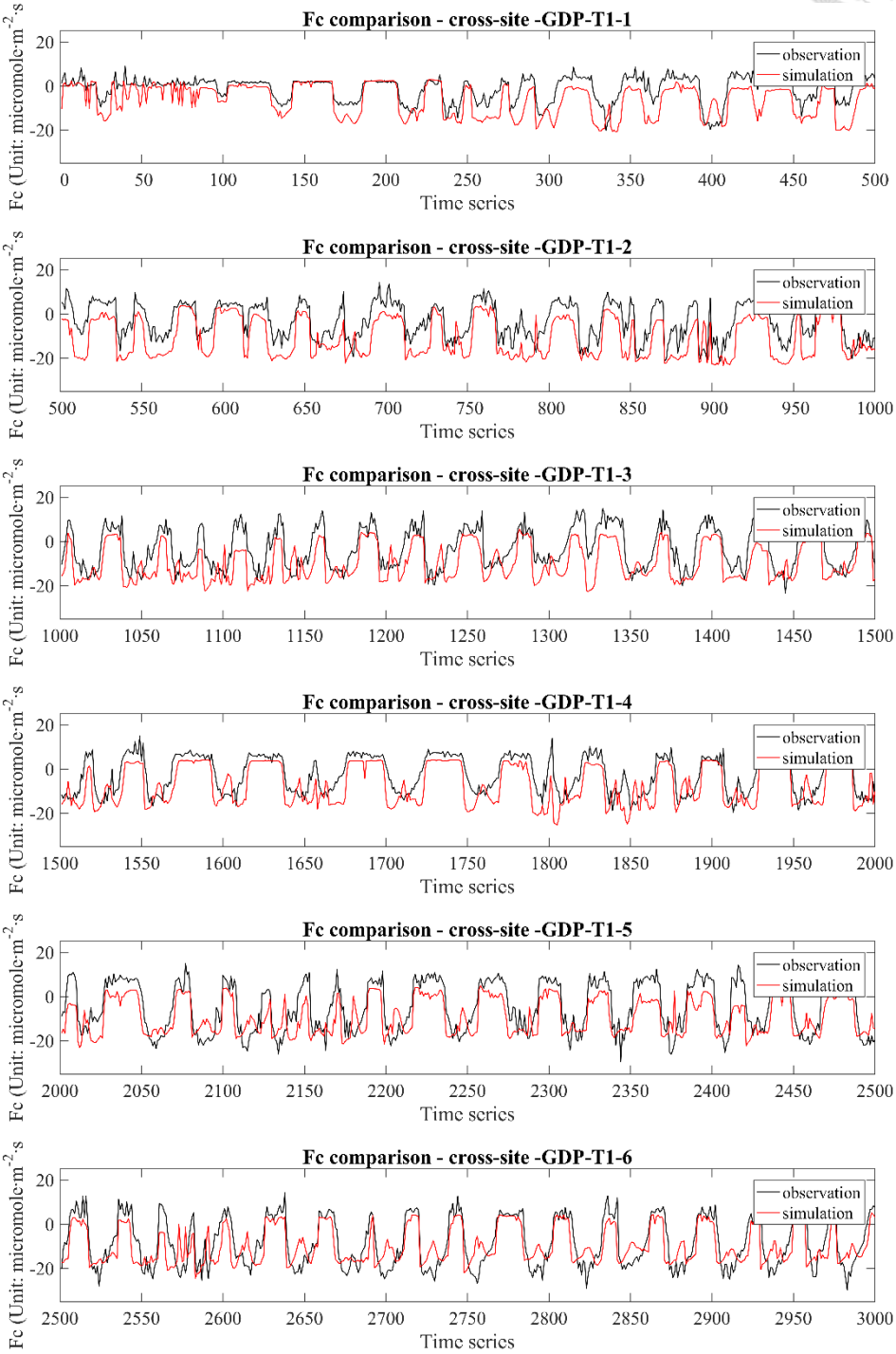




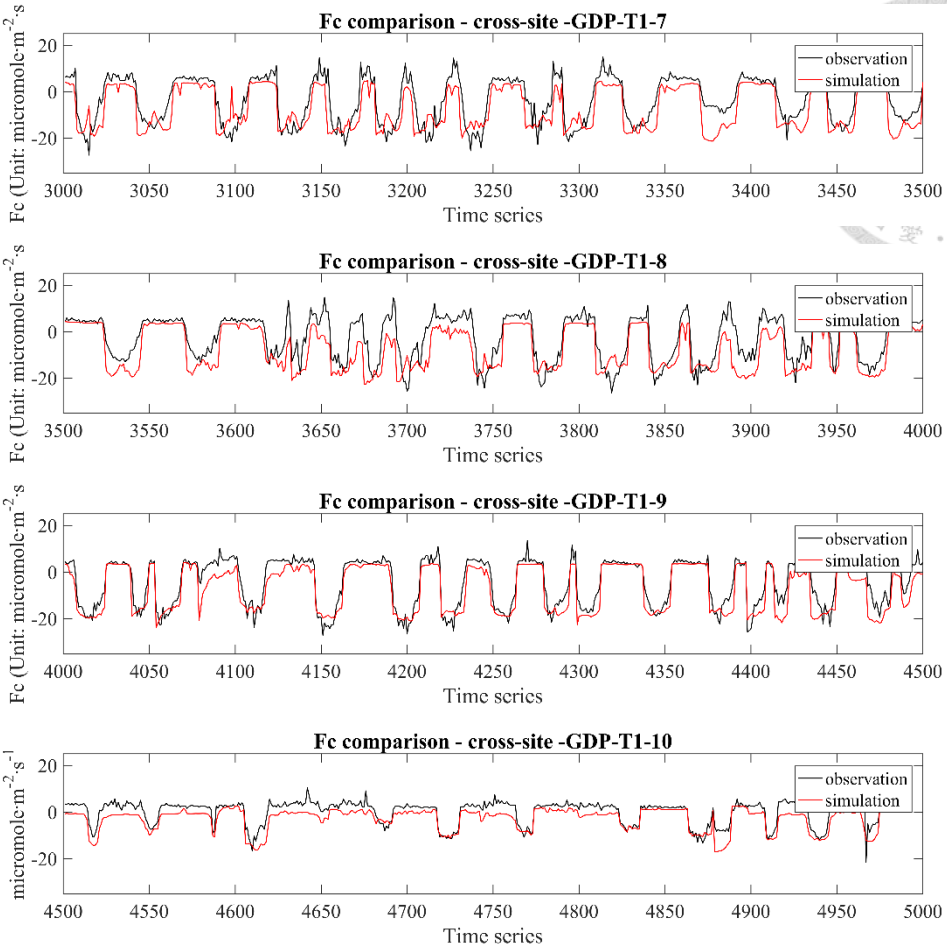
doi:10.6342/NTU201703598

Appendix IV: The cross-site simulation results.

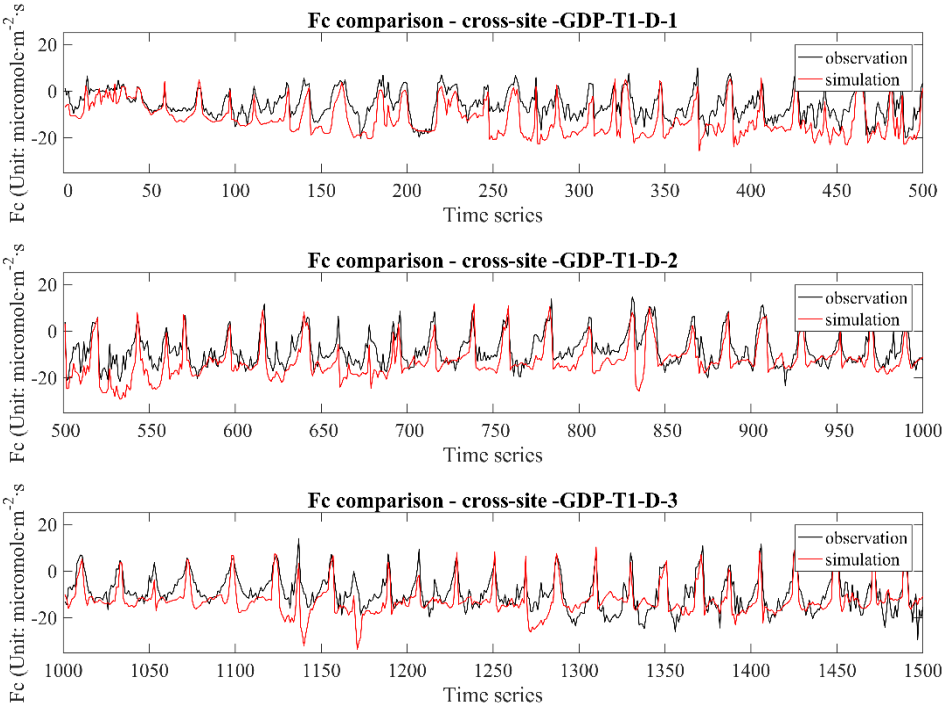
1. GDP-T1
2. All data



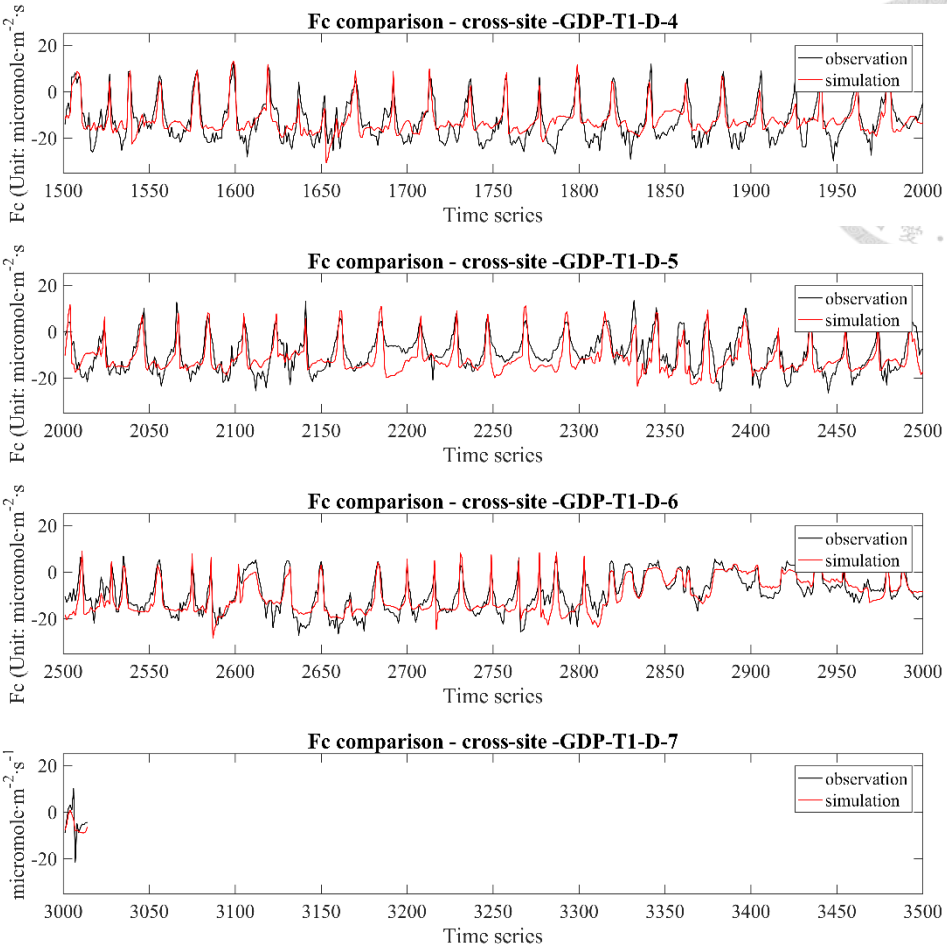
doi:10.6342/NTU201703598



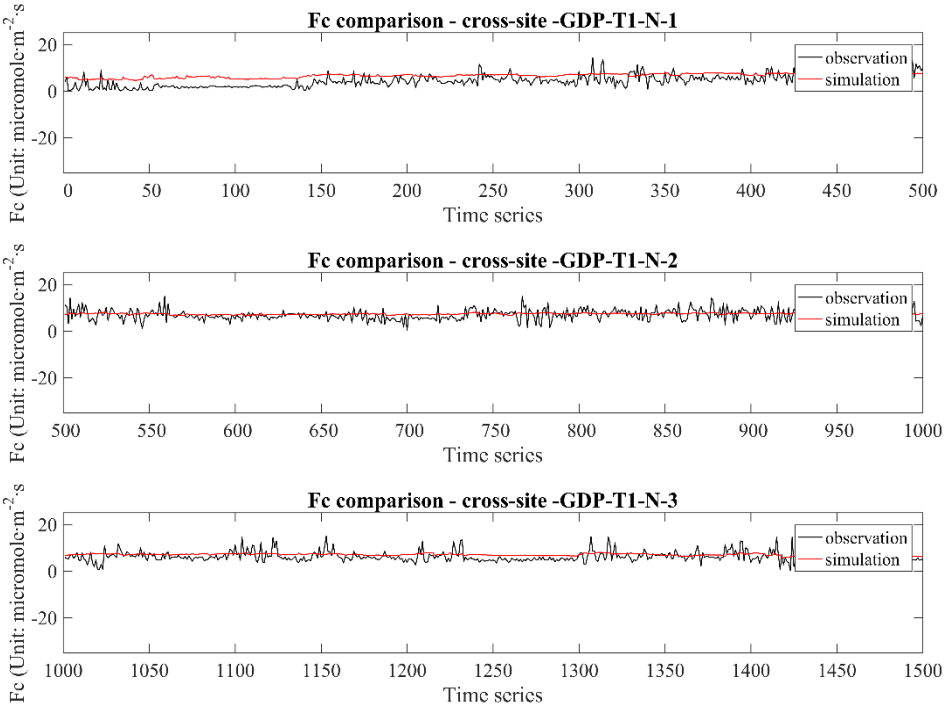
3. Day data



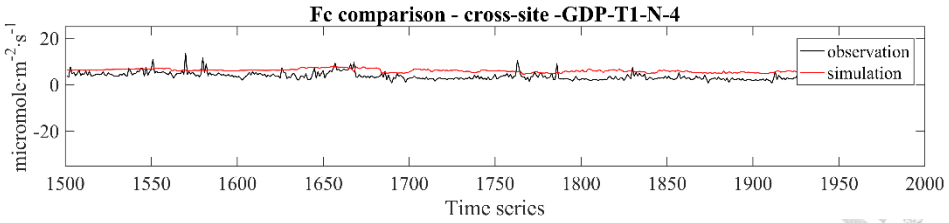
doi:10.6342/NTU201703598



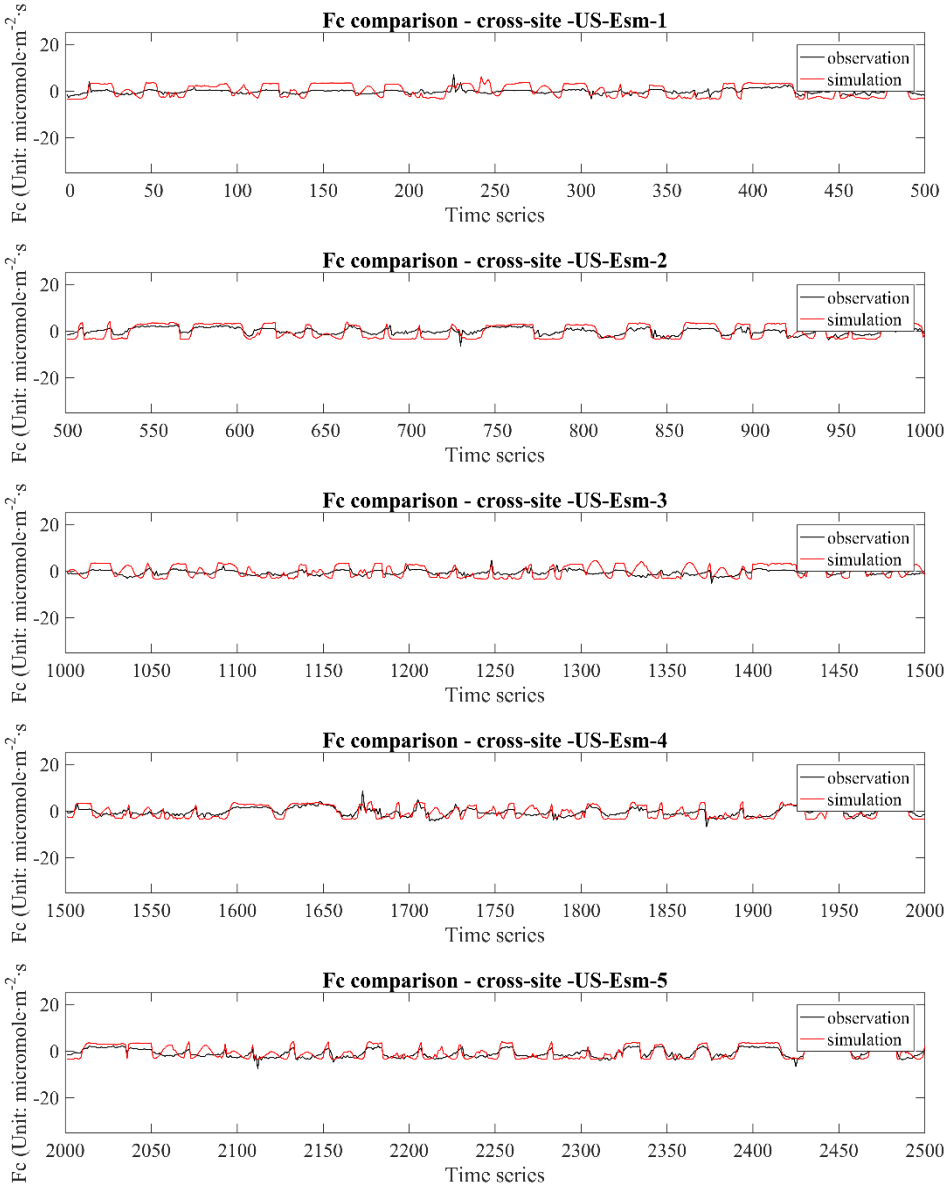
4. Night data



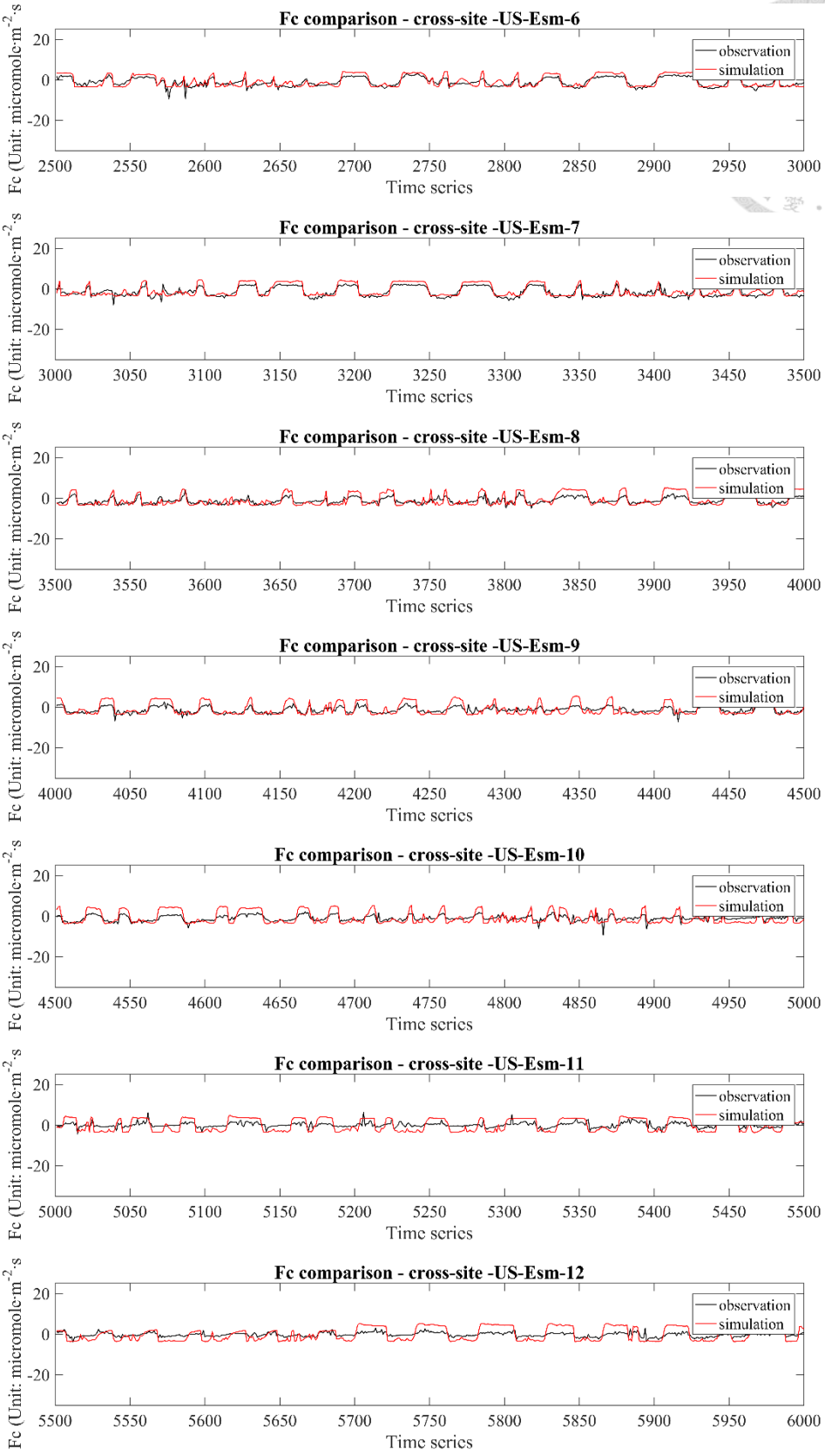
doi:10.6342/NTU201703598



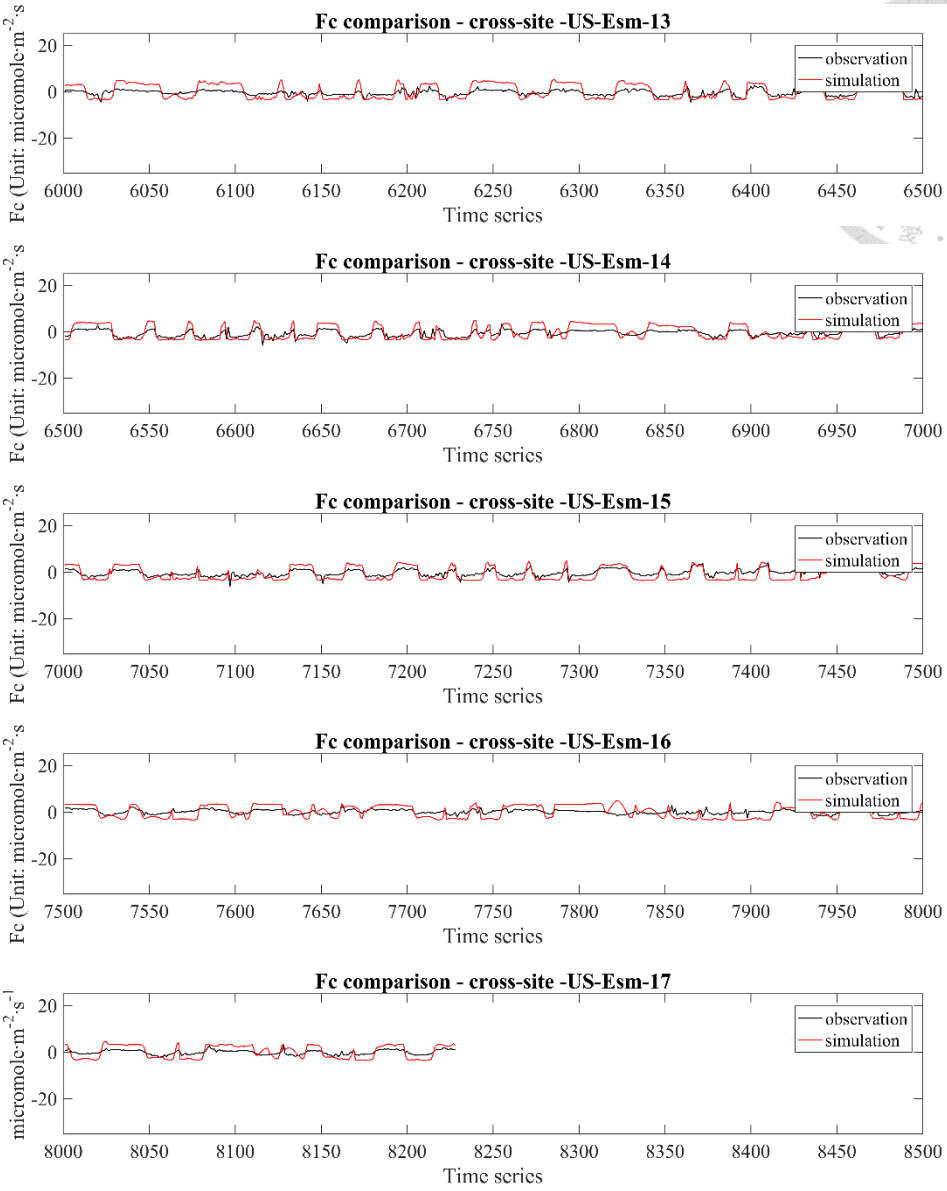
5. US-Esm
6. All data



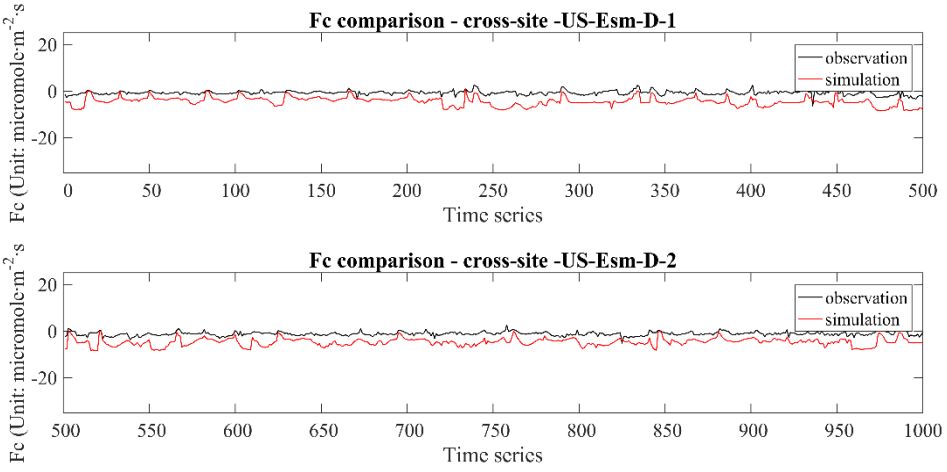
doi:10.6342/NTU201703598



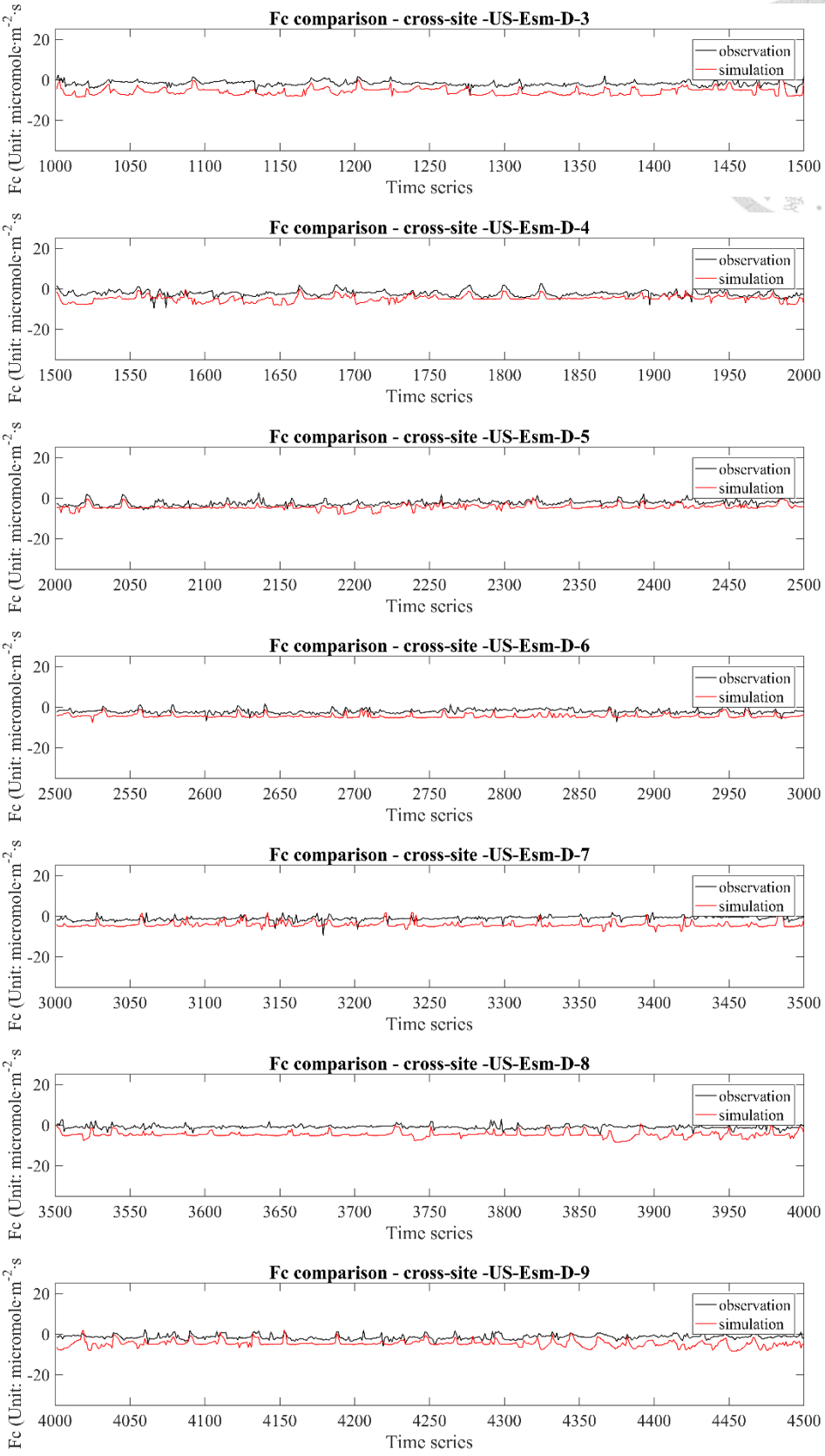
doi:10.6342/NTU201703598



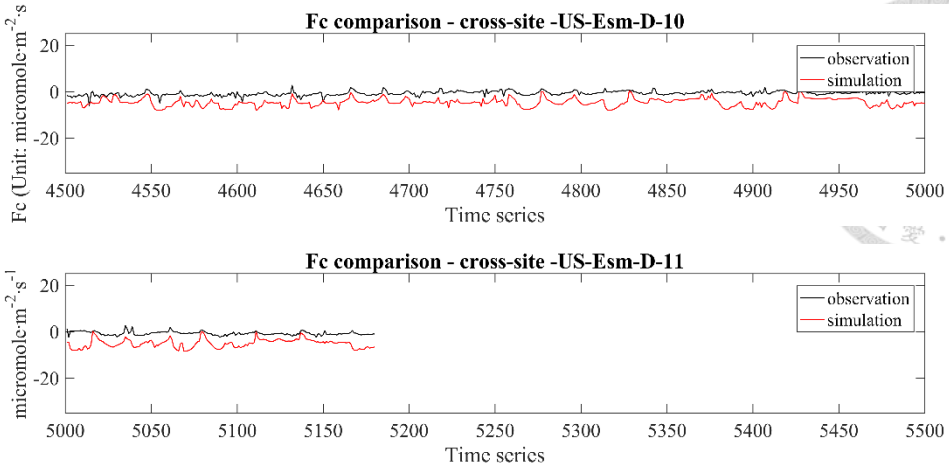
7. Day data



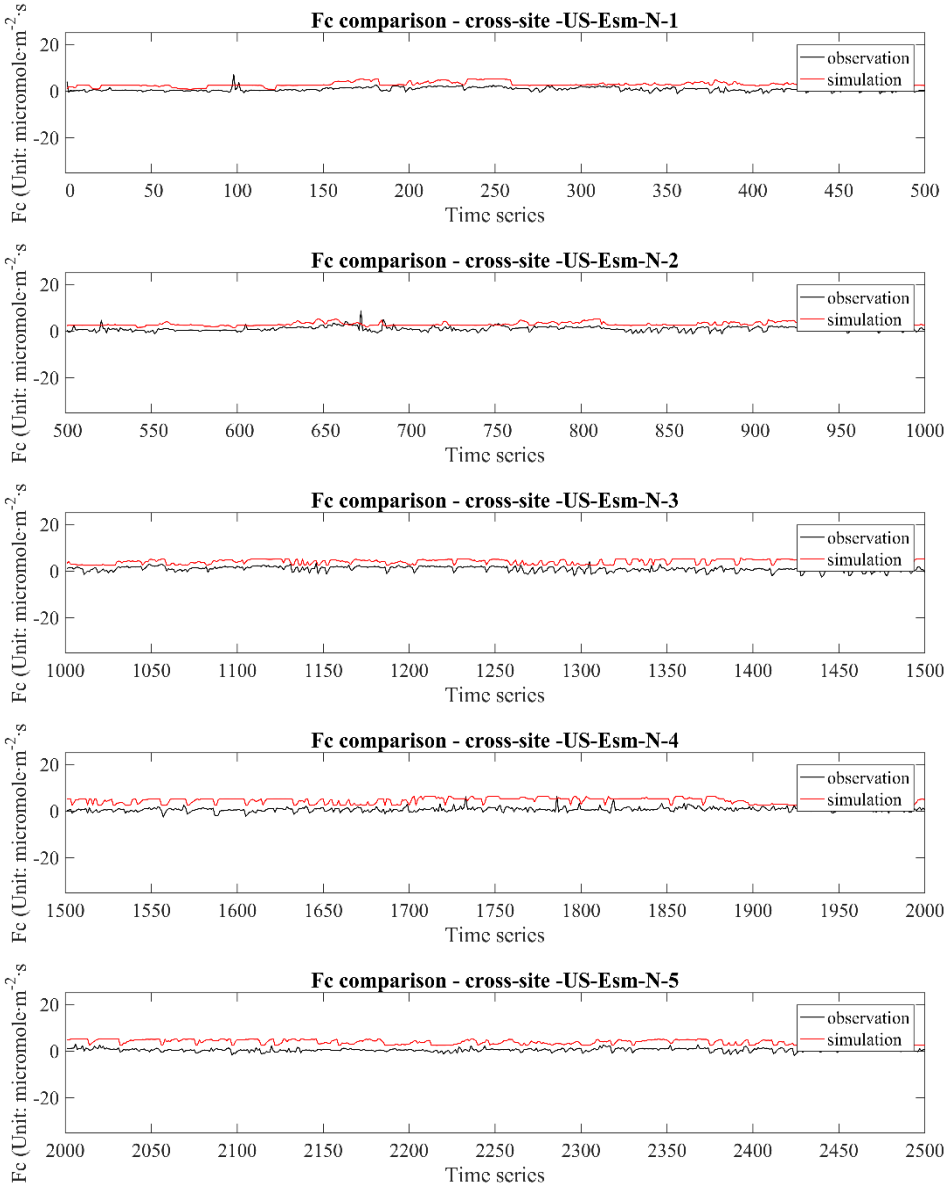
doi:10.6342/NTU201703598



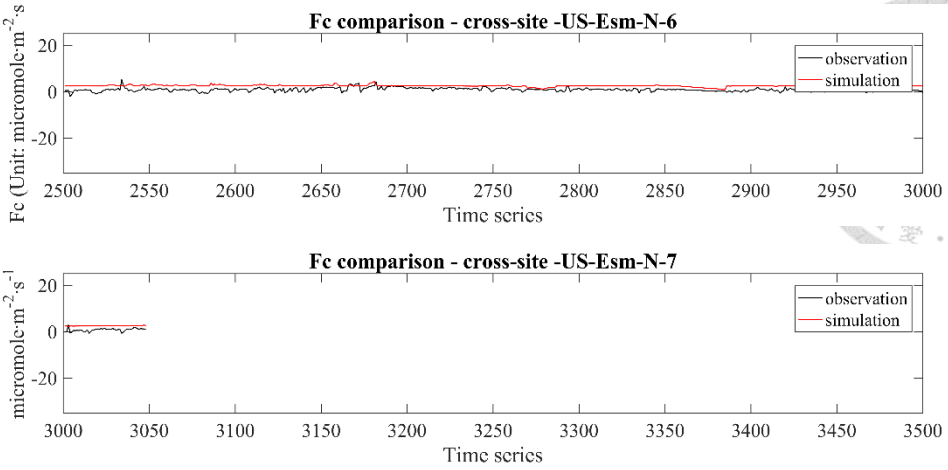
doi:10.6342/NTU201703598



8. Night data



doi:10.6342/NTU201703598



doi:10.6342/NTU201703598

Appendix V: The code used in the training phase.

```
clear all;
close all;
clc;


me='17th';    site='GDP-T1';    folder='GDP-T1';    folder2='All data';      % all data
% load the extrema
ex=load(strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',me,'_',site,'_extremum.txt'));
max=ex(5,1); min=ex(5,2);
% load the extrema
T = load(strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',me,'_',site,'_training.txt'));
INPUT=[T(:,1:4)]';OUTPUT=[T(:,5)]';


x=785; y=922; z=982;
training_A=INPUT(:,1:x);validation_A=INPUT(:,x+1:y);testing_A=INPUT(:,y+1:z);
TrainOut_A=OUTPUT(1,1:x);ValOut_A=OUTPUT(1,x+1:y);TestOut_A=OUTPUT(1,y+1:z)
;


for n = 1:1:15; % the number of neuron in the hidden layer
  for m = 1:1:10;
    net=feedforwardnet(n,'trainlm');
    net=configure(net, INPUT, OUTPUT);
    net.layers{1}.transferFcn = 'logsig';
    net.layers{2}.transferFcn = 'purelin';
    % the data division of training, validation, and testing group
    net.divideFcn='divideind';
    net.divideParam.trainInd=[(1:x)];
    net.divideParam.valInd=[(x+1:y)];
    net.divideParam.testInd=[(y+1:z)];

    net.inputweights{1,1}.initFcn='rands';
    net.biases{1}.initFcn='rands';
    net=init(net);
```


doi:10.6342/NTU201703598

```
% parameters in the training phase
net.trainParam.show=1000;
net.trainParam.epochs=5000;          % iteration
net.trainParam.goal=0.000001;        % threshold of error

net=train(net,INPUT,OUTPUT);

wt_in=net.IW{1,1};
wt_out=net.LW{2,1};
bias_in=net.b{1};
bias_out=net.b{2};

zh1=sim(net,training_A);
zh2=sim(net,validation_A);
zh3=sim(net,testing_A);

figure(m)
plotregression(TrainOut_A,zh1,strcat(site,'-TrainData n=',num2str(n),' and
    m=',num2str(m)));
saveas(gcf,[strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-R-
    square TrainOut-n=',num2str(n),' and m=',num2str(m),'.tif')]);
plotregression(ValOut_A,zh2,strcat(site,'-ValData n=',num2str(n),' and m=',num2str(m)));
saveas(gcf,[strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-R-
    square ValOut-n=',num2str(n),' and m=',num2str(m),'.tif')]);
plotregression(TestOut_A,zh3,strcat(site,'-TestData n=',num2str(n),' and
    m=',num2str(m)));
saveas(gcf,[strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-R-
    square TestOut-n=',num2str(n),' and m=',num2str(m),'.tif')]);

% denormaliztion
zh11=zh1*(max-min)+min;
zh21=zh2*(max-min)+min;
zh31=zh3*(max-min)+min;

TrainOut_A1=TrainOut_A*(max-min)+min;
ValOut_A1=ValOut_A*(max-min)+min;
```


doi:10.6342/NTU201703598

```
TestOut_A1=TestOut_A*(max-min)+min;


% RMSE
rmse=[sqrt(mse(TrainOut_A1-zh11)) sqrt(mse(ValOut_A1-zh21)) sqrt(mse(TestOut_A1-
    zh31))];
r=[regression(TrainOut_A,zh1) regression(ValOut_A,zh2)
    regression(TestOut_A,zh3)];%regression value

% MAPE
m1=sum(abs((TrainOut_A-zh1)./TrainOut_A))./length(zh1);
m2=sum(abs((ValOut_A-zh2)./ValOut_A))./length(zh2);
m3=sum(abs((TestOut_A-zh3)./TestOut_A))./length(zh3);
mape=[m1 m2 m3];

w=[wt_in,wt_out'];
save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-weight-
    ',num2str(n),'-',num2str(m),'.txt')],'w','-ascii','-double');

b=[bias_in',bias_out];
save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-bias-
    ',num2str(n),'-',num2str(m),'.txt')],'b','-ascii','-double');

Fc_o=T(:,5)'; Fc_s=[zh1 zh2 zh3];
m4=sum(abs((Fc_o-Fc_s)./Fc_o))./length(Fc_o);
rmse_all=sqrt(mse(Fc_o-Fc_s));
r_all=regression(Fc_o,Fc_s); All=[m4 rmse_all r_all];

training=[Fc_o' Fc_s' T(:,6:7)];

save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-
    training data-',num2str(n),'-',num2str(m),'.txt')],'training','-ascii','-double');
save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-All
    stat-',num2str(n),'-',num2str(m),'.txt')],'All','-ascii','-double');
  save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-
    Fc_s normalized-',num2str(n),'-',num2str(m),'.txt')],'Fc_s','-ascii','-double');
save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-
    RMSE-',num2str(n),'-',num2str(m),'.txt')],'rmse','-ascii','-double');
```


doi:10.6342/NTU201703598

```
    save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-R-
        ',num2str(n),'-',num2str(m),'.txt')],'r','-ascii','-double');
    save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-
        MAPE-',num2str(n),'-',num2str(m),'.txt')],'mape','-ascii','-double');
    save([strcat('D:\Guandu ANN\',me,' meeting\',folder,'\',folder2,'\training\',site,'-
        ',num2str(n),'-',num2str(m)),'.mat'],'net');


    close all
  end
end
```


doi:10.6342/NTU201703598

Appendix VI: The code used in the simulation phase.

```
clear all;
close all;
clc;

for z=1:1:3
    for j=1:1:3
        me='17th';num='500';    fn=['GDP-T1';'GDP-T2';'US-Esm'];
        folder=fn(z,:);

        nn=[7 5 7;5 4 3;7 5 6];                    n=nn(z,j);
        LL=[8 5 2;5 2 1;15 9 5];                   L=LL(z,j);

        fn2=['All data';'Day data';'Nig data']; folder2=fn2(j,:);

        if folder2=='Nig data'
            folder2='Night data';
        end
        type=['-A';'-D';'-N']; site=type(j,:);
        if site=='-A'
            site=''
        end

        S = load(strcat('D:\Guandu ANN\',me,'
            meeting\',folder,'\',me,'_',folder,site,'_simulation.txt'));
        ex=load(strcat('D:\Guandu ANN\',me,'
            meeting\',folder,'\',me,'_',folder,site,'_extremum.txt'));
        max=ex(5,1); min=ex(5,2);

        for m=1:1:10;
            for i=1:L;
            load(strcat('D:\Guandu ANN\',me,'
                    meeting\',folder,'\',folder2,'\training\',folder,site,'-',num2str(n),'-
                    ',num2str(m),'.mat'));
            INPUT=[S(1+str2num(num)*(i-1):str2num(num)+str2num(num)*(i-
                    1),1:4)]';
```


doi:10.6342/NTU201703598

```
        OUTPUT=[S(1+str2num(num)*(i-1):str2num(num)+str2num(num)*(i-
                1),5)]';
        zh=sim(net,INPUT);

        zh1=zh*(max-min)+min;
        OUTPUT1=OUTPUT*(max-min)+min;
        en=[OUTPUT' zh'];

        close all
        end
end

for m=1:1:10;
        for i=1:1:L
            T=load(strcat('D:\Guandu ANN\',me,'
               meeting\',folder,'\',folder2,'\simulation\comdata-',folder,site,'-
               n=',num,'-',num2str(m),'-',num2str(i),'.txt'));
            Ob=T(:,1); Si=T(:,2)'      % normalized data
            TT(1,1+str2num(num)*(i-1):str2num(num)*i)=Ob;
            TT(2,1+str2num(num)*(i-1):str2num(num)*i)=Si;
            Fc_all=TT';
        end

            [R,r]=corrcoef(Fc_all);
            OUTPUT=Fc_all(:,1);
            OUTPUT1=OUTPUT*(max-min)+min;
            zh=Fc_all(:,2);
            zh1=zh*(max-min)+min;
            M=length(Fc_all);
            b= sum(abs((OUTPUT-zh)./OUTPUT))/M;
            % MAPE calculated by normalized data

            c= (sum((OUTPUT1-zh1).^2)/M)^0.5;
            %RMSE calculated by de-normalized data
            d= R(1,2);
            e= r(1,2);
            MRRP=[b c d e];
```


doi:10.6342/NTU201703598

```
        Fc_alls=[Fc_all S(:,6:7) OUTPUT1 zh1];
        save([strcat('D:\Guandu ANN\17th
          meeting\',folder,'\',folder2,'\simulation\',folder,site,'-',num,'-all simulation
          data Fc-m=',num2str(m),'.txt')],'Fc_alls','-ascii','-double');
        save([strcat('D:\Guandu ANN\17th
          meeting\',folder,'\',folder2,'\simulation\',folder,site,'-',num,'-MRRP-
          m=',num2str(m),'.txt')],'MRRP','-ascii','-double');

    end
        for m=1:1:10;
            MRRP=load(strcat('D:\Guandu ANN\',me,'
                    meeting\',folder,'\',folder2,'\simulation\',folder,site,'-',num,'-
                    MRRP-m=',num2str(m),'.txt'));
            MAPE=MRRP(1,1);
            RMSE=MRRP(1,2);
            R=MRRP(1,3);
            P=MRRP(1,4);

            final(m,1)=MAPE;
            final(m,2)=RMSE;
            final(m,3)=R;
            final(m,4)=P;
        end
    save([strcat('D:\Guandu ANN\17th
      meeting\',folder,'\',folder2,'\simulation\',folder,site,'-',num,'-Final MRRP of
      simulation-m=',num2str(m),'.txt')],'final','-ascii','-double');

    clearvars -except z j me num fn folder

    end
end
```


doi:10.6342/NTU201703598