Department of Environment Systems
Graduate School of Frontier Sciences
The University of Tokyo

2021
Master's Thesis

New Parametrization of Mixing Length in Urban Canopy Model

(都市キャノピーモデルにおける混合距離の新しいパラメタリゼーション)

Submitted Feb 25, 2022

Adviser: Associate Professor Tomohiko Ihara

Xinyao Zhang

Table of contents

List of Figures and Tables

CHAPTER.1 Introduction

1.1 Climate change

1.1.1Global warming

In the last 50 years, there is no doubt that the air temperature is increasing around the whole world and the topic of global warming has been becoming hotter and hotter. The global temperature rise is apparently a piece of pivotal evidence demonstrating the fact that the global warming phenomenon is deteriorating ceaselessly and the trend is unalterable. To be specific, the increased greenhouse gas emission, mainly carbon dioxide, has driven the planet's average surface temperature increase by 0.9° C since the late 19th century. [1] However, given the fact that the water is with higher heat capacity ratio, the ocean temperature rise can be even more persuasive evidence. [1] That is why it is strongly called that the counter measure should be taken to controlling global warming below 2℃. Otherwise, we will risk causing destructive damage to the earth. [2]

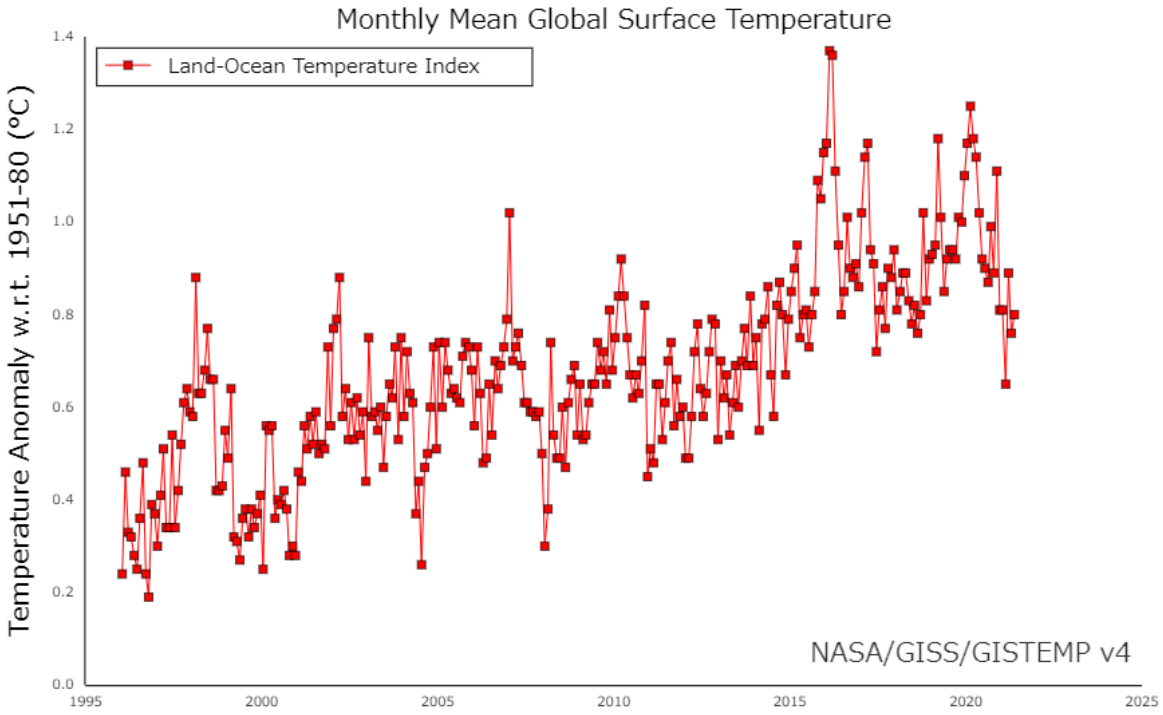

Figure 1.1. Monthly Mean Global Surface Temperature (National Aeronautics and Space Administration) [1]

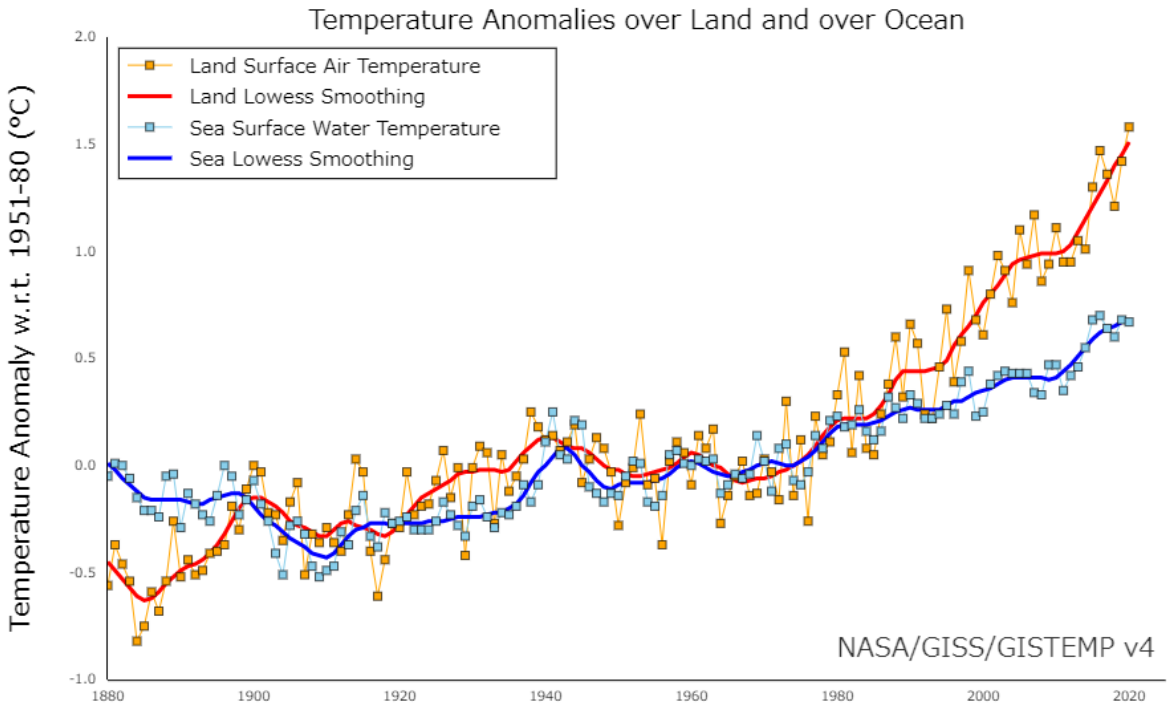

Figure 1.2. Temperature Anomalies over Land and over Ocean (National Aeronautics and Space Administration) [1]

1.1.2 Urbanization

And then, to make matters even worse, the temperature increase is really a force to be reckoned with in citizens' everyday life. As it is universally acknowledged, City is the manifestation of human the prosperity of human civilization. citizens have witnessed urban areas' prolonged economic boom and superfluous urbanization in the last century. Needless to say, the chief characteristic of urbanization is population aggregation. Actually, the consequent population shift from rural to urban area is quite conspicuous: United Nations maintained that 64% of the developing world and 86% of the developed world would be urbanized area by 2050. Also, the United Nations pointed out that nearly all global population growth from 2017 to 2030 will be by cities and the dramatic urban population increase will be about 1.1 billion. [3] As a result, the energy demand to address the enormous challenges in ASEAN has been increasing at an alarming rate and local citizens have no choice but an increasing generation from all sources. Due to coals' abundance in the region, scalability, reliability, and lower costs, power plants currently attach much importance to coals' consumption in ASEAN. [3] Thus, innumerous greenhouse gas emissions and anthropogenic

heat generation make the negative effect of temperature increase only make things worse in the urban area.

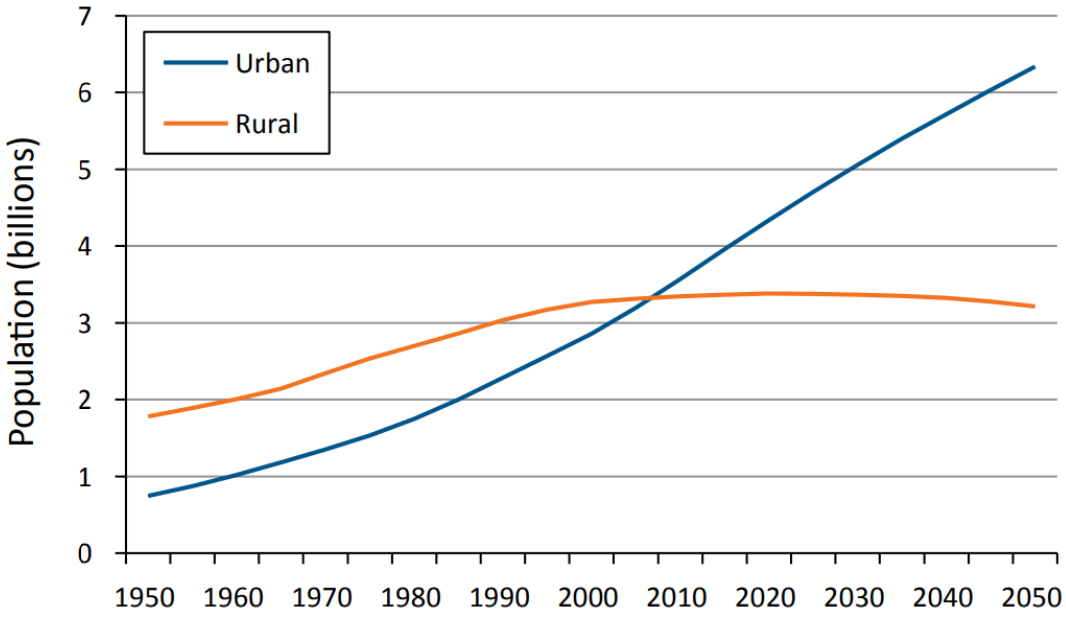

Figure 1.3. Estimated and projected populations in urban and rural settings, 1950–2050 [3]

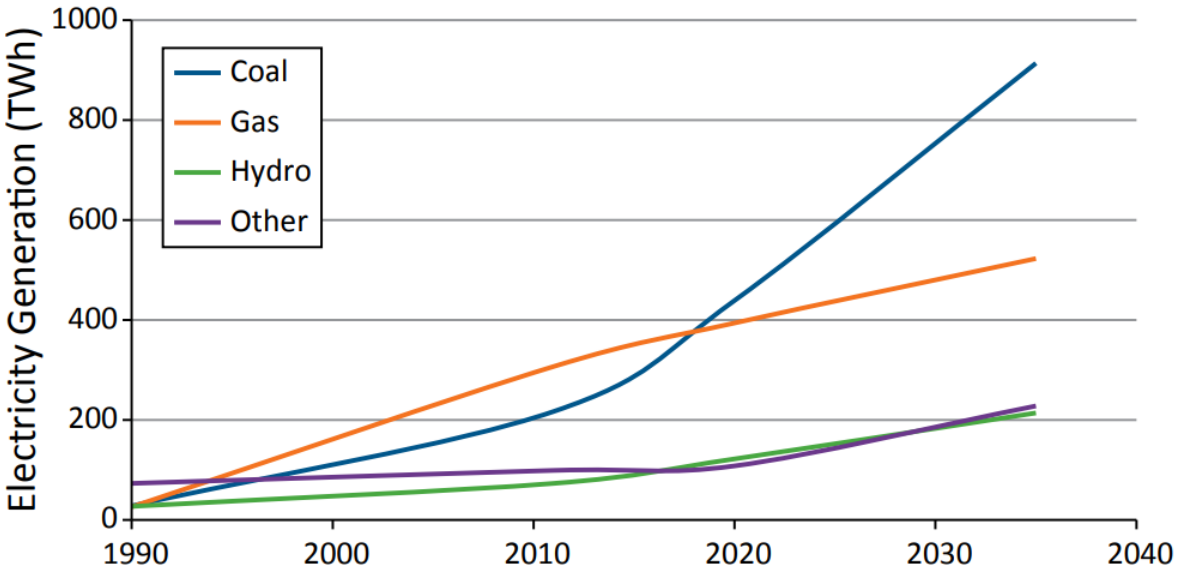

Figure 1.4. Historical and projected electricity generation in ASEAN, by source [3]

1.1.3 Urban heat island

As it is introduced previously, the contribution of global warming and urbanization to worldwide temperature has hastily increased and the significant disastrous effect on citizens is not far-reaching. As showed in Figure 1.1, from 1990 to 2010, the annual average global temperature has gone up by 0.9℃. However, things get serious in a developed country. To be specific, Japan, widely accepted as one of the successfully modernized countries, has witnessed an impressive temperature increase in the last decays. [4]

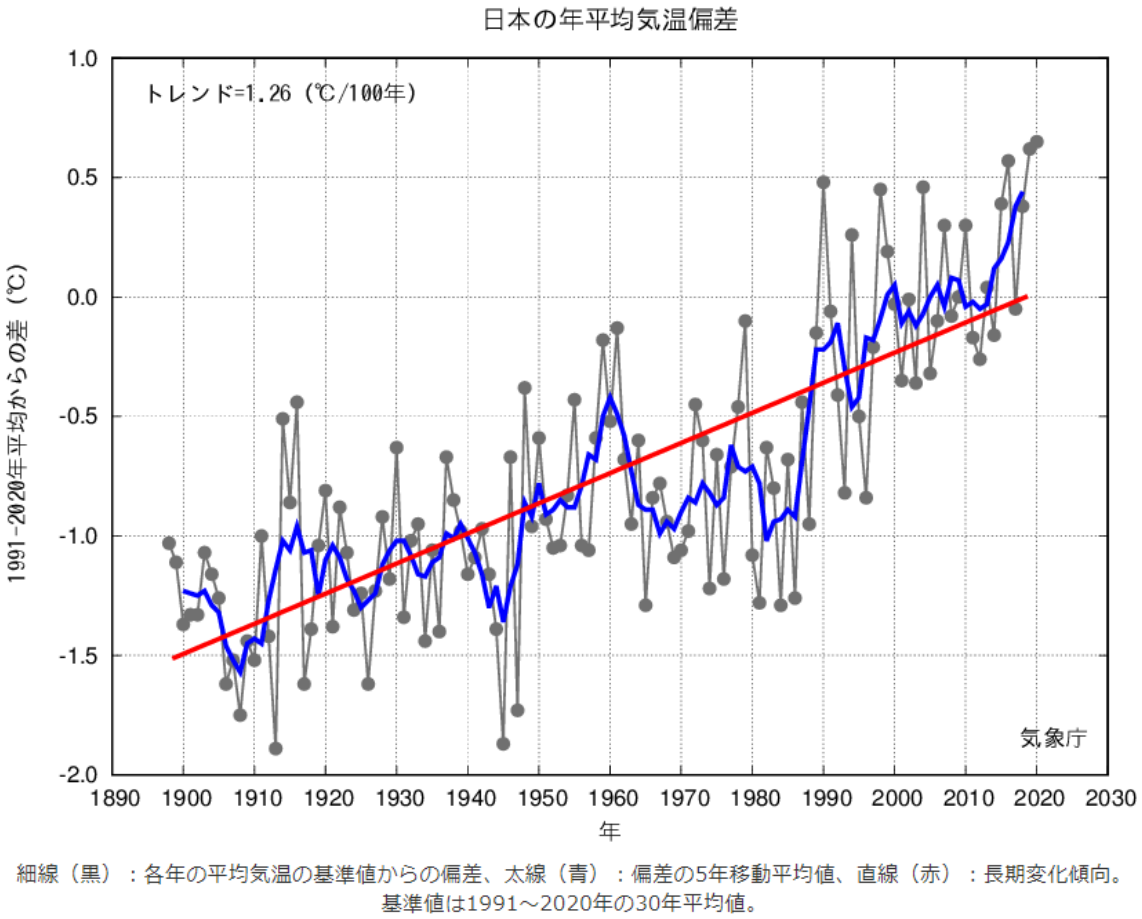

Figure 1.5. Variation of annual Japan temperature from 1890 to 2020 (Japan Meteorological Agency) [4]

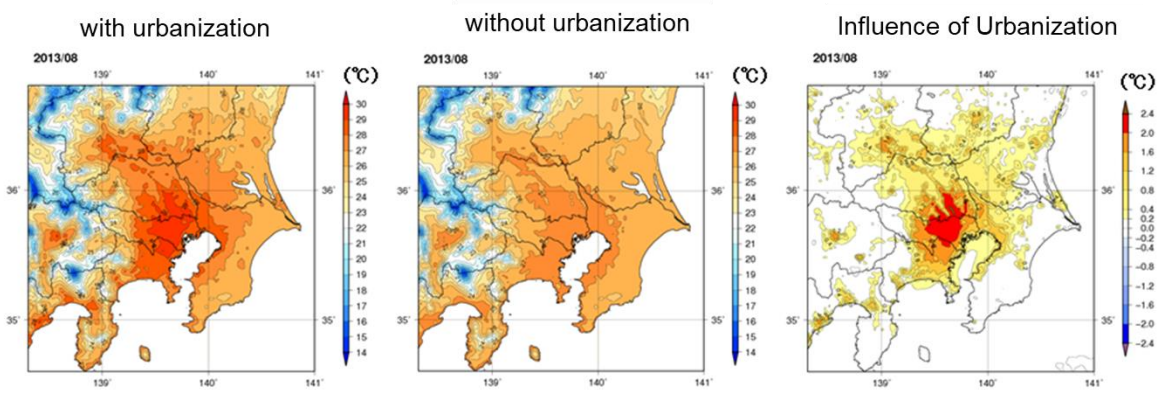

Figure 1.6. Change of the temperature due to the urbanization in Kanto district, August 2013 (Japan Meteorological Agency, 2013) [5]

Researchers tend to attribute the significant temperature increase in the urban district to UHI. The urban heat island (UHI), described as an urban area or metropolitan area that is significantly warmer than its surrounding rural areas due to human activities, has become a pivotal environmental problem in Japan [5] and conversely result in very profound negative effects on human activities, such as hyperthermia, increased CO_2 emissions due to increased cooling demand, etc. To our relief, Various local activities, especially air-conditioning, against UHI have been increasingly promoted after the Outline of Countermeasures to Urban Heat Island [6] was brought into law in March 2004.

1.2 Simulation System Introduction

1.2.1 Canopy Model and Building Energy Model

To our surprise, UHI countermeasures' influence on energy demand has been profound. Once they are installed, energy consumption, especially air-conditioning demand, will change due to the decrease in air temperature and the change in thermal insulation. Without a doubt, countermeasures' capability to guarantee a pleasant living area is a manifestation of a competitive proposal. Nevertheless, it is highly required that the energy demand increase caused by this countermeasure is the last to leave behind.

Ihara et al [7] developed CM-BEM [8] for the purpose of evaluating both the changes in thermal environment and energy consumption resulting from the installation of various UHI countermeasures from the viewpoints of UHI and global warming using annual meteorological and building energy models.

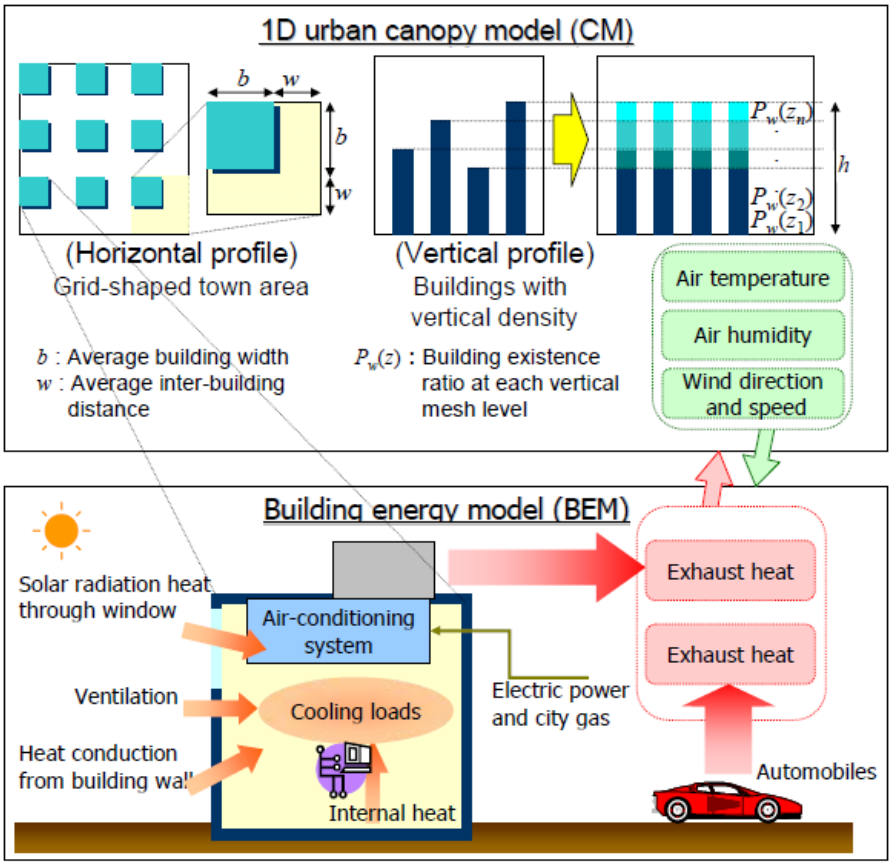

Figure 1.7. Overview of CM-BEM [7]

CM-BEM is composed of two sub-models—the canopy model (CM) [9] and the building energy model (BEM). The model describes the feedback process which is composed of the impact on a building's air-conditioning energy demand from the weather inside an urban canopy and the effects of exhaust heat on the external environment.

As implied by its name, BEM incorporates the atmospheric parameters, such as radiation, temperature, and wind speed, and calculates the cooling load in the buildings and the energy consumption of air conditioners. Consequently, the waste heat generated integrates with the air surrounding the building. To support BEM's energy consumption simulation and ensure solid periodic atmospheric conditions, CM is developed as a 1D model simulating the atmospheric parameters. The canopy model has all the atmospheric parameters initialized by WRF, accommodates the anthropogenic heat produced by the building, and utilizes a semiempirical model to describe the eternally continuing heat exchange process happening in the canopy area.

1.3 Constituents of Flux in Canopy Area

As it is universally acknowledged, the Earth's atmosphere is a kind of mixture whose three major constituents are nitrogen, oxygen, and argon. Water vapor accounts for roughly 0.25% of the atmosphere by mass. However, there is no harm in regarding the air as a kind of typical newton fluid to analyze.

The Reynolds number is widely accepted as a helpful element to predict general newton fluid flow patterns in different fluid flow situations. In the case of lower atmospheric layers, the air there is stable, and the flow tends to be dominated by laminar flow. Still, as elevation rises, the strong wind usually develops and the Reynolds number will be at a relatively high level. Under this condition, the atmospheric layer is no longer in a stable situation, which means that turbulence gains an advantage over laminar flow. Therefore, we can draw the conclusion safely that the flux exchange process happening in canopy scale is composed of mean wind speed wind shear contribution and perturbation wind speed contribution. The equation reveals the relationship between wind gradient as follows:

$$\tau = \tau_1 + \tau_t = \mu \frac{du}{dz} + \mu_t \frac{d\bar{u}}{dz}.$$

τ_1: Eddy viscosity
τ_t: Turbulent eddy viscosity
z: Elevation
μ: Eddy viscosity coefficient
μ_t: Turbulent eddy viscosity coefficient
u: Horizontal wind speed
$\bar{u}$: Horizontal wind perturbation

Further research was needed to throw light upon the ratio of perturbation contribution to the flux exchange process, happening in the canopy area. To our relief, Schubauer [10] measured turbulent shear stress's trend according to elevation. Certain statistical properties of turbulence observed in a boundary layer and fully developed pipe flow were compared. Difference and likenesses are shown in Figure 1.8. It was a somewhat shocking result that laminar flow only gained an advantage over turbulent flow in the button area of the boundary layer and turbulent flow's dominant role is as significant as what it is for a pipe that was filled by turbulent water.

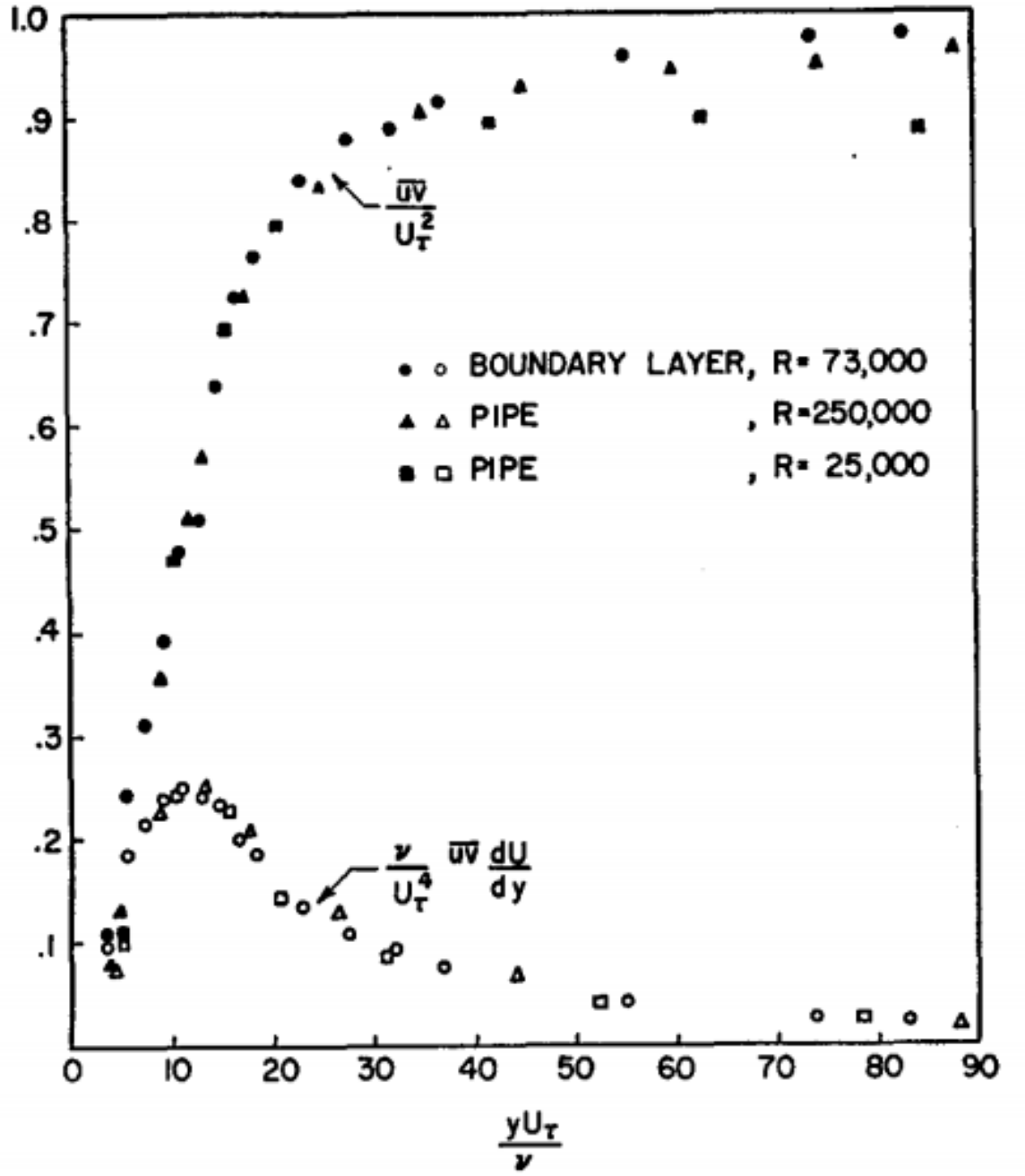

Figure 1.8. Turbulent shearing stress and rate of turbulence [10]

1.3.1 Mixing length theory

The perturbation contribution plays a lead role in canopy flux exchange. Thus, it is natural that efforts should be paid to quantify fluid's irregular fluctuations. Still, in turbulent flow, the speed of the fluid at a point is continuously undergoing changes in both magnitude and direction. The flow of wind and rivers is generally turbulent in this sense, even if the currents are gentle. In other words, it is almost impossible to observe the turbulent flow and excavate the intrinsic relations between it and other atmospheric elements.

Having collected data from the experiment, hypotheses have been put forward to reveal the intrinsic relations between Reynolds stress and time-averaged physical elements in the flow field. That's the so-called semiempirical theories. Among these semiempirical theories, the mixing length model is one of the most developed and most widely used.

$$\bar{u} = l_m \left|\frac{du}{dz}\right|,$$
$$K = \bar{u} l_m = {l_m}^2 \left|\frac{du}{dz}\right|,$$
$$\tau_t = K\frac{du}{dz} = {l_m}^2 \left|\frac{du}{dz}\right|\frac{du}{dz}.$$

l_m: The variable, mixing length, describes the linear relationship between turbulent velocity scale and wind speed gradient.
K: eddy viscosity, also eddy diffusivity.

We can simply regard mixing length as a coefficient, rooted in experimental data, describing the intrinsic relationship between time-averaged wind speed gradient and wind's fluctuate component. [11] This semiempirical hypothesis is derived from turbulent fluid's fundamental characteristics and is capable of describing the turbulent fluid's key features. Besides, there is literature [12] having proved this simplified method to be a quite accurate algorithm to simulate the complex flux exchange process.

1.4 Objectives

1. Canopy height approximation

As is introduced previously, mixing length theory suggests that there is a linear relationship between the square of mixing length and Reynolds stress at a certain altitude. The only thing I want to add is that compared to the area inside the canopy, there are no canopy elements beyond the canopy boundary and the flux exchange process is relatively free. However, at the aspect canopy district, the flux transfer process is under the influence of canopy elements' effect. Since CM is utilized to simulate the flux exchange process inside the canopy's drag effect, CM researchers should never have data collected in free wind district to simulate mixing length describing the canopy's situation. MLP will be utilized to give a better solution to canopy height approximation.

2. Mixing length parametrization

In terms of mixing length, Kondo H et al [13] attached much importance to canopy elements density profile and validated that the profiles of mixing length, mean wind speed, and plane area index were closely related. Based on Navier-Stokes equations, this study will simulate the wind speed profile for a particular mesh district and approximate the local mixing length profile. Overall characteristics describing the homogenous features for a particular canopy together with strict mathematical deduction will be utilized to describe the drag effect and accelerating effect derived from the pressure-gradient force as critically as possible. Also, LES-urban [26] describing the flux in Tokyo and Nagoya district will be referred to improve simulation performance.

3. CM-BEM performance evaluation

Since the flux transfer process is primarily based on momentum diffusivity, the performance of velocity simulation becomes one of the most important metrics to prove simulation performance. Analysis of the trend of deviation between velocity observed and velocity simulated will be utilized to validate the numeric simulation model's robustness.

The flux transfer process is rather important for a simulation model. To expound this view, I have introduced the two structures of CM-BEM briefly. Nevertheless, a further explanation should be made to describe the flux exchange process in detail and illustrate the method to improve simulation performance by utilizing atmospheric data collected from Japanese urban districts.

The numerical simulation system exploring the interaction between outdoor thermal conditions and energy consumption of air conditioning in buildings was realized by Kikegawa. [14]

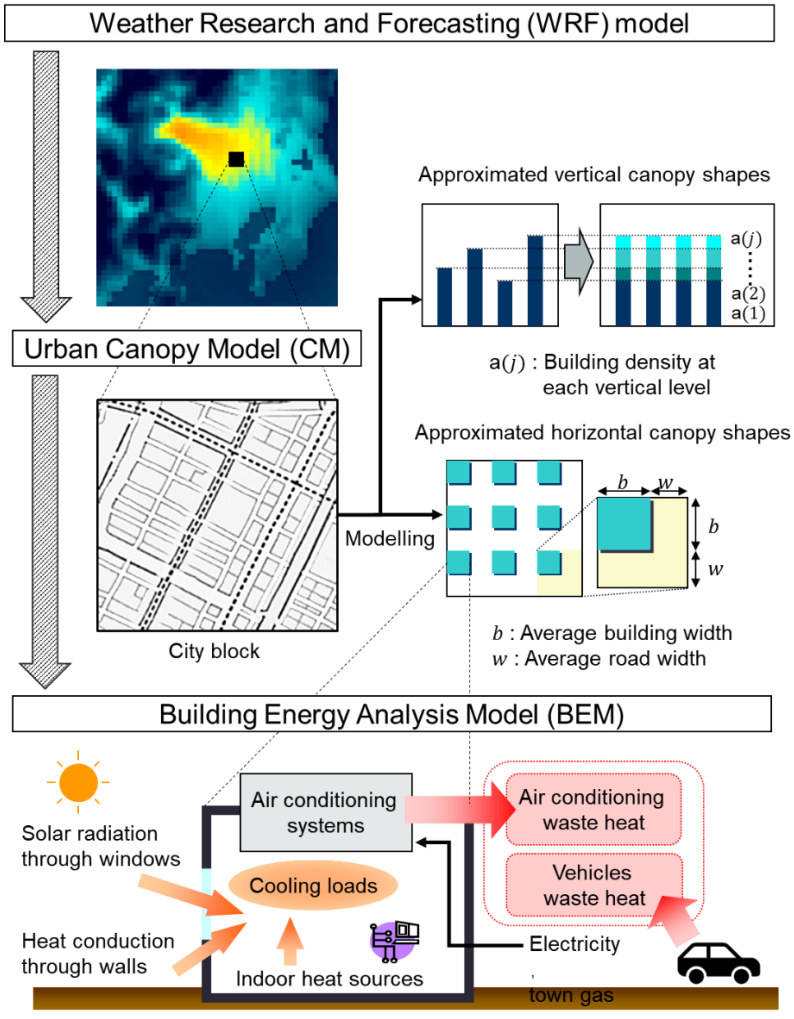

Figure 2.1 Composition of the simulation models [14]

2.1The framework of the simulation system

The structures of simulation systems were shown in Figure 2.1. [14] The simulation system consists of three sub-models: the mesoscale meteorological model, the urban canopy model, and the building energy simulation model. The Weather Research and Forecasting (WRF) model, which is developed by the National Center for Atmospheric Research (NCAR), was a next-generation mesoscale numerical weather prediction system designed for both atmospheric research and operational forecasting applications. Apparently, it is an outstanding model to simulate urban climate worldwide. However, given the fact that the mesh size of this model is too big compared to ordinary blocks in the center areas of cities, the multi-layer urban canopy model (CM) was adopted for the meteorological forecast on a city-block-scale. At last, the building energy model (BEM) was developed to explore the intrinsic relation between outdoor thermal conditions and air-conditioning energy consumption.

Figure 2.2 shows the calculation flow of the simulation system. Firstly, initial conditions and the upper boundary conditions of the CM atmosphere are obtained from WRF outputs. Thus, CM has the value of meteorological parameters initialized. Secondly, the building constructions information is imported from GIS data. This step is also very important. There is no doubt that the atmospheric layers in the upper districts are commonly with higher wind speed and richer kinetic energy. In that way, upper atmospheric layers are capable of accelerating the layers beneath. However, the canopy elements' density distribution is the last thing to ignore. Resulting from high building density, the flux exchange process is ordinarily moderated and that is why the flux transfer phenomenon is relatively fierce in rural areas. According to the building density at a certain altitude, mixing length, the coefficient to describe the wind speed exchange efficiency, is approximated firstly so that momentum diffusivity at certain elevation could be calculated consequently. Next, potential temperature and wind speed are fully utilized to estimate the fluid's situation. That is because the diffusive efficiency of heat and potential temperature is swift and violent under unstable conditions. Then, having real-time thermal components and momentum diffusivity involved, CM simulates the heat diffusivity and potential temperature diffusivity to provide a solid foundation for describing the thermal flux exchange process outdoor. At last, real-time wind speed is calculated and the thermal condition is simulated. It is really impressive that CM-BEM has the architectural response of air-conditioning energy consumption and its consequent waste heat emitted into the urban atmospheric layers, we can draw the conclusion safely that it is a strong numerical simulation system.

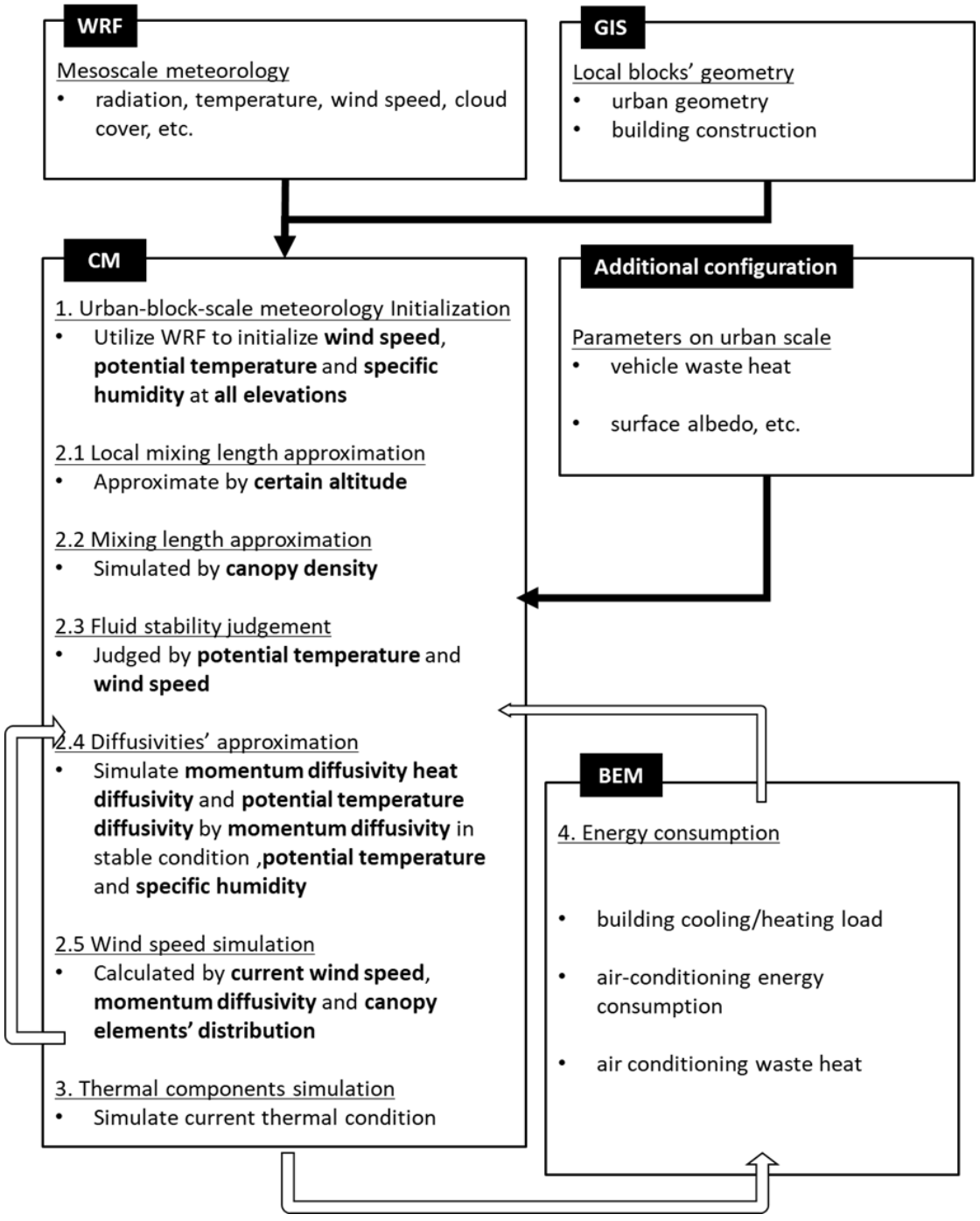

Figure 2.2 Composition of the simulation models

2.2 Diffusivity in the Multilayer UCM

The basic multilayer UCM equation for the conservation of momentum [15] is

$$\frac{\partial u}{\partial x} = \frac{1}{m}\frac{\partial}{\partial z}\left(m\,K_m\,\frac{\partial u}{\partial z}\right) - cau\left(\sqrt{u^2+v^2}\right) + g_1,$$
$$\frac{\partial v}{\partial t} = \frac{1}{m}\frac{\partial}{\partial z}\left(m\,K_m\,\frac{\partial v}{\partial z}\right) - cau\left(\sqrt{u^2+v^2}\right) + g_2,$$

As it was shown in Figure 2.1, u and v are east-west and north-south wind velocity components, respectively, z and t are the vertical coordinate and time, respectively. g_1 and g_2 represent arbitrary external forcing.b is average building width and w is average road width. λp is the mean plane area where and $P_b(z) = \frac{\lambda_p(z)}{\lambda p\,(0)}$. The urban area under consideration is divided into small rectangular areas that are equivalent to each grid in a mesoscale model, with a grid width of 500m.

The height of the local buildings is not uniform, but the distribution of the height can be considered. It is defined that the rate of actual buildings that occupy level z to all the buildings (floor density distribution, ($0 \le P_b(z) \le 1$). Where $P_b(z) = 0$, there is no building at all and thus building density is zero. Contrarily, $P_b(z) = 1$ means that the entire building area at a certain elevation, z, is actually entirely occupied by buildings. The grid including buildings is considered as a porous medium. Since the equation is one-dimensional, volume porosity is equal to surface permeability. Then volume porosity is defined as, [15]

$$m = 1 - \left\{\frac{b^2}{(w+b)^2}\right\} P_b(z),$$

Here, c is the effective drag coefficient depending on $\frac{b^2}{(b+w)^2}$ and derived from wind-tunnel experiments. [16] [17] a is a similar parameter to A_t in Uno et al. [18] But, here, the building floor density distribution Pb(z) was taken into account, with

$$a = \frac{b\ P_b(z)}{(b+w)^2 - b^2 P_b(z)},$$

K_m is vertical diffusivity for momentum. As it is introduced in Chapter 1, K_m is given by

$$K_m = {l_m}^2 \left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|,$$

The equations for potential temperature and specific humidity are,

$$\frac{\partial \theta}{\partial t} = \frac{1}{m} \frac{\partial}{\partial z} \left(K_h\, m\, \frac{\partial \theta}{\partial z} \right) + \frac{1}{c_p\, \rho}\, Q_{AS}(z,t),$$
$$\frac{\partial q_v}{\partial t} = \frac{1}{m} \frac{\partial}{\partial z} \left(K_q\, m\, \frac{\partial q_v}{\partial z} \right) + \frac{1}{l\rho}\, Q_{AL}(z,t),$$

Where h, q_v, Q_{AS}, and Q_{AL} are the potential temperature, specific humidity, anthropogenic sensible heat, and anthropogenic latent heat, respectively. Both the sensible and latent anthropogenic heat depends on the time and height, and are dynamically calculated with building energy consumption models in Kondo and Kikegawa (2003) [14]. Here, c_p is the specific heat of air, q , the air density, and l, the latent heat of vaporization.

At the aspect of the canopy atmospheric layer, turbulent flow contribution gains an advantage over laminar flow except for the surface layer. Thus, Gambo (1978) [19] put forward a formula to judge the atmospheric layer's situation. The critical Richardson number R_{fc} is 0.29 and flux Richardson number R_f is compared with it.

For the case that $R_f \geq R_{fc}$, the atmospheric situation is stable. Under this condition, momentum diffusivity is capable of describing the flux exchange process of potential temperature and specific humidity,

$$K_m = K_{\boxtimes} = K_q = {l_m}^2 \left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|,$$

On the other hand, when $R_f < R_{fc}$, the situation is not stable anymore. S_M and S_H are the function of R_f,

$$K_M = {l_m}^2 \left| \frac{d\sqrt{u^2 + v^2}}{dz} \right| \frac{{S_M}^{1/2}}{\sqrt{C}} S_H \left(1 - R_f\right)^{1/2},$$

$$K_q = K_h = {l_m}^2 \left| \frac{d\sqrt{u^2 + v^2}}{dz} \right| \frac{{S_M}^{3/2}}{\sqrt{C}} \left(1 - R_f\right)^{1/2},$$

2.3 Mixing length in the Multilayer UCM

As is introduced in chapter 2.2, the momentum diffusivity coefficient, describing the capability of the momentum flux exchange process, depends on the wind speed gradient and mixing length. When the atmospheric situation is stable, the momentum diffusivity is a rather competitive candidate to describe the flux transfer process of potential temperature and specific humidity. However, when the Richardson number is at a low level, it is assumed that the atmospheric layers are under unstable conditions. While it is maintained that the equation to calculate momentum diffusivity is still capable of describing the efficiency of the momentum exchange process, the Richardson number is utilized to modify the formula to calculate the diffusivity of potential temperature and specific humidity. All in all, it is momentum diffusivity in a stable fluid condition, depending on mixing length, not any other element that is the decisive factor to flux exchange process happening in Multilayer UCM.

Lettau [20] has described one way of describing the mixing length profile. Attempts by the author to use this and another method on observations from several different sources have yielded a bewildering variety of vertical distributions of mixing length. The most common type of distribution is an increase up to 200 or 300 meters followed by a more or less constant value, usually of less than 100 meters, at higher levels. This kind of distribution results from Lettau's analysis of the Leipzig wind profile (Fig. 2.3)

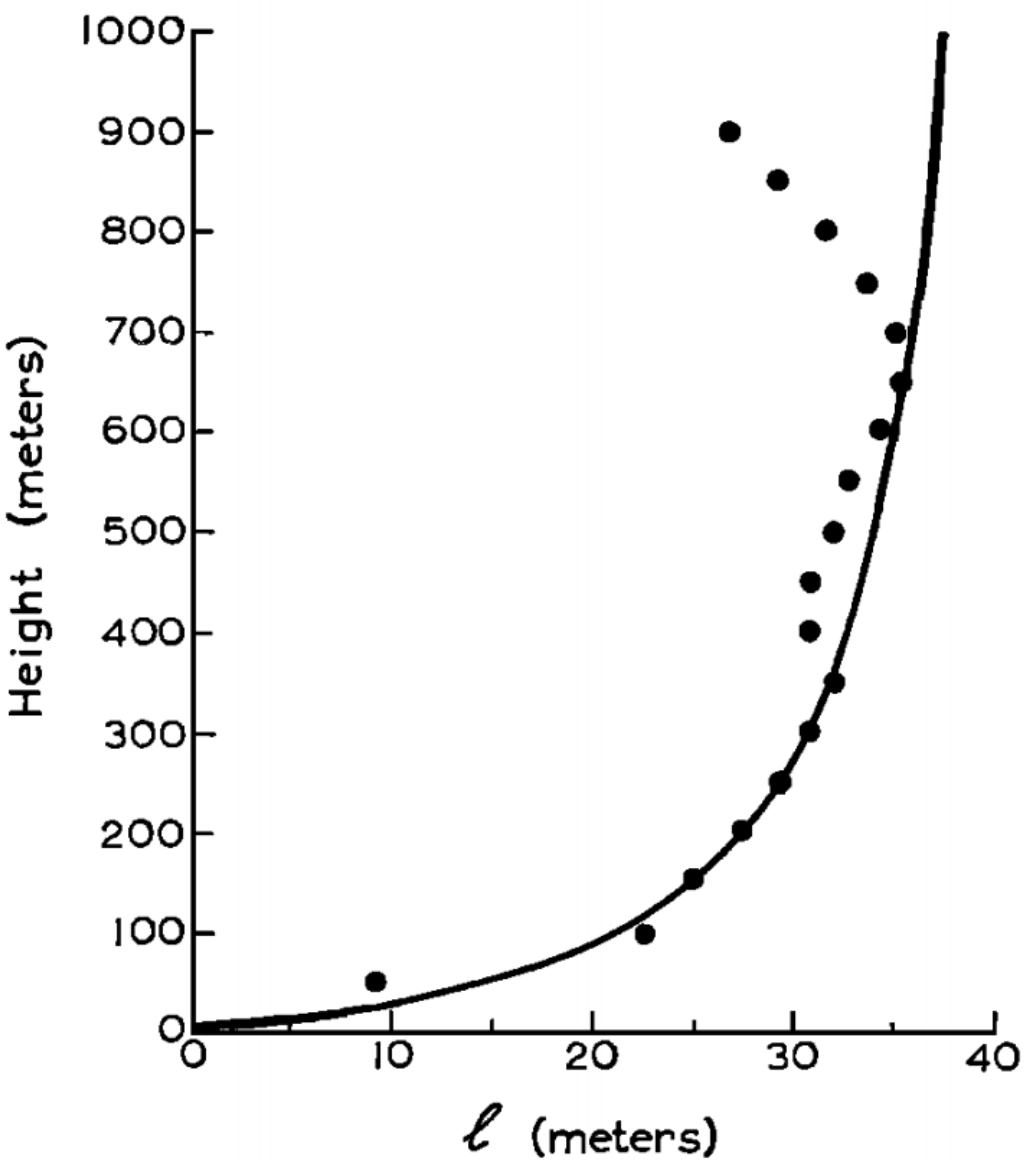

Figure 2.2 Vertical distribution of mixing length for Lettau's [20] analysis of the Leipzig wind profile

The proposal has been put forward by Watanabe and Kondo [21] to simulate mixing length. The author regards the canopy area as a vertically homogeneous canopy.

a) For the surface layer without any canopy element, the turbulent motion is limited by the ground surface only, and thus the characteristic length scale consists only of the height z. k is von Kármán constant

$$\frac{l_m(z)}{z} = k,$$

b) For the tall and dense canopy, the whole canopy is occupied by canopy elements. In this case, the turbulent motion is modified by the canopy elements through the action of the form drag. Thus, z is no longer important. However, the term $(c_d a_0)^{-1}$ where a_0 is the leaf area density, is the significant length scale. From dimensional analysis, the author predicted that

$$l_i \, c_d \, a_0 = \alpha,$$

Kondo and Akashi (1976) [11] applied von Kármán 's similarity hypotheses to the exponential wind profile (Inoue, 1963) [22], and obtained a value of $\alpha = 2k^3$ (=0.128) for a general canopy.

c) For the general case, the canopy is assumed to be a vertically homogeneous canopy. Under this case, turbulent motion is limited by the ground surface as well as the canopy elements, so that both z and l_m are significant. The author normalized mixing length and elevation length scale

$$f_{(\eta)} = \frac{l_m(z)}{l_i},$$

$$\eta \equiv \frac{z}{l_i} = \frac{c_d \, a_0 \, z}{\alpha},$$

As η increases, the limitation due to the canopy elements becomes dominant, and the mixing length approaches its intrinsic value. As η decreases, on the other hand, the limitation due to

the ground surface becomes dominant, and the mixing length approaches the value of the atmospheric surface layer. These are expressed by

$$\eta \to 0:$$
$$\begin{cases} f(\eta) \to k\eta \\ df/d\eta \to k' \end{cases}$$

And

$$\eta \to \infty:$$
$$\begin{cases} f(\eta) \to 1 \\ df/d\eta \to 0' \end{cases}$$

And at last, the formula for the general canopy was given as

$$l_m(z) = 2k^3 \frac{1}{ca}(1 - e^{-\eta}) = k l_i (1 - e^{-\eta}),$$
$$\eta \equiv \frac{z}{l_i} = \frac{caz}{2k^2},$$

Here, where a is leaf area density from Watanabe and Kondo (1990) [21]. In the urban canopy, a is usually the ratio of the total frontal area at level z, included in the small range dz, to the total fractional volume occupied by air in the canopy (Coceal and Belcher 2004). [22] Then, a may be modified to

$$\mathrm{a} = \frac{\left(\lambda_p(z)/\lambda_p(0)\right) b\, dz}{\left\{(b+w)^2 - b^2\left(\lambda_p(z)/\lambda_p(0)\right)\right\} dz} = \frac{\left(\lambda_p(z)/\lambda_p(0)\right) b}{(b+w)^2 - b^2\left(\lambda_p(z)/\lambda_p(0)\right)},$$

and c is a function of $\lambda_p(0)$, which was derived from the wind-tunnel experiment of Maruyama(1993). [23]

As it is described in Figure 2.2, mixing length has a tendency to decrease at high altitudes. Thus, Blackadar(1962) [24] has set a limitation to mixing length: $L_{2(z)} = kz\left(1 + \frac{kz}{L_B}\right)^{-1}$.

Blackadar originally gave $L_B = 0.00027G/f$, where G is the geostrophic wind magnitude and f is the Coriolis parameter. Mellor and Yamada (1974) [25] suggested $L_B = 0.3u_*/f$, where u_* is the friction velocity under neutral stability.

2.4 Summary for Current Mixing length Approximating Method

Apparently, Watanabe and Kondo (1990) [21] have given a very constructive solution. As early as the 1950s, researchers have put forward proposals to canopy whose surface layer is without any canopy element and canopy occupied with canopy elements. Both of them were rather extreme and only existed under the ideal situation. The author synthetically considered these two situations and give the porous medium canopy a general solution. After a rigorous mathematical deriving, the solution reappeared the feature of mixing length profile extrapolated by analysis of the Leipzig wind profile.

However, the are several negative points that remain. Firstly, the solution is given for the canopy area, where the canopy element's density is always positive, while the free wind district is generally involved in the simulation domain. Since the top boundary of the multilayer UCM is usually set to a much higher level than canopy height, we should take the boundary of the urban canopy seriously. In the canopy area, the atmospheric layers are accelerated by the upper atmospheric layers distinguished with relatively richer kinetic energy and moderated by buildings' drag effect. Nevertheless, the area beyond the canopy boundary is the so-called free wind area, where there is no canopy element, and momentum exchange processes are generally more significant there. Thus, having this area involved in atmospheric components simulated may risk resulting in an overestimation of flux exchange efficiency. Also, we should keep in mind that the mixing length method is essentially a kind of semi-empirical solution. No matter how intricate the mathematical derivation is, it cannot guarantee a precise result for the urban commercial district since the data source is not obtained from a real modernized district.

In the next chapter, I will introduce how the dataset derived from urban areas is utilized to approximate canopy height for urban canopies.

CHAPTER.3 Database Utilized to Simulate Flux Exchange Process

3.1 Database Introduction

As is introduced in chapter 2, the objective of this study is to simulate the flux exchange process happening in the canopy area. In other words, I am in badly need of a huge dataset of vertical profiles of turbulent flow statistics, while collecting the real-time meteorological data in upper atmospheric layers remains a difficult problem for researchers. To our relief, recently, Kanda et al. (2013; KA13 hereinafter) [26] performed a large-eddy simulation (LES) for each of 100 areas of 1 km^2, using real morphology in Tokyo and Nagoya. They constructed a database for profiles of horizontally-averaged wind speed, turbulent kinetic energy (TKE), and vertical momentum flux. KA13 derived a new parametrization for aerodynamic roughness length and mainly investigated phenomena above the urban canopy. In the present work, the same LES database is used to investigate the mixing length within the urban canopy.

3.2 Numeric Simulation Model

The parallelized large-eddy simulation model (PALM) was used in Kanda's [26] database simulation (Inagaki et al. 2012). [27] The numerical schemes used were the second-order Piacsek–Williams form C3 scheme for advection and the third-order Runge–Kutta scheme for time integration. The fractional step method ensures incompressibility, and the Temperton algorithm for the fast Fourier transform (FFT) was used to solve for the resulting Poisson equation for the perturbation pressure (Raasch and Schröter 2001). [28] An implicit filtering of the governing equations followed the Schumann volume-balance approach, while turbulence closure for the LES was based on the modified Smagorinsky model, with the flux–gradient relationships of the 1.5-order Deardorff scheme.

The mask method used in PALM to explicitly resolve solid obstacles on a rectangular grid, which was based on the method of Kanda et al. (2004), [29] proceeded as follows: numerical computation was executed at each grid point as if there were no obstacles, and forcing induced by physical boundary conditions was introduced to grid points corresponding to obstacle surfaces, wherein zero wall-normal velocities define the wall positions (Letzel et al. 2008). [30] The simplified and optimized mask method used in PALM reduced a three-dimensional obstacle into two-dimensional topography, improving the performance and minimizing the computational load. The wall function was based on Monin–Obukhov similarity theory and prescribed a Prandtl layer for each wall surface (Letzel 2007). [31]

3.3 Computational Set-Up

The streamwise (x), spanwise (y), and vertical (z) sizes of the computational domain were 1,000 m (L_x), 1,000 m (L_y), and 600 m (L_z), respectively, with a uniform spatial resolution of 2 m. The total grid number was 500 by 500 by 300 along the x, y, and z axes, respectively. The bottom surface consisted of a realistic building geometry or idealized simple arrays of buildings, with the local topographic relief ignored in order to purely focus on the urban geometrical effect. The streamwise direction was set from west to east for all runs. Because of the large number of districts used and the inherent diversity and weak regulation of building construction in Tokyo, the surface structures, such as the major street angles relative to the given wind direction, were variable.

The neutrally-stratified atmospheric boundary layer was simulated by initially setting a uniform streamwise velocity (u) of 3 m s^{-1} with zero surface heat flux. All surfaces were non-slip, whereas the top boundary was slip. Cyclic conditions were set for both horizontal directions, and the volume flux of flow was conserved by adjusting the streamwise static pressure gradient. The simulation was continued until the flow reached a fully-developed quasi-steady turbulent state. The integration time for each LES run was variable (around 5 hours; results from the last two hours were used for all investigations).

Conventional aerodynamic parametrizations are based on the logarithmic velocity law and are thus derived from surface-layer scaling with neutral stratification. Real urban boundary layers are often composed of two layers: a surface layer in which mechanical turbulence is dominant and a mixed layer in which thermal turbulence is dominant. Therefore, a fully developed urban atmospheric boundary layer with neutral stratification up to 600 m should rarely exist in practice. The current extreme numerical set-up without thermal effects can vertically extend the surface layer, thereby ensuring the existence of a logarithmic wind profile region or inertial sublayer. This set-up is used simply for the derivation of aerodynamic parameters for urban surfaces, which is indispensable under the framework of fluid dynamics.

The outer-layer fluctuation, whose horizontal scale is much larger than that of surface-layer eddies, has little influence on the momentum transport and logarithmic wind profile in the surface layer, as observed in wind-tunnel experiments (Hattori et al. 2010), [32] outdoor experiments (Inagaki and Kanda 2008, 2010), [33] [34] and numerical simulations (Castillo et

al. 2011). [35] Therefore, the aerodynamic parameters will be valid even under outer-layer fluctuations.

3.4 Morphology Information

The original building data, MAPCUBE, were commercially provided from the CAD Centre Corporation in Japan. The original building data format was a two-dimensional array of building heights with a horizontal resolution of 1 m. One file covered 4,000 m in the west-east direction (x) and 3,000 m in the south-north direction (y). This file was divided into 12 areas of 1,000 × 1,000 m^2 with downsizing into a 2-m resolution for use in PALM, which reads this data format and automatically converts the values to either 0 (= air) or 1 (= solid buildings) integer values of a three-dimensional field for the masking method. Among all Tokyo (622 km^2) and Nagoya (322 km^2) files, 107 representative districts (97 in Tokyo and 10 in Nagoya) were selected for the LES runs. As an example, Fig. 3.1 shows maps of building height for three distinctive surface geometries: (a) a cluster of skyscrapers, (b) business or commercial districts with mostly mid-height buildings, and a few isolated towers, and (c) a low residential area. Only buildings were considered and objects such as vegetation and automobiles were ignored.

To ensure robust parametrization, 23 simple arrays of buildings were added to the LES database. The arrays used were only square or staggered. Among 23 cases, 16 were arrays of homogeneous cubes or cuboids, and 7 were arrays of cuboids of variable height. The additional LES results for these simple arrays of buildings could be used for comparison with experimental data for the same geometries from the literature. Moreover, the results are useful for clarifying the differences in statistics among real urban geometries and simple artificial building arrays.

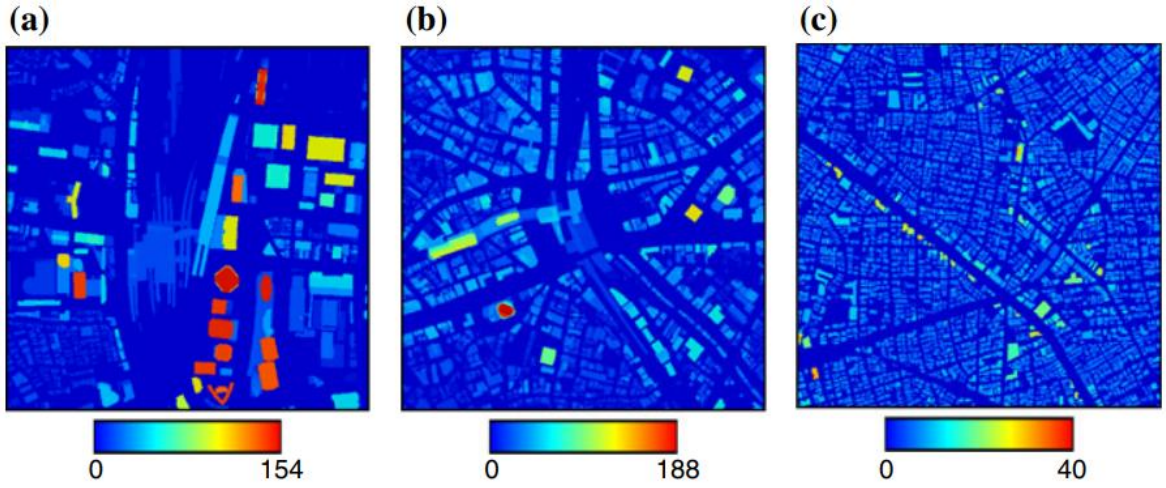

Figure 3.1 Maps of building height for three different urban surfaces. a Skyscrapers (ID97), b business district (ID 96), and c residential area (ID76)

3.5 LES Result

The LES results for 107 different urban surfaces, together with 23 simple arrays of buildings, provide a database of urban surface properties and turbulent flow statistics. Hereafter we call this database LES-Urban, which is composed of three different files for each urban surface: the color map of building height to approximate bulk geometric parameters, the header file containing the bulk aerodynamic and geometric variables (Table 1), and the profile file containing the horizontally-averaged turbulent statistics and corresponding layered geometric parameters. The horizontally-averaged turbulence statistics include wind velocity (u), the standard deviations of u, v, and w, normalized by the friction velocity u^*, turbulent kinetic energy normalized by u^*, and total vertical momentum flux (Reynolds stress + dispersive stress).

Table 3.1 Example for the list of the header files of LES-Urban, containing the bulk geometric and aerodynamic variables [26]

ID	FG	λ_f	λ_p	H_{ave} (m)	H_{max} (m)	σ_H (m)	H_{uw} (m)	u^+ (ms^{-1})	$d(1)$ (m)	$z_0(1)$ (m)	$d(2)$ (m)	$z_0(2)$ (m)	$k(2)$	C_{ld}	D_s/D_f
1	1	0.15	0.25	11.3	39.9	7.9	36	0.23	13.2	0.7	9.8	1.2	0.36	0.043	0.231
2	1	0.17	0.21	15.6	53.5	10.8	52	0.25	19.5	1.1	14	1.9	0.35	0.051	0.16
3	1	0.15	0.29	14.7	128.1	10.7	52	0.26	15.4	2.2	18.8	1.9	0.41	0.097	0.124
4	1	0.18	0.32	12.2	53.7	8	44	0.25	19.7	1.3	13.2	2.4	0.34	0.081	0.118
5	1	0.15	0.27	9.5	43.9	5.5	34	0.26	12.7	1.2	10	1.6	0.37	0.102	0.161
6	1	0.23	0.25	10.9	94.7	8.5	90	0.27	24.7	1.5	16.7	2.5	0.36	0.096	0.091
7	1	0.15	0.2	6.3	16.7	2	16	0.18	5.6	0.3	3.9	0.5	0.36	0.087	0.356
8	1	0.29	0.32	9.3	47	6.3	40	0.24	16.3	1.1	10.9	1.9	0.35	0.125	0.112
9	1	0.22	0.27	11.1	49.3	8.5	46	0.27	19.5	1.4	13.8	2.5	0.35	0.119	0.099
10	1	0.43	0.36	13.1	217	10.7	52	0.27	37.3	2	28.6	2.5	0.38	0.339	0.07
11	1	0.44	0.38	11.1	93.7	9.1	78	0.31	28.6	2.7	22	3.9	0.37	0.16	0.052
12	1	0.39	0.35	11.5	64.2	8.9	56	0.28	21.6	2.2	16.6	3.3	0.36	0.158	0.068
13	1	0.37	0.36	11.2	101.8	9.7	100	0.31	31.4	2.3	33.3	2.1	0.41	0.199	0.034
14	1	0.36	0.39	8.5	44.1	4.8	38	0.24	13.8	1.2	11.3	1.6	0.37	0.263	0.107
15	1	0.38	0.4	9.2	90.5	8.8	90	0.32	22.8	2.4	24.1	2.2	0.41	0.397	0.071
16	1	0.34	0.36	11.9	91.4	10.7	58	0.3	17.2	2.7	19.2	2.4	0.41	0.111	0.058
17	1	0.28	0.33	8.2	75.4	5.4	62	0.26	20.3	0.6	14	1.1	0.44	0.161	0.105
18	1	0.31	0.39	8.3	75.4	6	76	0.24	18.5	1.2	16.2	1.4	0.39	0.303	0.107
19	1	0.29	0.35	8.1	48.6	4.6	38	0.22	12.6	1	9.5	1.4	0.37	0.204	0.127
20	1	0.25	0.31	10.4	41.7	6.7	40	0.25	14	1.7	11.6	2.2	0.37	0.191	0.095

ID is the sequential number of the dataset and denotes the surface data category: 1: real three-dimensional geometry, 20: square array of homogeneous building, 21: staggered array of homogeneous building, 30: square array with variable building height, 31: staggered array with variable building height. λ_f is the frontal area index, λ_p is the plane area index, H_{ave} is the average building height, H_{max} is the maximum building height, σ_H is the standard deviation of building height, H_{uw} is the height of the momentum flux peak, u^* is the friction velocity, $d(1)$ is the displacement height (Method 1), $z_0(1)$ is the roughness length (Method

1), $d(2)$ is the displacement height (Method 2), $z_0(2)$ is the roughness length (Method 2), $k(2)$ is the variable von Karman constant (only Method 2), and D_s/D_f is the floor drag relative to the total drag.

3.6 The Value of This Dataset to My Study

Firstly, it is important to stress the size of the computational domain describing the urban aerodynamic objective law. The streamwise (x), spanwise (y), and vertical (z) sizes of the computational domain were 1,000 m (L_x), 1,000 m (L_y), and 600 m (L_z), respectively, while the horizontal mesh size of CM is 500m*500m. Kanda simulated the profile of mean atmospheric parameters on a relatively larger scale thus we can draw the conclusion safely that a robust solution was given to the urban canopy, which is composed of the homogeneous porous medium of canopy elements. Also, the atmospheric component is meticulously simulated in a vertical direction: intervals are as small as 2m and that is rather impressive.

Secondly, I want to emphasize easy access. Once researchers decide to build a tower in the center area of a big city, like Tokyo, there is seldom an approach to lessen the financial burden. Besides, the tower itself will result in the overestimation of canopy density and the pressure gradient force remains hard to ignore.

Next, the database has a solid physical background. The intricately detailed scale model, parallelized large-eddy simulation model (PALM), was used to guarantee a good simulation performance. The morphology information was imported from MAPCUBE. This building dataset was commercially provided by the CAD Centre Corporation in Japan. This original building data format was a two-dimensional array of building heights with a horizontal resolution of 1m. Tus this dataset describes the buildings' distribution in vivid detail for Tokyo and Nagoya district and reveals the urban geometry in Japanese district. The author simplified and optimized mask method used in PALM reduced a three-dimensional obstacle into two-dimensional topography, improving the performance and minimizing the computational load.

At last, LES-urban provides sufficient support for canopy momentum transform simulation. As is introduced in chapter2, CM describes the flux transform process happening in the urban canopy area. Compared to laminar flow contribution, the urban canopy boundary layer is usually well developed and the turbulent flow plays a pivotal role. Kanda simulated the neutrally-stratified atmospheric boundary layer by initially setting a uniform streamwise velocity (u) of 3 m s^{-1} with zero surface heat flux. As the momentum flux transfer process is fundamental for CM and the heat and potential temperature diffusivity is adjusted according to flow situation, Kanda's simulation was conducted with zero surface heat flux and thus minimized the influence of secondary factors on the flux exchange process. Whereas the top boundary was slip, it is concluded from the experiment that the fluctuation derived from

periodical boundary could hardly contribute to the momentum transfer efficiency at a certain altitude.

All in all, the aerodynamic factors collected from the fully-developed quasi-steady turbulent state are capable of providing sufficient support for my research.

CHAPTER 4 Mathematical Methodology

4.1 The feature of boundary layer

According to Kanda et al. (2013; KA13 hereinafter) [26], although atmospheric thermal condition variety goes on and on, the momentum diffusivity coefficient tends to be a constant at every certain elevation. Therefore, it will do no harm to imagine the feature of a fully-developed boundary to acquire a general idea of the canopy boundary. Primarily, I create a common scenario where wind flow with the uniform streamwise velocity at all altitudes goes through a rough plane area.

For the case that the atmosphere flow is regarded as ideal flow, then the rough plane will not have any influence on the flow and the atmospheric layers are with the same velocity at every altitude.

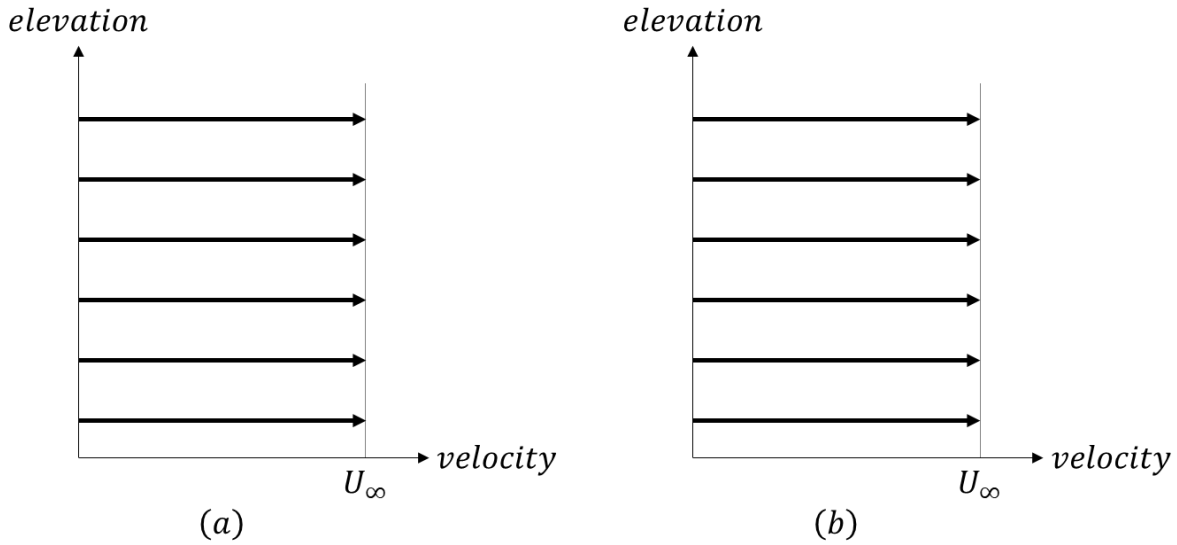

Figure 4.1 Ideal flow's profile in boundary layer (a)Initial condition (b)Well developed boundary

However, the atmospheric flow is not ideal in practice, which means that it is always with viscosity. When the fluid goes through the ground surface, the fluid practices adhere at the ground surface and thus the atmospheric layer in the bouton of boundary share the same speed with the plane. That's why it is suggested that the velocity of the atmospheric layer near the ground is zero. Due to the influence of viscosity, the internal friction resistance between fluid practices, the lower atmospheric layers are moderated by the atmospheric layer

and thus great velocity gradient appears near the ground. Hence, the area with velocity gradient is the so-called boundary.

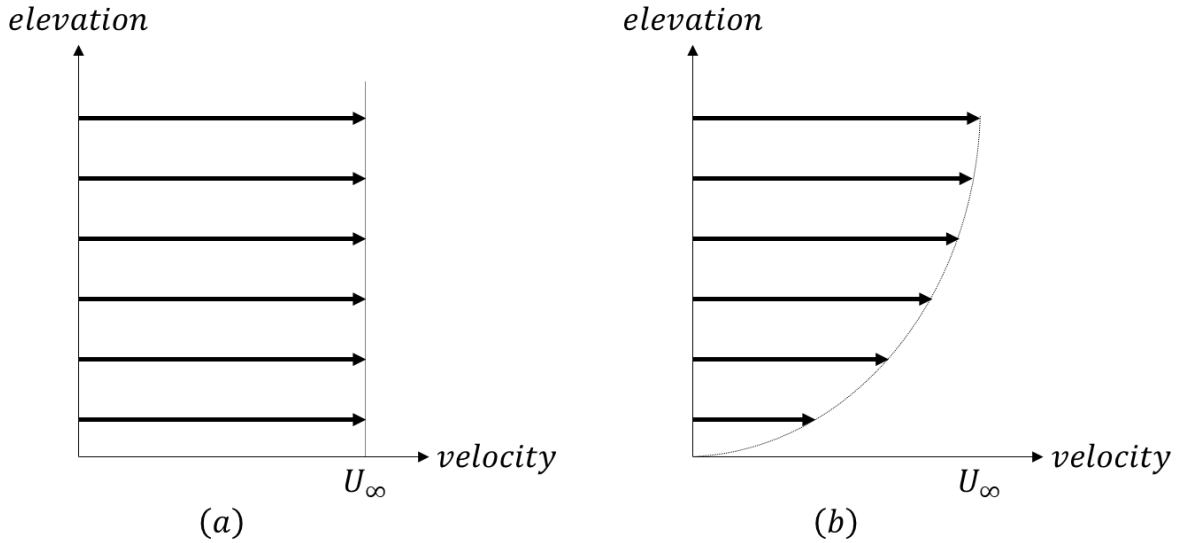

Figure 4.2 Real flow's profile in boundary layer (a)Initial condition (b)Well developed boundary

Distinguished with canopy elements, the flux exchange process is usually hampered by obstacles. Since the basic attributions of canopy elements various significantly as elevation changes within a particular canopy, it is nearly impossible to give a solution to the differential equation of the turbulent boundary layer. Researching canopy's boundary can help us get to know the flux exchange process under the influence of canopy elements and put forward an approximate solution to fit the result based on LES, to the greatest extent possible.

4.2 Canopy Height Approximate

4.2.1 Nominal Thickness

In the last chapter, I introduced that the most fundamental feature of canopy boundary momentum transformation is wind speed gradient. Due to the viscosity, atmospheric layers are accelerated by upper atmospheric layers with higher kinetic energy and moderated by the fluid beneath.

$$\tau = \mu \frac{du}{dz},$$

τ: General eddy viscosity
z: Elevation
μ: Eddy viscosity coefficient
u: General horizontal wind speed

Thus, it is natural that the canopy height is to be set at which altitude the wind speed is very close to the free wind ($0.99u_{\infty}$) and the general eddy viscosity produced by wind gradient is infinitely small. This boundary height δ is universally accepted as nominal thickness. Beyond this district, the drag effect derived from surface ground and canopy elements is negligible thus it is not wise to have data observed at this altitude to be involved in our simulation.

4.2.2 Competitive Candidate for Boundary height

Kondo (2015) [13] obtained a simple relationship between elevation in the urban canopy and mixing length from the LES database. Figure 4.3 is a scatter plot of the relationship between mixing length (L) derived from the LES database and elevation.

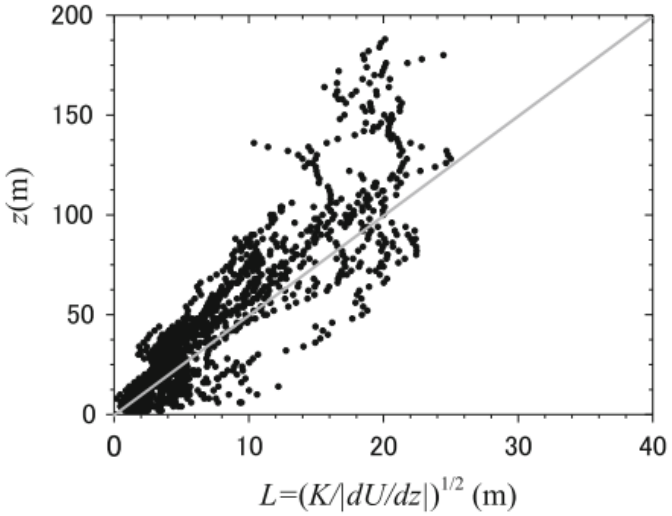

Figure 4.3 Scatter plot of mixing length L and elevation [13]

Here, mixing length was calculated using (Monin and Yaglom 1971) [12],

$$K_m = \frac{-\langle u'w' \rangle}{\left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|} = l_m^{\,2} \left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|,$$

$$l_m = \frac{\sqrt{-\langle u'w' \rangle}}{\left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|},$$

where $\langle\ \rangle$ signifies the horizontal average of each LES domain at the same level.

As the physical meaning of mixing length is describing the intrinsic relationship between time-averaged wind speed gradient and wind's fluctuate component, it is strongly implied it

is the profile of vertical momentum flux $-\langle u'w' \rangle$ rather than wind speed gradient that is the competitive candidate to simulate mixing length. LES-urban is based on MAPCUBE and the building density distribution is involved in the result. The simulation result is compared with experimental data for the same geometries from other pieces of literature [36][37] and the performance for various types of urban surfaces is impressive.

To have a general idea of aerodynamic profile in Japanese urban area, the typical specific urban district is selected from LES-urban: skyscrapers ID97, business district ID96, and residential area ID76, each of which corresponds to the area in Fig. 3.1.

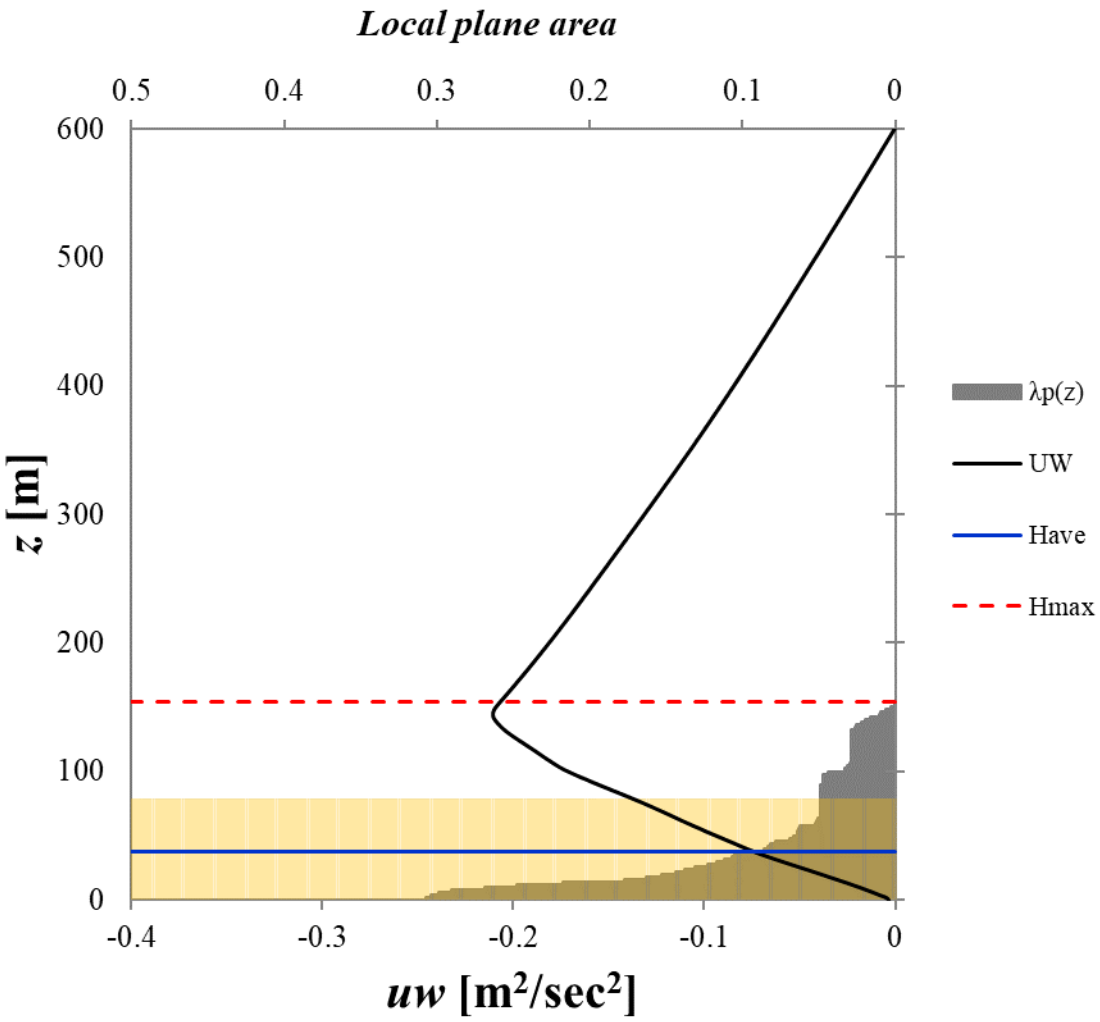

Figure 4.4(a) Profile of horizontally-averaged momentum flux for Skyscrapers ID97 [26]

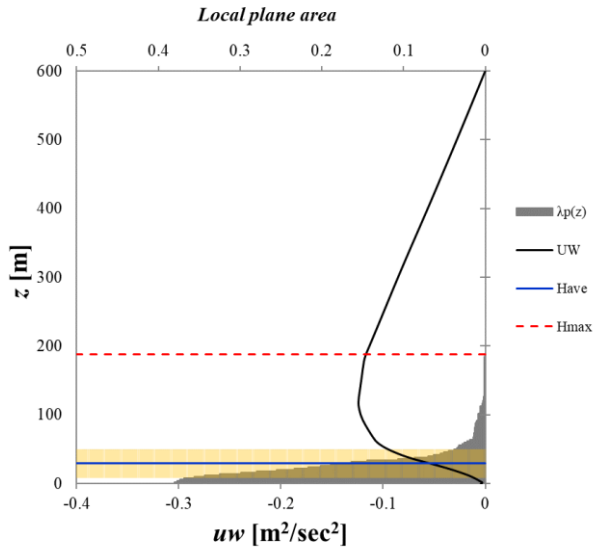

Figure 4.4(b) Profile of horizontally-averaged momentum flux for business district ID96 [26]

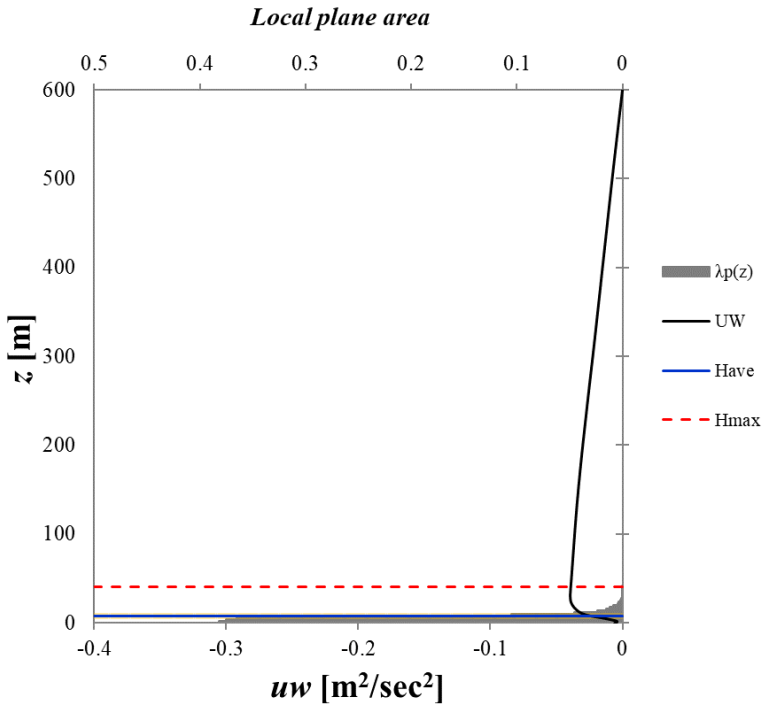

Figure 4.4(c) Profile of horizontally-averaged momentum flux for residential area ID76 [26]

The densely shaded area (grey) is the local plane area density λ_p for each 2-m layer, the lightly shaded area (yellow) is the standard deviation of building height σ_H around the average building height H_{ave} (horizontal line), and the maximum building height H_{max} is shown by the dotted line.

The business district ID96 is distinguished with maximum building height. However, as the ID97 is with higher local plane area density in the upper area and higher building height deviation, the cluster of sky scrapers actually have a higher canopy element's density in the upper area, which means that a large amount of kinetic energy is generated there and the vertical momentum exchange process is enhanced. As a result, H_{uw} is close to H_{max}.

Although the business district ID96 has several scattered tall buildings and thus a large H_{max}, the majority of local buildings are of medium height. That is why the building height deviation is relatively small comparing to ID97. Consequently, in the business district, the momentum flux peak was lower than that of the skyscrapers, and H_{uw} existed around the middle of H_{max} and $H_{ave} + \sigma_H$.

The low-storey residential area has homogeneous building density distribution. Thus, it is with the smallest H_{max}, σ_H, and H_{ave}. As a result, the vertical momentum flux is not accelerated that much. As a result, the H_{uw} is pretty small there.

Taking these results into consideration, we can draw the conclusion safely that H_{uw} existed between H_{max} and $H_{ave} + \sigma_H$. Also, these three length scales are sufficient parameters to describe vertical momentum flux profile and thus capable of simulate H_{uw}.

On the other hand, the negative peak of vertical momentum flux is a manifestation of the momentum flux efficiency. Beyond this height, the momentum exchange process is relatively moderated in accordance with the absence of kinetic energy generated by high-rise buildings. That makes H_{uw} a very canopy boundary height candidate.

According to the horizontal flow conservation equation, the mass of fluid that goes through the simulation unit per unit time should be a constant. Now that the atmospheric flow in the lower area is moderated, it is not strange to see the air flow is accelerated in the upper area to compensate for the momentum loss in the lower area. Consequently, there is a certain wind speed gradient existing in this area.

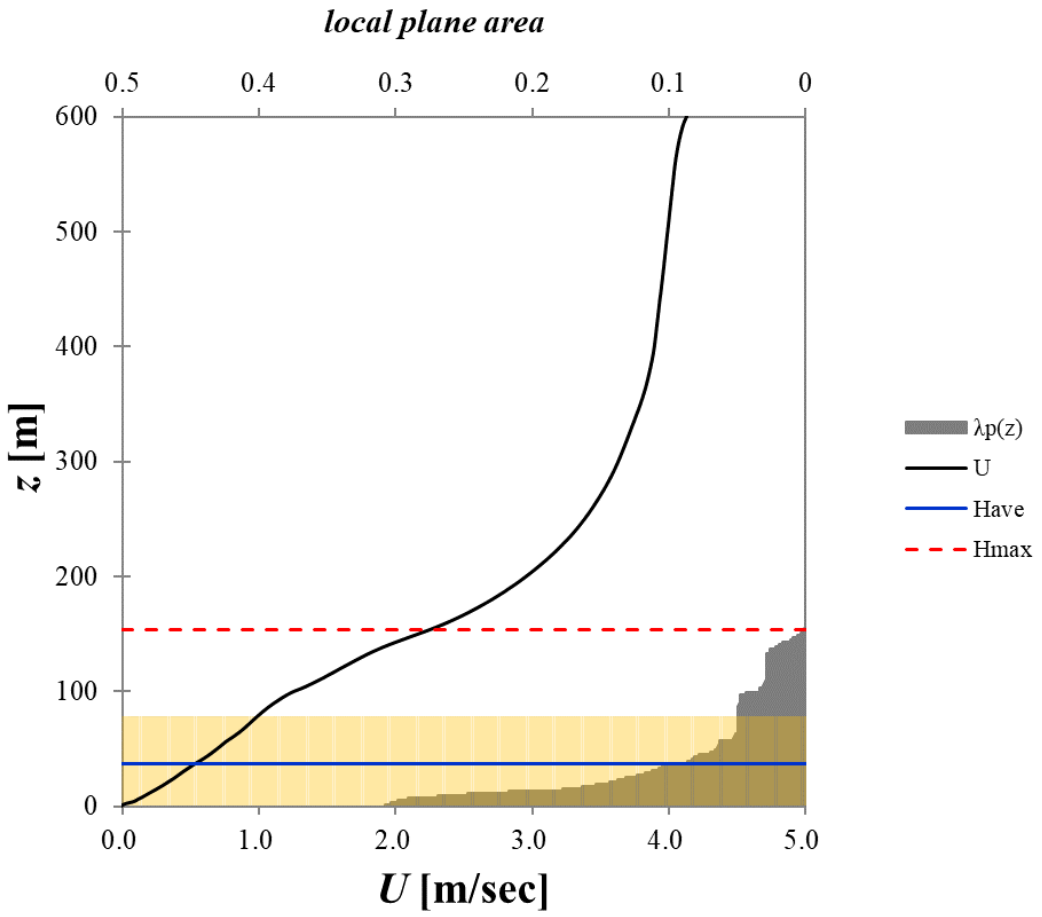

Figure 4.5 (a) Profile of horizontally-averaged wind speed for Skyscrapers ID97 [26]

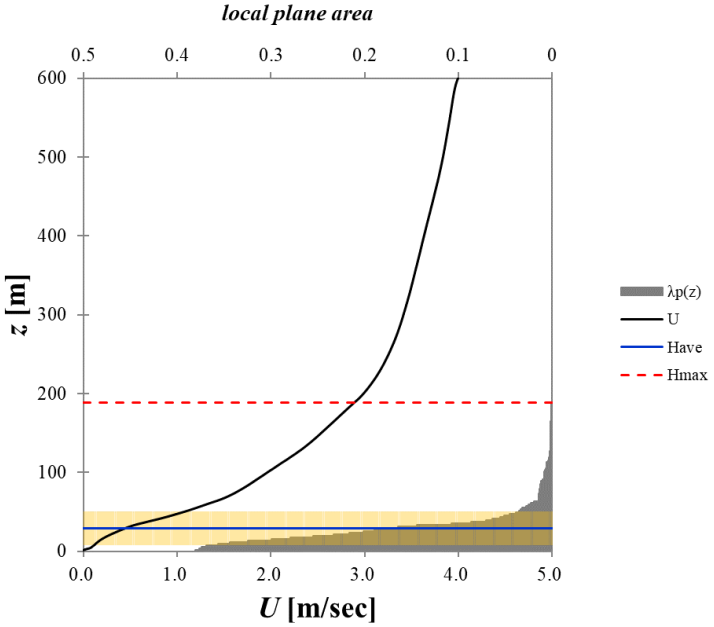

Figure 4.5(b) Profile of horizontally-averaged wind speed for business district ID96 [26]

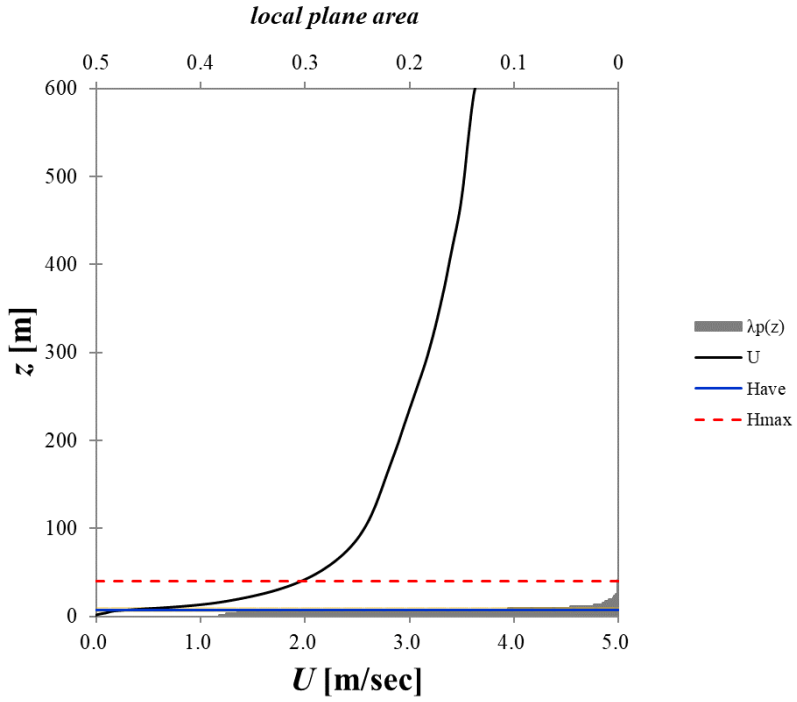

Figure 4.4(c) Profile of horizontally-averaged wind speed for residential area ID76 [26]

4.3 Normalized Wind Speed in Canopy Area

4.3.1 Normalized Wind Speed for One-Dimension Canopy

Assuming that the wind pressure is hydrodynamic pressure and the Navier-Stokes equations is given as: [41]

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\gamma\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial z^2}\right),$$
$$\frac{\partial w}{\partial t}+u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}=-\frac{1}{\rho}\frac{\partial p}{\partial z}+\gamma\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial z^2}\right),$$
$$\frac{\partial u}{\partial x}+\frac{\partial u}{\partial z}=0,$$

Where γ is kinematic viscosity.

These equations have been widely accepted to be a good solution to the two-dimensional plane boundary layer. It takes the pressure gradient force in horizontal scale into consideration and thus makes it a competitive method to analyze the influence of obstacles in detail.

Next, I normalized the parameters and consequently find some factors to be infinitesimal of higher order, which means that they are negligible compared to primary factors. Then,

$$\frac{\partial u}{\partial t}+u\frac{\partial u}{\partial x}+w\frac{\partial u}{\partial z}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+\gamma\frac{\partial^2 u}{\partial z^2},$$
$$\frac{\partial p}{\partial z}=0,$$
$$\frac{\partial u}{\partial x}+\frac{\partial u}{\partial z}=0,$$

As this study focuses on the momentum flux exchange efficiency at a certain altitude for a particular canopy district under stable atmospheric conditions, it will do no harm to make a few assumptions.

1. The velocity of coming flow is a constant for atmospheric layers with various elevation

$$U_\infty=U,$$

2. The boundary layer is well developed. It implies the wind speed is also a constant at every elevation after a period of time. As a result, the turbulent flow is enhanced and thus capable of describing the flux exchange process in a canopy in practice

$$\frac{\partial u}{\partial t} = 0,$$
$$\frac{\partial w}{\partial t} = 0,$$

3. Resulting from canopy elements' drag effect, pressure-gradient force is generated as the manifestation of drag effect. Although atmospheric layers are moderated by pressure-gradient force, the air flow is with a constant rate because the coming flow compensates the momentum loss, which implies

$$-\frac{1}{\rho}\frac{\partial p}{\partial x} = U\frac{dU}{dx},$$

In that way, the equations to describe the boundary layer in stable conditions are given as

$$u\frac{\partial u}{\partial x} + w\frac{\partial u}{\partial z} = U\frac{dU}{dx} + \gamma\frac{\partial^2 u}{\partial z^2},$$

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial z} = 0,$$

However, CM is a one-dimension solution and takes the canopy elements as a homogeneous porous medium. Thus, much importance is attached to the density distribution of canopy elements.

To our relief, K.Pohlhause developed an approximate Solution to simulate the wind speed profile in the canopy with a two-dimensions obstacle of arbitrary shape.

4. Firstly, the vertical simulation scale is strictly defined. The elevation length scale η is normalized by canopy height δ as

$$\eta = \frac{z}{\delta},$$

5. Secondly, the wind speed profile must satisfy boundary conditions:

a) Where there is near the ground,

$$z = 0, \eta = \frac{z}{\delta} = 0,$$

Since all the surfaces except for the boundary are non-slip:

$$u_0 = 0,$$

Thus, the wind speed gradient peaks there, and the maximum Reynolds stress is given as

$$\eta = 0, \frac{\partial u}{\partial z} = \frac{\tau_0}{\mu},$$

pressure-gradient force is given as,

$$\gamma \frac{\partial^2 u}{\partial z^2} = \frac{1}{\rho} \frac{\partial p}{\partial x} = -U \frac{dU}{dx},$$

b) Where there is the boundary

$$z = \delta, \eta = \frac{z}{\delta} = 1,$$

Since the boundary is slip surface:

$$u_\delta = U,$$

Thus, the wind speed gradient is pretty small there and we could simply consider that

$$\eta = 1, \frac{\partial u}{\partial z} = 0, \frac{\partial^2 u}{\partial z^2} = 0,$$

6. There is no doubt that more model complexity results in a better fit. However, infinite model complexity leads to over-fitting. In that way, it is assumed that quartic polynomial

is complex enough to describe wind speed profile. Here, I have normalized wind speed $f(\eta)$ as [41]

$$\frac{u}{U} = f(\eta) = a\eta + b\eta^2 + c\eta^3 + d\eta^4,$$

7. Together with boundary conditions, I can have this equation solved as

$$\Lambda = \frac{\delta^2}{\gamma}\frac{dU}{dx},$$

And the constant coefficients are solved as

$$a = 2 + \frac{\Lambda}{6}, b = -\frac{\Lambda}{2}, c = -2 + \frac{\Lambda}{2}, d = 1 - \frac{\Lambda}{6},$$

Thus,

$$\frac{u}{U} = f(\eta) = (2\eta - 2\eta^3 + \eta^4) + \frac{\Lambda}{6}(\eta - 3\eta^2 + 3\eta^3 - \eta^4) = F(\eta) + \Lambda G(\eta),$$
$$F(\eta) = 2\eta - 2\eta^3 + \eta^4 = 1 - (1-\eta)^3(1+\eta),$$
$$G(\eta) = \frac{1}{6}(\eta - 3\eta^2 + 3\eta^3 - \eta^4) = \frac{1}{6}\eta(1-\eta)^3,$$

Λ is called *shape factor*. It primarily depends on the shape of canopy elements, which contributes mostly to $\frac{dU}{dx}$.

Derived by Bernoulli's equation, Λ can be described as [41]

$$\Lambda = \frac{\delta^2}{\gamma}\frac{dU}{dx} = -\frac{dp}{dx}\Big/\mu\frac{U}{\delta^2},$$

That means Λ could be regarded by the ratio of pressure-gradient force to viscous force. In that way, the positive and negative of Λ are capable of describing both drag effect and accelerating effects.

The simulation districts may have different vertical building density distributions and various canopy heights thus it is usually impossible to find their features in common. Yet, researching canopies sharing the same Λ can throw light upon their intrinsic similarity.

$f(\eta)$, it should be noted, is the equation reveals the relationship between normalized wind speed and normalized vertical length scale. That makes it a perfect solution to research the regularity of the flux exchange process happening in CM, a one-dimension model.

4.3.2 The Basic Properties of Λ

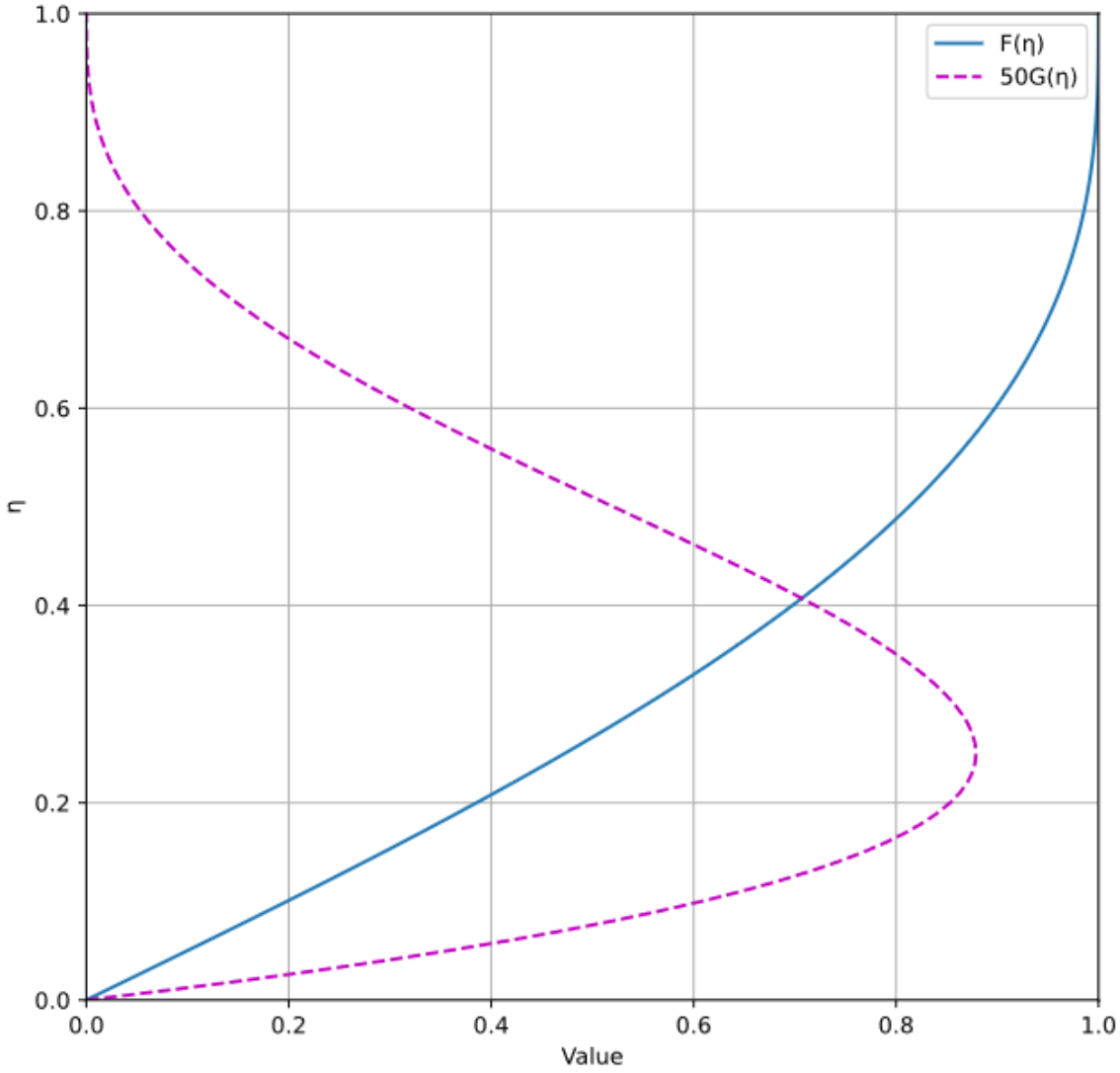

Figure 4.5 The function relationship between F, Gand η

Λ is given as

$$\Lambda = \frac{\delta^2}{\gamma}\frac{dU}{dx} = -\frac{dp}{dx}\Big/\mu\frac{U}{\delta^2}$$

Under the condition that Λ is zero, the contribution of $G(\eta)$ to $f(\eta)$ vanishes. It is suggested that $\frac{dU}{dx} = \frac{dp}{dx} = 0$ and thus the influence of pressure gradient force disappears in the circumstances. In other words, $F(\eta)$ left there is to describe a canopy without any canopy elements, also flat-plate boundary layer.

$G(\eta)$ is a positive parameter in the defined area. The value of $G(\eta)$ cuts no figure where it is near the ground and close to canopy boundary but peaks at $\eta = 0.25$. Compared to $F(\eta)$, its contribution to wind speed profile is relatively small in numerical terms. However, combining Λ , $G(\eta)$ is capable of describing the accelerating effect resulting from atmospheric with rich kinetic energy and drag effect moderating atmospheric layers. To be specific, positive Λ means accelerating effect in the whole vertical scale (Figure 4.6), and negative Λ drops a hint that the momentum flux is hampered by pressure gradient force and the wind is moderated (Figure 4.7).

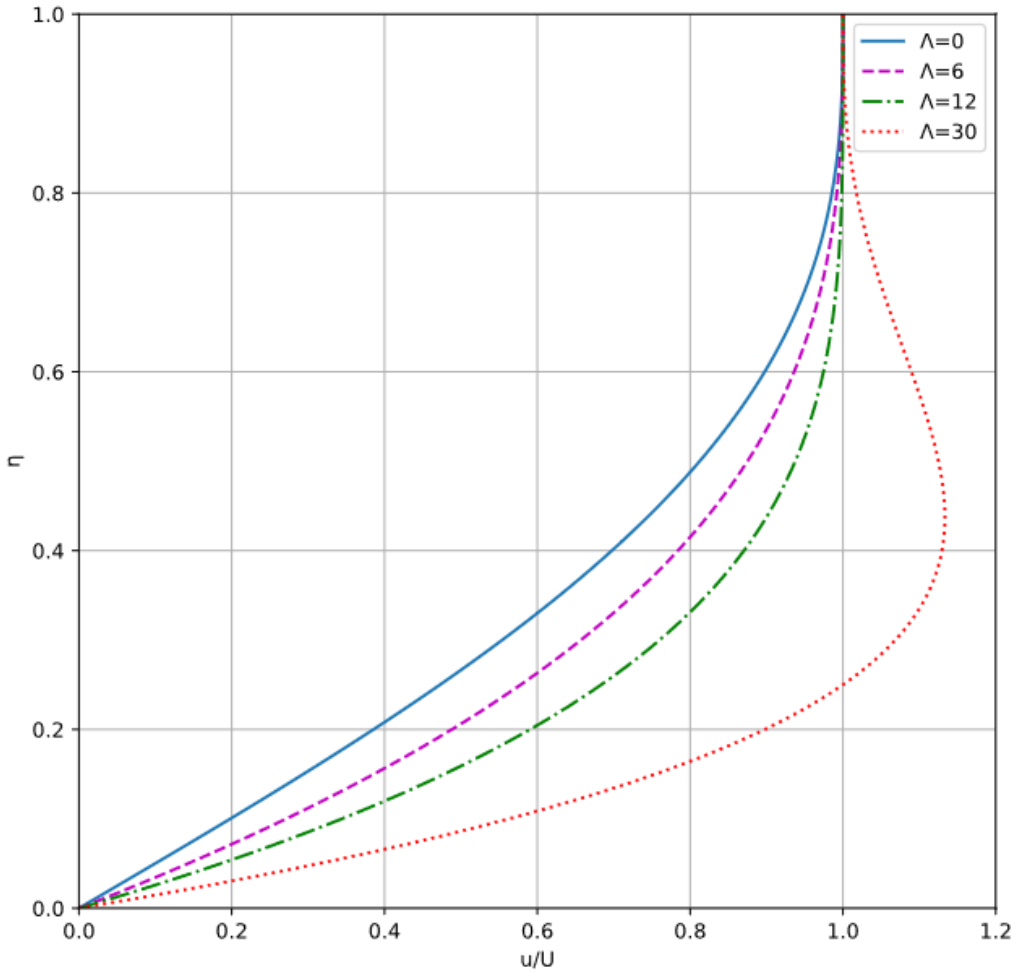

Figure 4.6 normalized wind speed profile with positive Λ

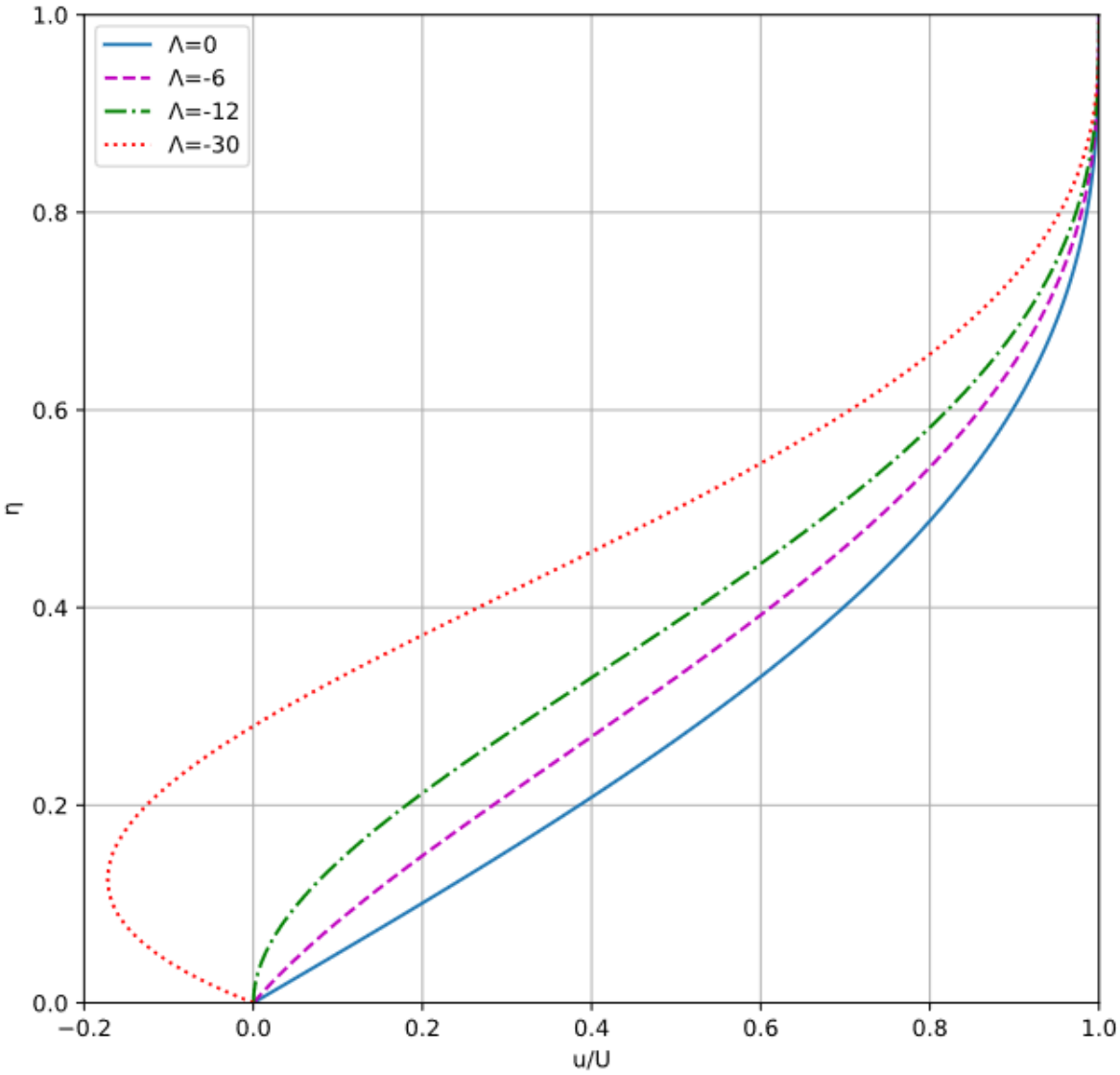

Figure 4.7 normalized wind speed profile with negative Λ

Notably, the maximum wind speed $f(\eta)$ are likely to exceed 1 when the numeric value of Λ surpasses 12, which should hardly happen due to the canopy's drag effect. Also, it is worth pointing out that constant parameter $a = 2 + \frac{\Lambda}{6}$ and vertical velocity gradient $\left(\frac{\partial u}{\partial y}\right)_{y=0}$ will become negative under the condition that $\Lambda < -12$. Under this circumstance, the velocity in the lower area will be in the opposite direction of the coming flow.

4.4 Mixing length Approximation

From Blackadar (1962) [24], I can acquire the universally acknowledged equation to simulate mixing length beyond canopy as,

$$l_m = kz\left(1 + \frac{kz}{L_B}\right)^{-1},$$

Kondo (2015) [13] demonstrated that 100m is a rather competitive candidate for L_B. As H_{uw} is utilized as the canopy boundary height, I can calculate mixing length at boundary height as,

$$l_{max} = kH_{uw}\left(1 + \frac{kH_{uw}}{L_B}\right)^{-1},$$

This study investigates the relationship between normalized mixing length, l_m, and shape factor Λ inside the canopy. Although the CM is a one-dimensional urban canopy model considering a simplified structure in a target area, I started from a simpler basic equation in which a horizontally homogeneous and stationary field is assumed (permeability is ignored). Under this condition, Kondo and Akashi (1976) [11] assumed

$$\frac{1}{\rho}\frac{d\tau}{dz} = \frac{d}{dz}\left(K\frac{du}{dz}\right) = \frac{d}{dz}\left(l_m^2\left(\frac{du}{dz}\right)^2\right) = \frac{1}{2}cau^2,$$

where τ is the vertical flux of momentum, ρ is the air density, c is the drag coefficient for the canopy element (assumed constant), a is leaf area density for the canopy. In the present case, which generates the drag force in a unit volume of the canopy, u is the longitudinal wind velocity component. This equation can be non-dimensionalized with scaling parameter, canopy height, for length scale and $F(\eta) + \Lambda G(\eta)$ for velocity, whereby

$$\frac{d}{d\eta}\left(l_{m+}^2\left(\frac{d\big(F(\eta) + \Lambda G(\eta)\big)}{d\eta}\right)^2\right) = \frac{1}{2}ca_+\big(F(\eta) + \Lambda G(\eta)\big)^2,$$

where subscripts "+" indicate non-dimensional parameters.

Because the parameter $\frac{1}{2}ca_+$, describing the flux exchange efficiency in the canopy, primarily depends on the shape of canopy elements. Thus, there is an excellent reason to suggest that it can be described by function $S_{(\Lambda,\eta)}$.

Then, I can have the equation to describe momentum transfer efficiency at a certain height η as

$$l_{m+}^2\left(\frac{d\big(F(\eta)+\Lambda G(\eta)\big)}{d\eta}\right)^2 = S_{(\Lambda,\eta)}\int_0^{\eta}\big(F(\eta)+\Lambda G(\eta)\big)^2 d\eta,$$

$$F(\eta) = 2\eta - 2\eta^3 + \eta^4 = 1-(1-\eta)^3(1+\eta),$$
$$G(\eta) = \frac{1}{6}(\eta - 3\eta^2 + 3\eta^3 - \eta^4) = \frac{1}{6}\eta(1-\eta)^3,$$

Where it follows that,

$$\frac{d\big(F(\eta)+\Lambda G(\eta)\big)}{d\eta} = 2 - 6\eta^2 + 4\eta^3 + \frac{\Lambda}{6}(1 - 6\eta + 9\eta^2 - 4\eta^3),$$

$$\begin{aligned}\big(F(\eta)+\Lambda G(\eta)\big)^2 &= \left(\frac{1}{36}\Lambda^2 + \frac{2}{3}\Lambda + 4\right)\eta^2 + \left(-\frac{1}{6}\Lambda^2 - 2\Lambda\right)\eta^3 + \left(\frac{5}{12}\Lambda^2 + \frac{4}{3}\Lambda - 8\right)\eta^4 \\ &+ \left(-\frac{5}{9}\Lambda^2 + \frac{5}{3}\Lambda + 4\right)\eta^5 + \left(\frac{5}{12}\Lambda^2 - 3\Lambda + 4\right)\eta^6 + \left(-\frac{1}{6}\Lambda^2 + \frac{5}{3}\Lambda - 4\right)\eta^7 \\ &+ \left(1 - \frac{1}{3}\Lambda + \frac{1}{36}\Lambda^2\right)\eta^8,\end{aligned}$$

Thus,

$$\int_0^{\eta}\left(F(\eta)+\Lambda G(\eta)\right)^2 d\eta$$
$$=\frac{1}{3}\left(\frac{1}{36}\Lambda^2+\frac{2}{3}\Lambda+4\right)\eta^3+\frac{1}{4}\left(-\frac{1}{6}\Lambda^2-2\Lambda\right)\eta^4+\frac{1}{5}\left(\frac{5}{12}\Lambda^2+\frac{4}{3}\Lambda-8\right)\eta^5$$
$$+\frac{1}{6}\left(-\frac{5}{9}\Lambda^2+\frac{5}{3}\Lambda+4\right)\eta^6+\frac{1}{7}\left(\frac{5}{12}\Lambda^2-3\Lambda+4\right)\eta^7$$
$$+\frac{1}{8}\left(-\frac{1}{6}\Lambda^2+\frac{5}{3}\Lambda-4\right)\eta^8+\frac{1}{9}\left(1-\frac{1}{3}\Lambda+\frac{1}{36}\Lambda^2\right)\eta^9+C$$
$$=\frac{1}{45360}\eta^3[140(\Lambda^2-12\Lambda+36)\ \eta^6+945(-\Lambda^2+10\Lambda-24)\ \eta^5$$
$$+540\ (5\Lambda^2-36\Lambda+48)\ \eta^4+840(-5\Lambda^2+15\Lambda+36)\eta^3$$
$$+756(5\Lambda^2+16\Lambda-96)\Lambda^2-1890(\Lambda^2+12\Lambda)\eta$$
$$+(420\Lambda^2+10080\Lambda+60480)],$$

$$\left(\frac{d\left(F(\eta)+\Lambda G(\eta)\right)}{d\eta}\right)^2=\frac{[(24\eta+12)-(4\eta-1)\Lambda]^2(\eta-1)^4}{36},$$

Therefore, I can find a tentative solution to calculate l_{m+}.

$$l_{m+}=\Big(\frac{1}{1260}[140(\Lambda^2-12\Lambda+36)\ \eta^6+945(-\Lambda^2+10\Lambda-24)\ \eta^5$$
$$+540\ (5\Lambda^2-36\Lambda+48)\ \eta^4+840(-5\Lambda^2+15\Lambda+36)\eta^3$$
$$+756(5\Lambda^2+16\Lambda-96)\eta^2-1890(\Lambda^2+12\Lambda)\eta$$
$$+(420\Lambda^2+10080\Lambda+60480)]\ \eta^3\ [(24\eta+12)-(4\eta-1)\Lambda]^{-2}(\eta$$
$$-1)^{-4}\ S_{(\Lambda,\eta)}\Big)^{0.5},$$

Notably, the equation to calculate l_{m+} should be valid in the defined domain.

A. Where there is the ground

$$\eta=0; l_{m+}=0,$$

$$\frac{d\left(F(\eta)+\Lambda G(\eta)\right)}{d\eta}=2+\frac{\Lambda}{6},$$
$$\left(F(\eta)+\Lambda G(\eta)\right)^2=0,$$

Thus,

$$C = 0,$$

At this altitude, it is obvious that $S_{(\Lambda,\eta)}$ and $\left(\frac{d(F(\eta)+\Lambda G(\eta))}{d\eta}\right)^2$ are not identically zero. I may as well calculate limits to throw light upon the component of l_{m+}

$$\frac{S_{(\Lambda,\eta)}}{\left(\frac{d(F(\eta)+\Lambda G(\eta))}{d\eta}\right)^2} = \lim_{\eta\to 0}\frac{l_{m+}^2}{\int_0^\eta (F(\eta)+\Lambda G(\eta))^2 d\eta} = \lim_{\eta\to 0}\frac{2l_m l_m'}{(F(\eta)+\Lambda G(\eta))^2},$$

As lowest order vertical length scale term of $(F(\eta)+\Lambda G(\eta))^2$ is $\left(\frac{1}{36}\Lambda^2+\frac{2}{3}\Lambda+4\right)\eta^2$, the lowest order term of l_m must be $g(\Lambda)\eta^{\frac{3}{2}}$, where $g(\Lambda)$ is a function independent from η.

B. Where there is in the middle of the canopy, the equation describing l_{m+} should always make sense, which means that $S_{(\Lambda)}$ must have $[(24\eta+12)-(4\eta-1)\Lambda]^2(\eta-1)^4$ included. In that way, it will do no harm to assume that

$$S_{(\Lambda,\eta)} = [(24\eta+12)-(4\eta-1)\Lambda]^2(\eta-1)^4\, f(\Lambda,\eta),$$

Where $f(\Lambda,\eta)$ is a function that depends on Λ and η.

Consequently,

$$\begin{aligned} l_{m+} = \Big(\frac{1}{1260}[&140(\Lambda^2-12\Lambda+36)\ \eta^6+945(-\Lambda^2+10\Lambda-24)\ \eta^5\\ &+540\ (5\Lambda^2-36\Lambda+48)\,\eta^4+840(-5\Lambda^2+15\Lambda+36)\eta^3\\ &+756(5\Lambda^2+16\Lambda-96)\eta^2-1890(\Lambda^2+12\Lambda)\eta\\ &+(420\Lambda^2+10080\Lambda+60480)]\,\eta^3\ f_{(\Lambda,\eta)}\Big)^{0.5}, \end{aligned}$$

$f_{(\Lambda,\eta)}$ should be totally independent of η, otherwise, the pre-condition that lowest order term of l_m must be $\eta^{\frac{3}{2}}$ will not be satisfied. In other words, $f(\Lambda)$ is a polynomial about Λ

$$S_{(\Lambda,\eta)} = [(24\eta+12)-(4\eta-1)\Lambda]^2(\eta-1)^4\, f(\Lambda),$$

C. Where there is at boundary height,

$$\eta = 1; l_{m+} = 1,$$

Thus,

$$1 = \frac{1}{1260}[140(\Lambda^2 - 12\Lambda + 36) + 945(-\Lambda^2 + 10\Lambda - 24) + 540\,(5\Lambda^2 - 36\Lambda + 48) + 840(-5\Lambda^2 + 15\Lambda + 36) + 756(5\Lambda^2 + 16\Lambda - 96) - 1890(\Lambda^2 + 12\Lambda) + (420\Lambda^2 + 10080\Lambda + 60480)]\ f(\Lambda),$$

Therefore,

$$f(\Lambda) = \frac{1260}{5\Lambda^2 + 426\Lambda + 26424},$$

$$S_{(\Lambda,\eta)} = [(24\eta + 12) - (4\eta - 1)\Lambda]^2(\eta - 1)^4 \frac{1260}{5\Lambda^2 + 426\Lambda + 26424},$$

Derived from shape factor Λ, the parameter describing the geometric similarity between canopies with vertical various length scales, and elevation length scale η, $S_{(\Lambda,\eta)}$ is capable of describing canopy elements' drag effect in the vertical direction.

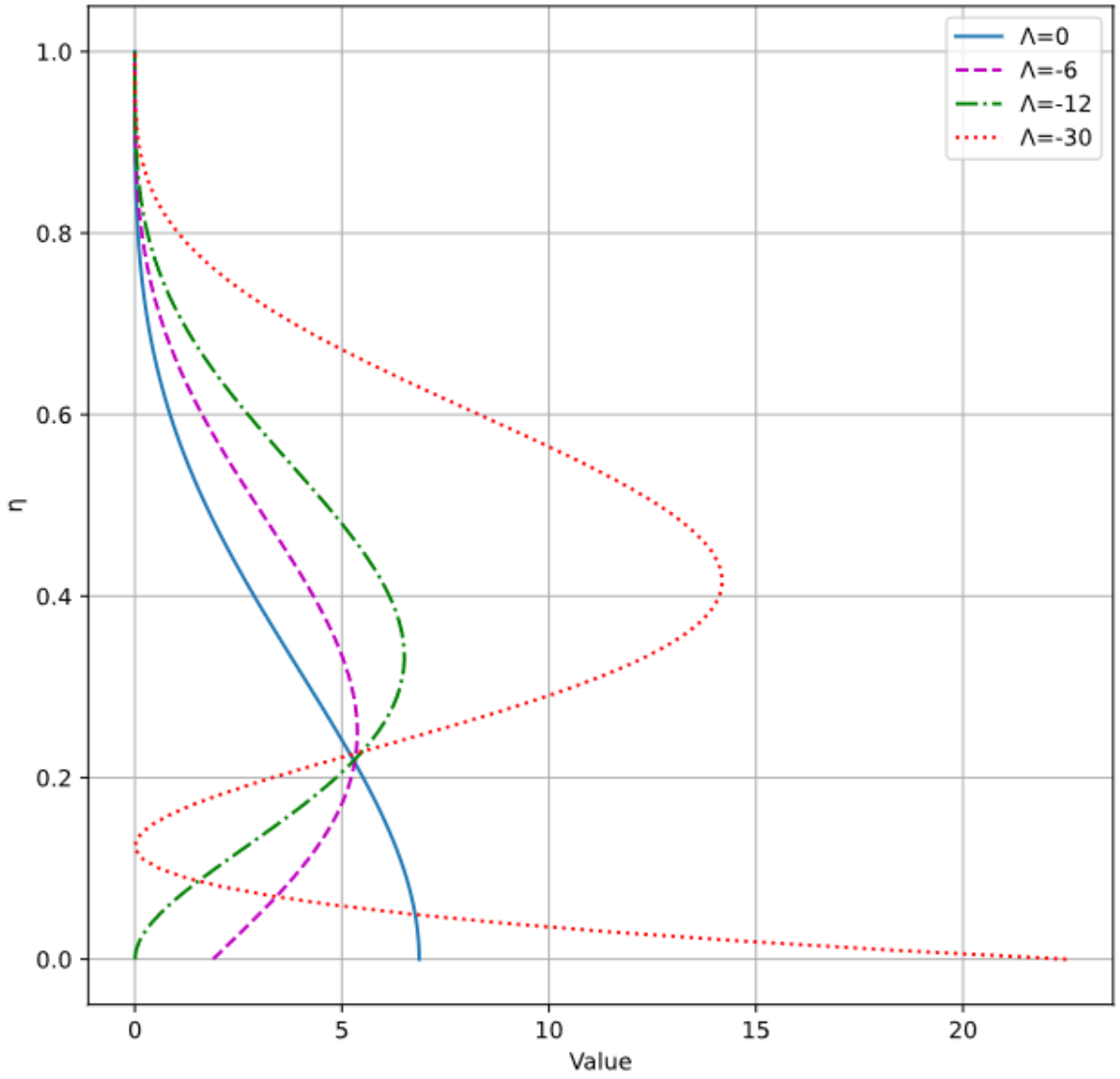

Figure 4.8 $S_{(\Lambda,\eta)}$ profile with negative Λ

Figure 4.8 shows the profile of $S_{(\Lambda,\eta)}$ with negative Λ. Under the circumstance that Λ is negative, the canopy density is pretty high and that results in drag effect described by $S_{(\Lambda,\eta)}$ in general. To be specific, the drag effect is significant near the ground under the condition that Λ is zero, which means the ground surface has a significant influence on a canopy with low density. Where η is close to 1, the elevation is very close to the canopy height, which suggests low canopy elements' density and thus the drag effect vanishes at this altitude. However, as the canopy density increases, Λ turns into a negative value, and the drag effect peaks at a relatively upper altitude.

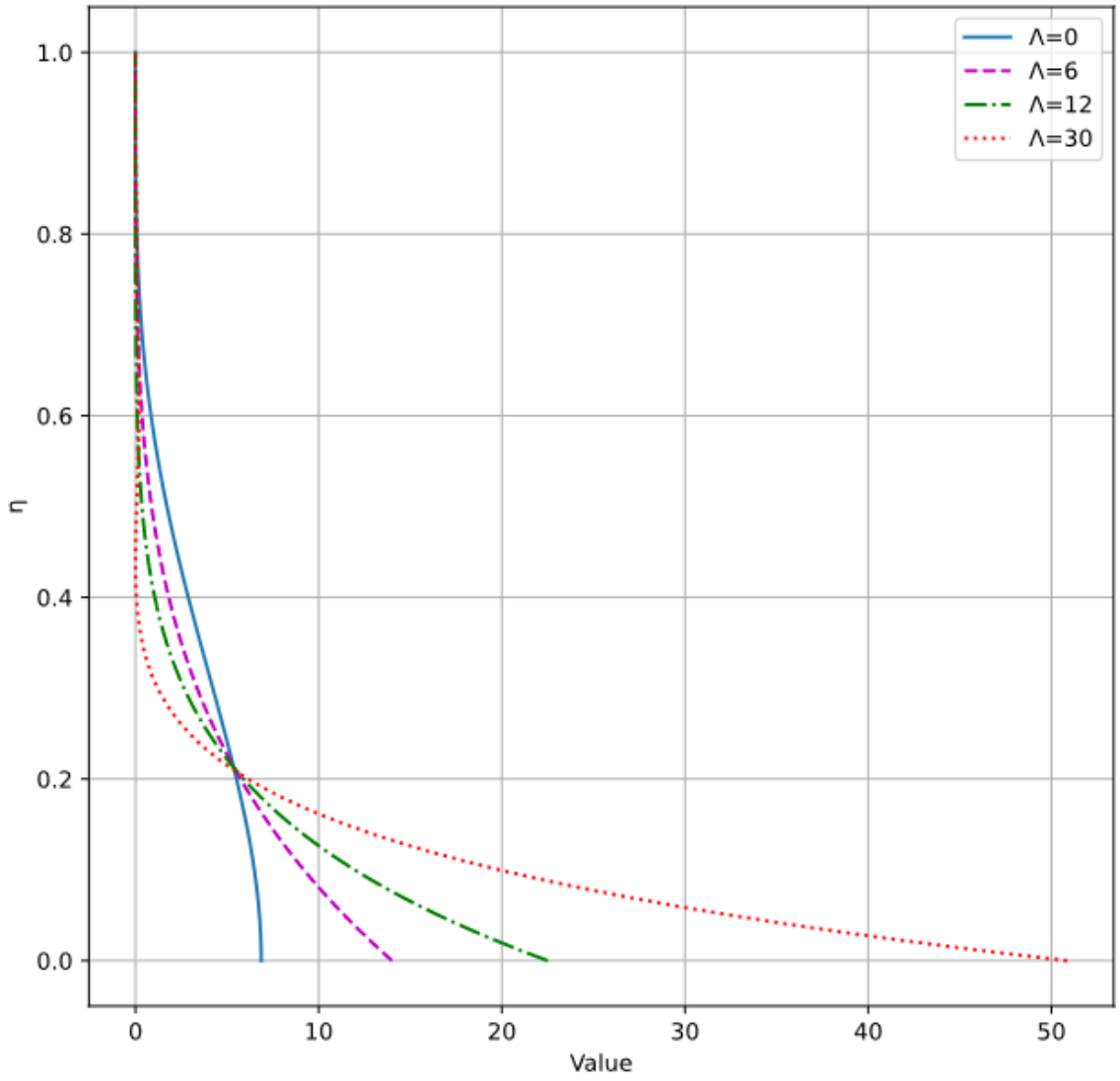

Figure 4.9 $S_{(\Lambda,\eta)}$ profile with positive Λ

Figure 4.9 shows the profile of $S_{(\Lambda,\eta)}$ with positive Λ. Where Λ is positive, the canopy elements are with low density and thus the atmospheric layers are more likely to be accelerated by upper atmospheric layers with rich kinetic energy. Under the circumstances, $S_{(\Lambda,\eta)}$ describes the accelerating effect. Notably, the accelerating effect is rather significant at the button of canopies on account of the fact that wind speed is universally lower where η is pretty small.

Having $S_{(\Lambda,\eta)}$ applied, I can have the equation to calculate l_m as,

$$l_{m+} = \Big(\frac{1}{5\Lambda^2 + 426\Lambda + 26424}[140(\Lambda^2 - 12\Lambda + 36)\ \eta^6 + 945(-\Lambda^2 + 10\Lambda - 24)\ \eta^5 + 540\ (5\Lambda^2 - 36\Lambda + 48)\ \eta^4 + 840(-5\Lambda^2 + 15\Lambda + 36)\eta^3 + 756(5\Lambda^2 + 16\Lambda - 96)\eta^2 - 1890(\Lambda^2 + 12\Lambda)\eta + (420\Lambda^2 + 10080\Lambda + 60480)]\ \eta^3\Big)^{0.5},$$

Figure 4.10 and Figure 4.11 shows the profile of l_{m+} with various Λ.

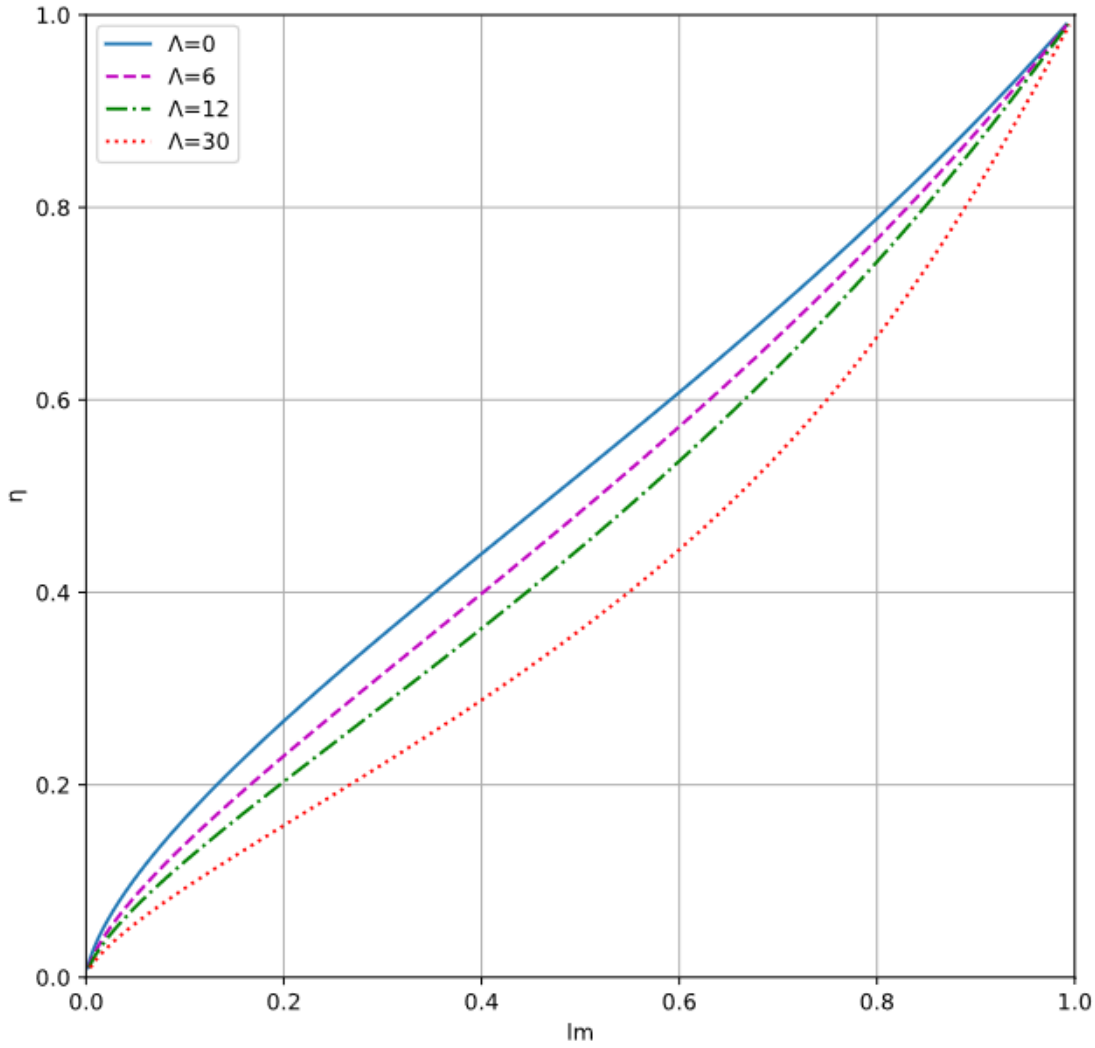

Figure 4.10 l_{m+} profile with positive Λ

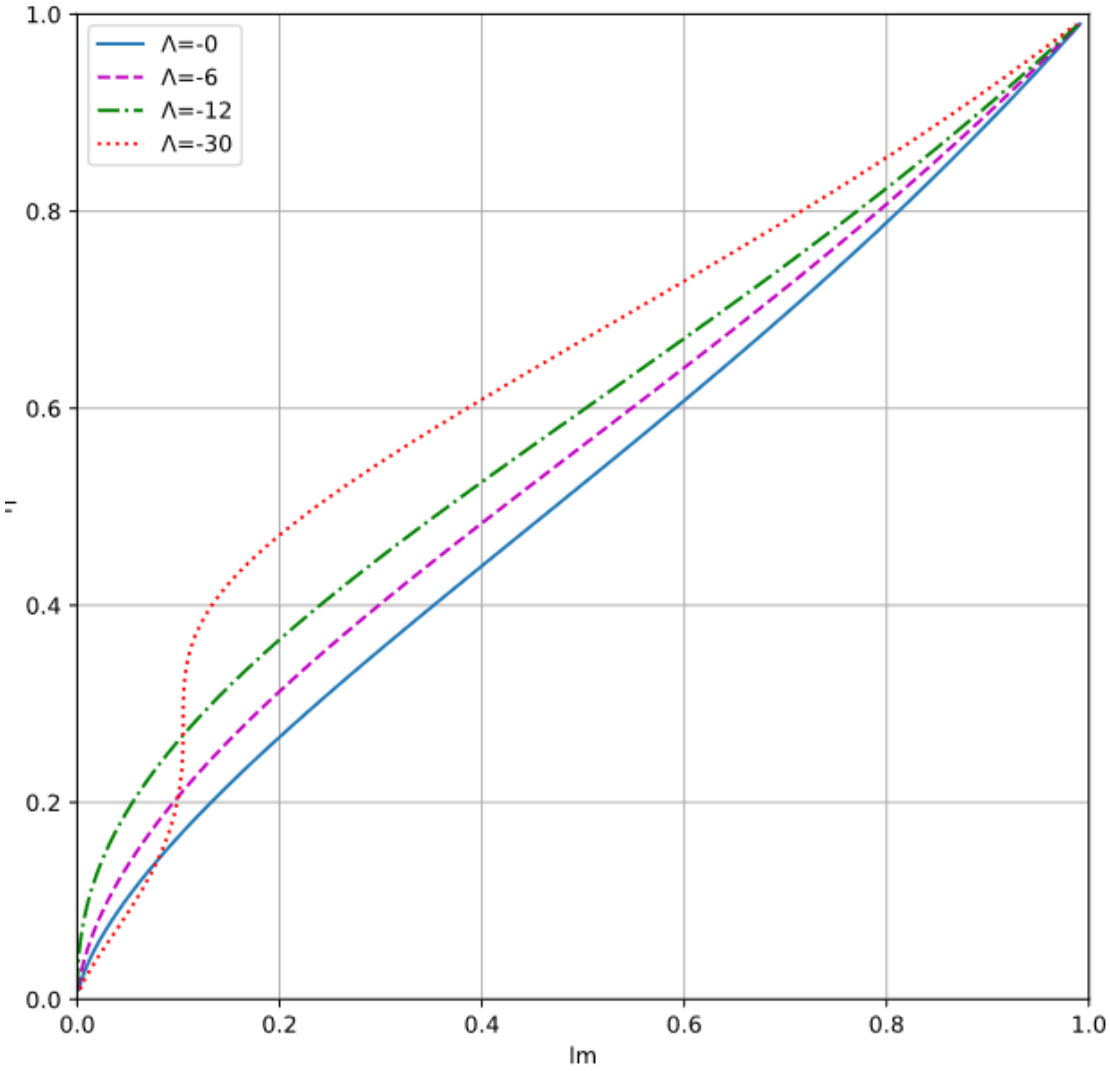

Figure 4.11 l_{m+} profile with negative Λ

4.5 Reflection

Firstly, the shape factor Λ significantly enlarged this study's scope of application. It is universally acknowledged that the influence of canopy elements on the flux exchange effect is a kind of drag effect. However, the influence of buildings on flux could be expressed by pressure-gradient force. Λ is introduced to describe the positive and negative of pressure gradient force. That makes it possible for the equations to describe not only the accelerating effects but also moderating effects in canopy scale.

$$\Lambda = \frac{\delta^2}{\gamma}\frac{dU}{dx} = -\frac{dp}{dx}\Big/\mu\frac{U}{\delta^2},$$

Secondly, the physical meaning of equation components should be emphasized.

$$\tau_t = K\frac{du}{dz} = {l_m}^2\left|\frac{du}{dz}\right|\frac{du}{dz},$$

As l_m describe the relationship between turbulent shear stress and wind speed gradient, the equation describing l_m should be totally independent of wind speed gradient and only in that way can l_m be capable of describing flux exchange efficiency under various wind gradients. After strict mathematical deduction the unique solution of $S_{(\Lambda,\eta)}$ is obtained and this solution eliminated the influence of wind speed gradient. Thus, it makes the equation of l_m a very competitive pattern to describe the momentum flux coefficient at a certain altitude with a given canopy shape factor, even the wind speed always changes and the momentum exchange process is never-ending.

Next, it is important to stress that the equation to describe l_{m+} is obtained by the exact math analysis, which makes the conclusion rather convincing. There is no denying that the traditional empirical formula [13] is capable of describing the wind speed contribution as

$$u_+ \propto \eta^p,$$

Where p is a constant. However, it doesn't really help explain the features of canopy boundary in essence: The canopy boundary is where the wind speed is so close to the geostrophic wind. Canopy's influence on the momentum flux exchange process is almost eliminated there and the velocity gradient is small enough to ignore. This study

takes boundary conditions seriously and numerical approximation was not utilized throughout the entire mathematical derivation process. As a result, the equation describing l_{m+} only have normalized length scale parameter η and shape factor Λ involved. Hence, the intrinsic relationship between atmospheric components' vertical distribution and shape factor is revealed to help explain atmospheric phenomena within the canopy.

Finally, the profile collected is very close to the previous study. [11] As is shown in figure 4.10 and figure 4.11, there is an apparent positive linear correlation between l_{m+} and vertical length scale η. Where the Λ is positive, l_{m+} become relatively larger, which means that the flux exchange process is enhanced in this situation. Where the Λ is negative, the drag effect is dominated by canopy elements' drag effect, which means relatively small l_{m+}. These qualitative analyses make this solution rather convincing.

CHAPTER.5 Data Science Methodology

5.1 Data Science to Calculate Canopy Height and Shape Factor

This study focuses on canopy elements' influence on the flux transfer process. In micro terms, the atmospheric particle rubs against non-slip surfaces. Consequently, kinetic energy is generated to promote the local momentum flux transferring process. Thus, canopy bulk geometric variables, (local average building height, local maximum building height, building height deviation) which describe the distribution of no-slip surface, are rather competitive variables to approximate canopy height H_{uw} and shape factor Λ . Unfortunately, they are observed from LES-urban, which means that analytic solution is not with easy access at this stage. It is natural to assume that there is a linear relationship between these parameters, while this study has MLP adopted to reveal the intrinsic relationship between canopy bulk geometric variables and H_{uw}, Λ and improve approximation performance.

5.1.1 Traditional Solution-- Linear Regression

There is no denying that Linear Regression is a competitive method for modeling the relationship between one or more independent variables -- canopy bulk geometric variables, and a dependent variable -- H_{uw} and Λ.

However, the traditional Linear Regression method is not always perfect. linearity implies the weaker assumption of monotonicity: that any increase in our feature must either always cause an increase in our model's output (if the corresponding weight is positive), or always cause a decrease in our model's output (if the corresponding weight is negative). there might exist a representation of our data that would take into account the relevant interactions among our features, on top of which a linear model would be suitable, we simply do not know how to calculate it by hand. With deep neural networks, we used observational data to jointly learn both a representation via hidden layers and a linear predictor that acts upon that representation.

5.1.2 Multilayer Perceptrons

We can overcome these limitations of linear models and handle a more general class of functions by incorporating one or more hidden layers. The easiest way to do this is to stack many fully-connected layers on top of each other. Each layer feeds into the layer above it, until we generate outputs. Assume that the total layer number is L. We can think of the first

$L - 1$ layers as our representation and the final layer as our linear predictor. This architecture is commonly called a multilayer perceptron, often abbreviated as MLP. Below, we depict an MLP diagrammatically.

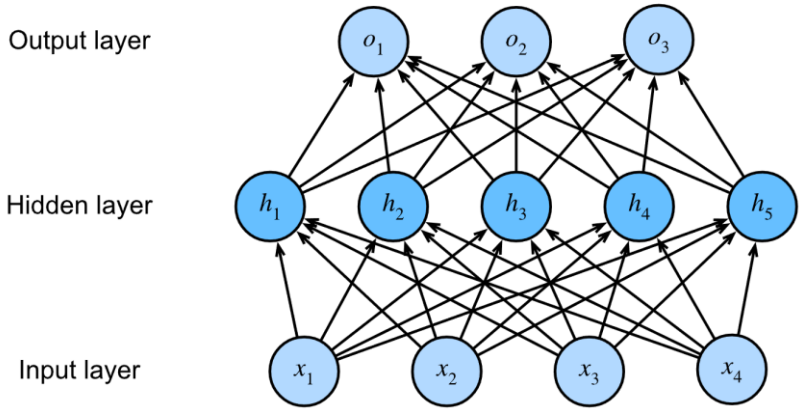

Figure 5.1 An MLP with a hidden layer of 5 hidden units [38]

This MLP has 4 inputs, 3 outputs, and its hidden layer contains 5 hidden units. Since the input layer does not involve any calculations, producing outputs with this network requires implementing the computations for both the hidden and output layers; thus, the number of layers in this MLP is 2. Note that these layers are both fully connected. Every input influences every neuron in the hidden layer, and each of these in turn influences every neuron in the output layer.

5.1.3 Nonlinear Layers to Activate Neurons

As it was shown in Figure 5.1, there is a multilayer perceptron Consisting of two linear regression layers to enhance modeling complexity, which is widely accepted as a competitive method to improve fitting performance. X is a minibatch of input data and O is the output data, while H is known as H own as a hidden-layer variable or a hidden variable. Besides, W and b are the weight and biases of linear regression layers respectively.

Formally, we calculate the outputs of the one-hidden-layer MLP as follows:

$$H = XW^{(1)} + b^{(1)},$$
$$O = HW^{(2)} + b^{(2)},$$

Note that after adding the hidden layer, our model now requires us to track and update additional sets of parameters, while we have got nothing in exchange.

$$O = \left(XW^1 + b^{(1)}\right)W^{(2)} + b^{(2)} = XW^{(1)}W^{(2)} + b^{(1)}W^{(2)} + b^{(2)} = XW + b,$$

We can view the equivalence formally by proving that for any values of the weights, we can just collapse out the hidden layer, yielding an equivalent single-layer model. In order to realize the potential of multilayer architectures, we need one more key ingredient: a nonlinear activation function σ to be applied to each hidden unit following the affine transformation. In general, with activation functions in place, it is no longer possible to collapse our MLP into a linear model:

$$H = \sigma\left(XW^{(1)} + b^{(1)}\right),$$
$$O = HW^{(2)} + b^{(2)},$$

5.1.4 Typical Activation Functions

Activation functions decide whether a neuron should be activated or not by calculating the weighted sum and further adding bias with it. They are differentiable operators to transform input signals to outputs, while most of them add non-linearity.

The most popular choice, due to both simplicity of implementation and its good performance on a variety of predictive tasks, is the rectified linear unit (ReLU). ReLU provides a very simple nonlinear transformation. Given an element x, the function is defined as the maximum of that element and 0:

$$ReLU(x) = max(x, 0),$$

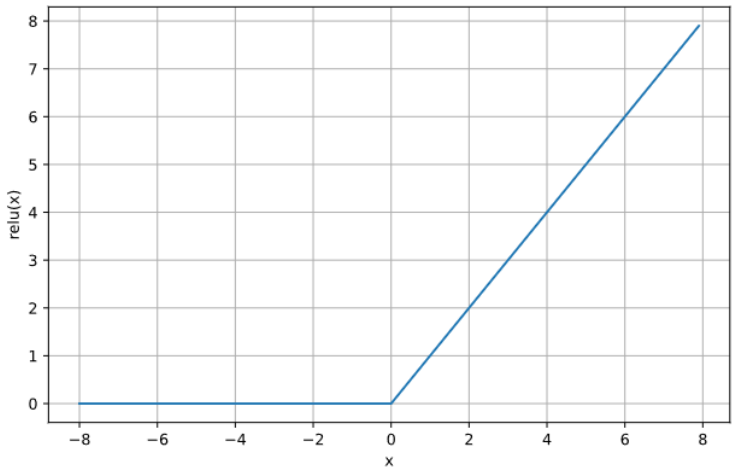

Figure 5.2 Profile of ReLU function [38]

The utilization of multilayer perceptrons can significantly improve modeling fitting performance, while this process is always accompanied by risking over-fitting.

Classical generalization theory suggests that to close the gap between train and test performance, we should aim for a simple model. Simplicity can come in the form of a small number of dimensions. Another useful notion of simplicity is smoothness, that the function should not be sensitive to small changes to its inputs. For instance, when we classify images, we would expect that adding some random noise to the pixels should be mostly harmless.

Srivastava et al (2014) [39] proposed to inject noise into each layer of the network before calculating the subsequent layer during training. They realized that when training a deep network with many layers, injecting noise enforces smoothness just on the input-output mapping.

Their idea, called dropout, involves injecting noise while computing each internal layer during forward propagation, and it has become a standard technique for training neural networks. The method is called dropout because we literally drop out some neurons during training. Throughout the training, on each iteration, standard dropout consists of zeroing out some fraction of the nodes in each layer before calculating the subsequent layer.

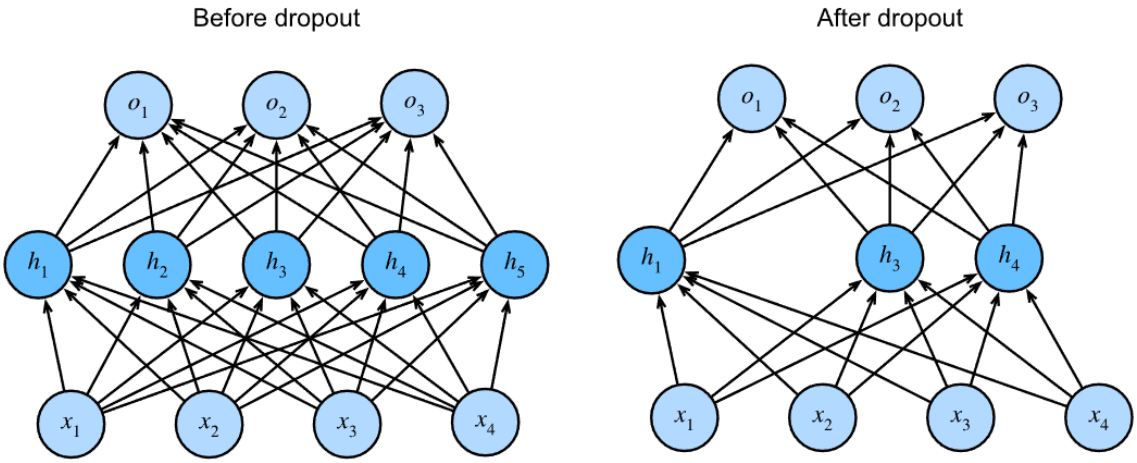

Figure 5.3 MLP before and after dropout [38]

In standard dropout regularization, one debiases each layer by normalizing by the fraction of nodes that were retained (not dropped out). In other words, with dropout probability p, each intermediate activation h is replaced by a random variable h' as follows:

$$h' = \begin{cases} 0 & with\ probability\ p \\ \dfrac{h}{1-p} & otherwise \end{cases}$$

By design, the expectation remains unchanged.

Recall the MLP with a hidden layer and 5 hidden units in Fig. 5.1. When we apply dropout to a hidden layer, zeroing out each hidden unit with probability p, the result can be viewed as a network containing only a subset of the original neurons. In Fig. 5.3, h_2 and h_5 are removed. Consequently, the calculation of the outputs no longer depends on h_2 or h_5 and their respective gradient also vanishes when performing backpropagation. In this way, the calculation of the output layer cannot be overly dependent on any one element of $h_1,\ldots,h_5$.

5.1.4 Training

With canopy height and shape factor, as with stock prices, we care about relative quantities more than absolute quantities. Thus, we tend to care more about the relative error $\frac{y-\hat{y}}{y}$ than the absolute error $y-\hat{y}$. One way to address this problem is to measure the discrepancy in the logarithm of the price estimates. In fact, this is also the official error measure used by the competition to evaluate the quality of submissions. After all, a small value δ for $|\log y - \log \hat{y}| \leq \delta$ translates into $e^{-\delta} \leq \frac{\hat{y}}{y} \leq e^{\delta}$. This leads to the following root-mean-squared-error between the logarithm of the predicted value and the logarithm of the observed value:

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\log y_i - \log \hat{y}_i)^2},$$

This study's training function will rely on the Adam optimizer. Compared to SGD, the main appeal of this optimizer is that, despite doing no better (and sometimes worse) given unlimited resources for hyperparameter optimization, people tend to find that it is significantly less sensitive to the initial learning rate.

As all canopy bulk geometric variables and observed data are from LES-urban, I am not even able to afford to hold out enough data to constitute a proper validation set. One popular solution to this problem is to employ K-fold cross-validation. Here, the original training data is split into K non-overlapping subsets. Then model training and validation are executed K times, each time training on K−1 subsets and validating on a different subset (the one not used for training in that round). Finally, the training and validation errors are estimated by averaging the results from the K experiments. The training and verification error averages are returned when we train K times in the K-fold cross-validation.

CHAPGTER. 6 Result

6.1 Canopy Height and shape factor Simulated by MLP

As is introduced in chapter 4, the negative peak of vertical momentum flux H_{uw} is assumed to be the most competitive for canopy height δ. Then, I normalized the vertical length scale z by H_{uw} and wind speed u by $u_{H_{uw}}$. At last, I calculated Λ at $0.25H_{uw}$ as the local shape factor. It seems to be not that reasonable at first glance. However, it should be taken into account that it is the contribution of ΛG_η rather than the role of Λ alone, which is able to influences the local velocity profile significantly. As it is shown in Figure 4.5, G_η is numerically significant where it is at 0.25δ. At other altitudes, especially where $\eta = 0$ and $\eta = 1$, its' contribution to the velocity becomes negligible. On the other hand, as it is shown in Figure 1.8, the influence of laminar flow is dominant only in the button area of the urban canopy($0 < \eta < 0.1\delta$). The turbulent flow is well developed at 0.25δ and that's why Λ at $0.25H_{uw}$ is capable of describing the geometric features for local blocks.

As there is no other data source to validate the MLP model's training performance, K-fold cross-validation was employed. The original dataset, training datasets, was split into 5 parts. In the 5 training stages, the 5 parts of the dataset are utilized in terms to do validation. Thus, the low validation log RMSE can prove that the modeling can perform pretty well when it is utilized to predict the canopy height and shape factor even for an urban canopy totally new to the train data.

To make a comparison, the performance of linear regression and MLP will be compared.

6.1.1 shape factor Simulated by MLP

The components of the traditional linear regression model are given as:

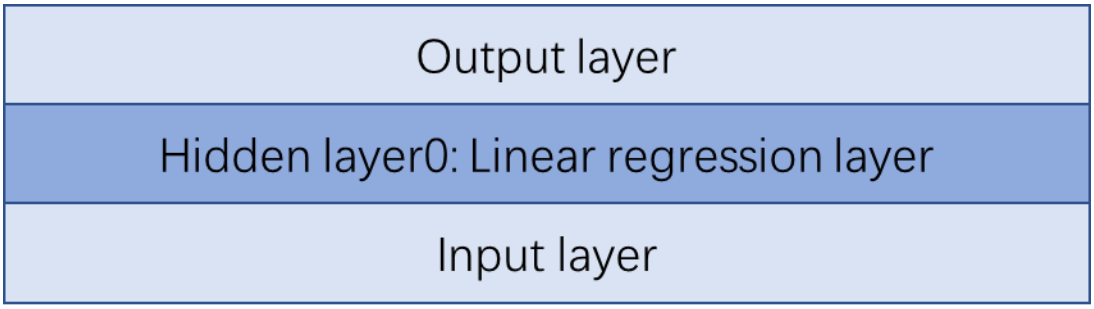

Figure 6.1 Linear Regression Model components

The epoch number for each training stage is set to be 500. Accordingly, the learning rate is set to be 0.001

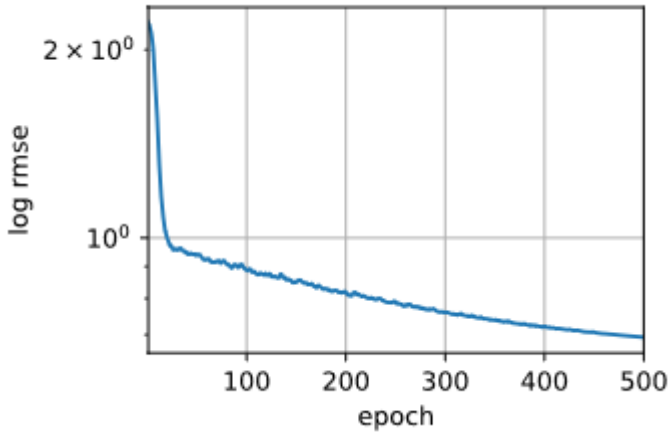

Figure 6.2 The bias between test data and predicted data

The bias decreases sharply before 200 times training and then the decrease becomes pretty slow. Finally, the bias maintains as 0.693551.

The model parameters are given as:

Layer 0: linear layer
Weight 0: [0.5052, 0.0241, -0.2976]
Bias 0: 2.3064

The components of MLP are given as

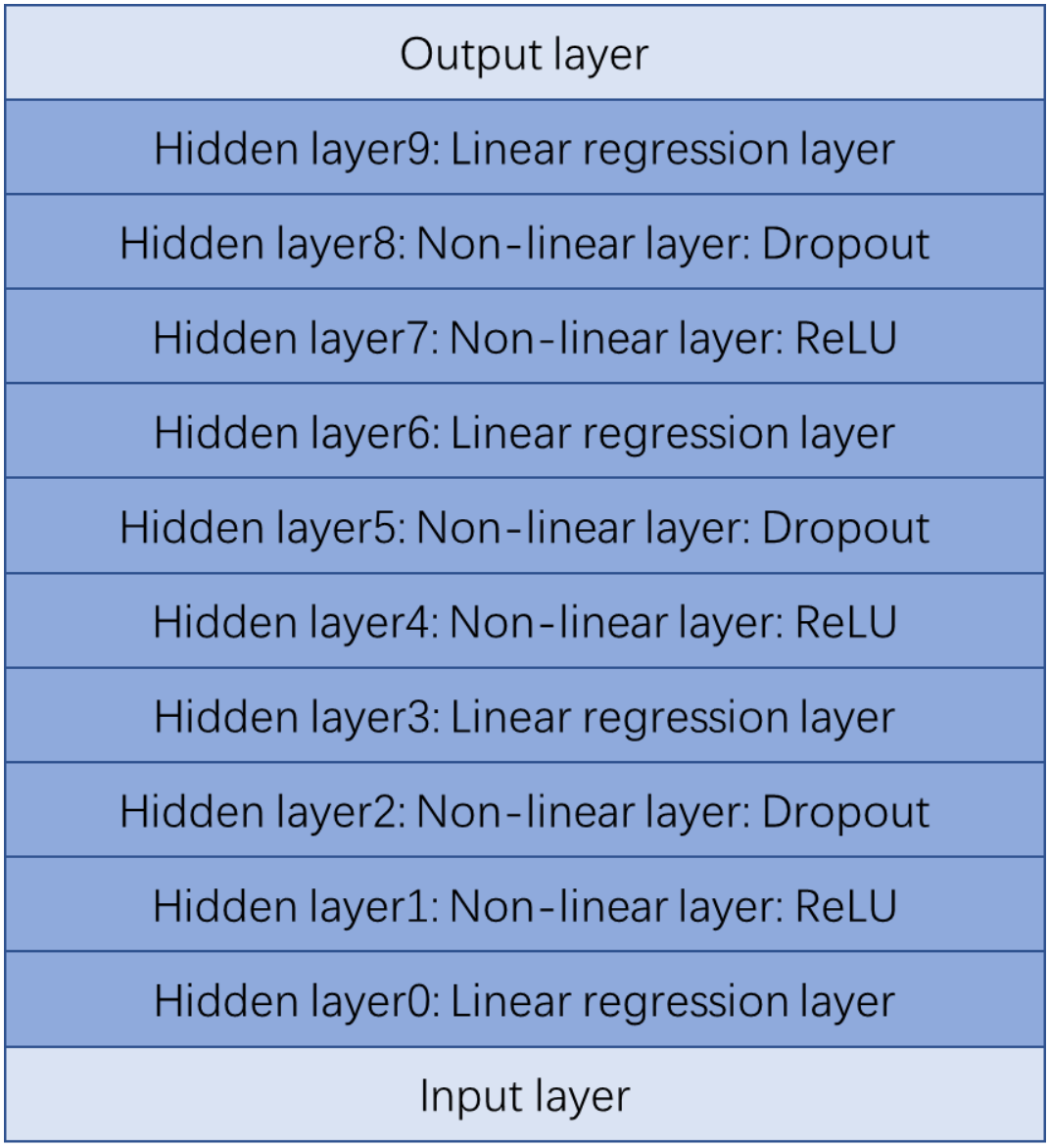

Figure 6.3 MLP components

The learning rate maintains 0.001. Also, the introduction of dropout layers has some noise injected into layers. As a result, the robustness is enhanced. Finally, the bias maintains as 0.655406.

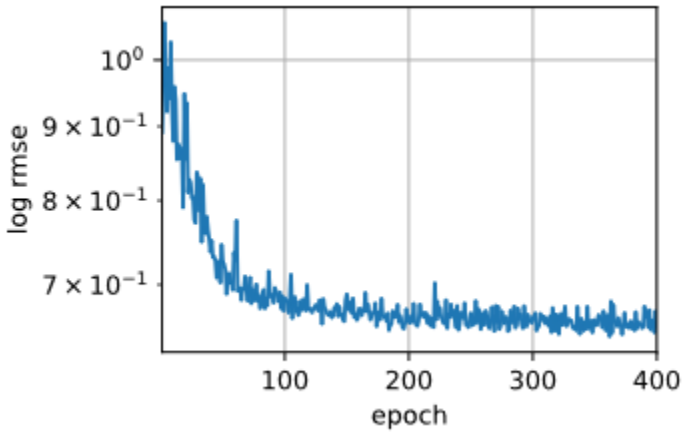

Figure 6.4 The bias between test data and predicted data

The model parameters are given as:

Layer 0: linear layer

Weight0:

[[3.1883e-01, 1.1630e-01, -5.8901e-01], [-5.3194e-01, -3.5612e-02, -5.0157e-01],
[3.9762e-01, 2.5054e-01, -6.1829e-01], [3.8939e-02, -1.1781e-01, -2.2913e-01],
[1.7379e-01, 1.7430e-01, 2.2754e-01], [-3.3412e-01, -4.0481e-01, -1.0282e-01],
[3.0190e-01, -4.6410e-01, -2.8157e-01], [-1.6452e-01, -1.1443e-01, 5.5956e-01],
[-3.0243e-01, -2.2352e-01, -2.8019e-01], [-7.6986e-02, 6.7840e-02, -3.2474e-01],
[-5.4202e-01, -2.2381e-01, 1.7839e-01], [5.0463e-01, -4.2084e-01, -5.3314e-01],
[-7.5232e-01, 2.1721e-01, 4.0491e-01], [3.4983e-01, -2.9771e-01, 4.1514e-01],
[1.3381e-01, 2.6990e-01, -4.1475e-01], [-2.0922e-01, -4.9653e-04, 4.1777e-01],
[2.9823e-01, 2.3948e-01, 1.6057e-01], [-1.1743e-02, 3.5408e-01, -6.4086e-01],
[-7.8270e-02, -1.5995e-01, -3.8884e-01], [8.3452e-02, 6.2681e-03, -2.2459e-01],
[-1.4837e-02, -3.2675e-01, 4.4351e-03], [3.1927e-01, -2.6191e-01, -1.8118e-01],
[3.0355e-01, -3.6933e-02, 1.5892e-01], [2.1200e-01, 8.2536e-02, -6.8460e-02],
[2.2817e-01, 8.8609e-02, -6.5848e-02], [1.7914e-01, -4.6667e-01, 2.0291e-01],
[-7.5970e-02, -8.9831e-02, 2.0126e-01], [4.8089e-01, -1.6736e-02, -3.5889e-01],
[9.7811e-02, -2.7037e-02, 4.4085e-01], [2.5879e-01, -1.4638e-01, -2.4663e-01],

[1.7419e-01, -2.6131e-01, 4.7237e-01], [1.6298e-01, 1.8511e-03, -3.1336e-01],
[2.4219e-01, -1.2651e-02, -2.9405e-01], [-4.7738e-01, 1.8357e-01, -2.8179e-02],
[-5.5151e-01, 2.3074e-01, -2.9715e-01], [-3.6551e-01, 4.5260e-03, 4.8313e-01]]

Bias0:
[0.0594, 0.4296, 1.1348, -0.0910, 0.8861, -0.4874, -0.1911, -0.5425,
  0.2088, -1.7427, 0.5117, 0.1694, -0.9521, -0.0516, 1.0688, 0.4992,
  0.3804, 1.5019, -0.3914, 1.6235, -0.5072, 0.0469, 1.7512, 0.3994,
  0.0840, 0.3211, 0.1509, 0.8237, 0.4443, 0.2007, 0.3675, 1.7152,
  1.8500, 1.4030, 0.4135, -1.2858]

Layer 1: ReLU

Layer 2: Dropout

Layer 3: Linear layer

Weight3 :
[[0.0610, -0.0092, 0.1431, ..., 0.0180, -0.0219, 0.1357],
 [0.0421, -0.0843, 0.0652, ..., 0.0933, 0.0302, -0.1632],
 [-0.1478, 0.0363, -0.1360, ..., 0.1154, 0.0117, 0.0623],
 ...,
 [0.0527, -0.0708, -0.1409, ..., -0.0501, 0.0349, -0.1473],
 [0.0429, 0.0832, 0.0497, ..., 0.1353, -0.0791, -0.2811],
 [0.0086, 0.0455, -0.0145, ..., 0.1798, 0.1608, 0.0726]]

Bias3:
[1.2437, 1.2949, -0.0999, 1.3931, 0.6267, 1.1190, 1.3892, 0.0576,
  1.4471, 1.5556, 1.5903, 1.1809, 0.2126, 1.2505, 1.0710, 1.4447,
  0.0451, 1.4975, 1.0408, 0.0412, 1.4850, -0.0869, 1.5063, 1.5969,
  1.7023, 1.5098, 1.4677, 0.0068, -0.1442, 0.1959, 0.1560, 1.5150,
  1.3991, 1.0086, 1.1120, 1.4943, 1.2248, 0.1094, 1.6768, -0.1497,
  0.1436, -0.1731, 1.5654, 0.8960, 1.3053, 1.6458, 1.3493, -0.1969,
  1.2725, -0.0195, -0.2857, 1.6460, 1.0070, 1.3184, 1.5878, -0.3357,
  0.1544, -0.0864, 0.0705, 0.1230, 1.1906, 0.9752, 1.6593, 0.2868,
  0.1110, 1.5692, 1.2533, 1.3392, 0.6902, 0.1308, 0.0594, 1.5452,
 -0.1676, 1.1706, -0.3209, -0.0288, -0.0917, 1.2923, 1.3014, 1.6133,
 -0.5448, 0.7286, 0.0355, -0.4072, 1.6396, 0.2267, 1.3661, 1.5300,

0.4674, -0.0686, -0.0615, 1.3360, 1.4873, 1.2682, 0.2598, -0.0475,
-0.4176, 0.0977, 1.1994, 1.2815, 1.4772, 0.9800, -0.0473, 0.6060,
-1.1143, 1.1060, 0.1224, 0.9529, -0.0862, 1.7523, 0.0214, 0.1130,
1.4465, 0.0383, 1.2736, 0.7910, 1.3099, 1.4951, 1.3730, 1.6808,
0.9495, 1.0245, 1.3160, 0.0083, 0.1279, 1.0819, 1.3888, 1.0815]

Layer 4: ReLU

Layer 5: Dropout

Layer 6: Linear layer
Weight6:

[[0.1258, 0.2024, 0.0418, ..., 0.0749, 0.0482, 0.0729],
[-0.0030, 0.1573, 0.0166, ..., 0.1400, 0.0472, 0.1268],
[0.0854, 0.1644, -0.0354, ..., -0.0965, 0.0994, -0.0114],
...,
[-0.0209, 0.0480, 0.0160, ..., 0.0208, -0.0086, 0.0365],
[0.1363, 0.1306, 0.0354, ..., 0.0304, 0.1019, 0.0963],
[-0.0791, 0.0600, -0.0402, ..., -0.0022, -0.0454, -0.0657]]

Bias6:

[1.1033e+00, 1.1929e+00, -1.7325e-01, 1.3990e+00, 1.2059e+00,
1.2203e+00, -1.9923e-04, 1.1198e+00, 1.1477e+00, 4.3964e-02,
1.6162e-02, 1.1650e+00, 1.0324e+00, 1.0832e+00, -2.0320e-01,
1.1120e+00, -9.6822e-02, -1.0106e-01, 1.2024e+00, 1.1946e+00,
5.3448e-02, 1.2836e+00, 1.1909e+00, 6.4903e-02, 1.1289e+00,
9.6680e-01, -8.3870e-02, 1.6582e+00, 2.2928e-02, 1.2209e+00,
-5.7140e-03, 1.3135e+00, 1.0366e+00, -9.1442e-02, 1.3685e+00,
-5.8625e-02])),

Layer7: ReLU

Layer8: Dropout

Layer9: Linear layer

Weight9:

[[0.1554, 0.0947, -0.1193, 0.2280, 0.1241, 0.1027, -0.0052, 0.2126,
  0.2325, 0.1347, 0.0042, 0.1136, 0.1471, 0.2050, -0.0618, 0.2237,
 -0.0226, -0.0758, 0.1343, 0.1418, -0.0025, 0.1381, 0.1162, -0.0879,
  0.1628, 0.1821, -0.1443, 0.1645, -0.1151, 0.1128, -0.0727, 0.1592,
  0.1578, -0.1057, 0.1122, -0.0002]]

Bias9:
[1.0839]

There are 85 sets of LES-urban used as training dataset and 37 sets of LES-urban used for validation. The observed shape factor and predicted shape factor are summarized in Table 6.1.

Compared to the traditional linear regression method, with log RMSE 0.693551, MLP improved predicting performance and the log RMSE becomes 0.655406. The result is not surprising: the shape factor describes the interior of a specific urban canopy. What we can foresee is that the vertical building density distribution will be a great addition to shape factor prediction. In the absence of the parameters capable of showing the interior construction, employing MLP is an impressive method to analyze data and the prediction bias is reduced by 7.7%.

Table 6.1 The comparison between observed shape factor and simulated shape factor, containing the bulk geometric variables

ID (m)	H_{ave} (m)	H_{max} (m)	σ_H (m)	Λ (observed)	Λ (Linear)	Λ (MLP)
93	9.3	39.9	5.1	13.52545	6.450537	10.30474
94	14.3	65.9	9.6	13.95535	8.26522	9.780249
95	16.8	55.4	11	5.454791	8.858326	9.365337
96	29.1	188.5	21.3	15.18284	15.22023	15.36781
97	37.4	154.1	41.3	12.43098	12.63231	13.07601
98	29.9	247	35	3.95537	12.95973	11.33288
99	26.3	115	19.4	15.44886	12.59703	12.42078
100	30.9	140	16.9	19.7788	16.2684	16.61418
101	13.8	52	10.6	11.99762	7.379571	9.696035
103	8.7	100	6.4	8.283385	7.211047	9.360319
104	11.3	85	8.9	7.210191	7.418758	9.681379
105	10.8	93	9.7	0.000562	7.121165	9.254066
106	16.5	70	14.5	4.968719	8.01766	10.41188
107	8.3	27	4.5	12.07715	5.812496	10.70492
108	8	8	0	5.203503	6.541386	9.795265
109	12	12	0	6.517325	8.658887	14.88993
110	16	16	0	6.297902	10.77639	14.18448
111	24	24	0	11.49662	15.01139	17.9156
112	36	36	0	17.46504	21.3639	19.4256
113	56	56	0	16.20505	31.9514	28.65522
114	10	10	0	53.0041	7.600137	12.54029
115	8	8	0	4.94069	6.541386	10.74481
116	12	12	0	7.859308	8.658887	11.35256
117	16	16	0	9.697388	10.77639	14.26003
118	24	24	0	21.32107	15.01139	15.69247
119	36	36	0	28.64552	21.3639	21.35461
120	56	56	0	25.77689	31.9514	18.53194
121	10	10	0	53.06726	7.600137	13.45432
122	10	10	0	50.69894	7.600137	10.57106
123	10	10	0	46.0651	7.600137	10.19259
124	11.8	17	3.1	16.85585	7.756077	12.21958
125	14.6	28	7.8	7.357745	8.037702	9.16894
126	11.5	16	2.6	30.25451	7.72915	12.00841
127	14.5	28	7.8	15.32348	7.987178	8.301858
128	11.5	16	2.6	28.12182	7.72915	11.68564
129	14.5	28	7.8	19.0787	7.987178	9.740983
130	14.6	28	7.8	20.50841	8.037702	9.573129

6.1.2 Canopy height simulated by MLP

The components of the traditional linear regression model are given as:

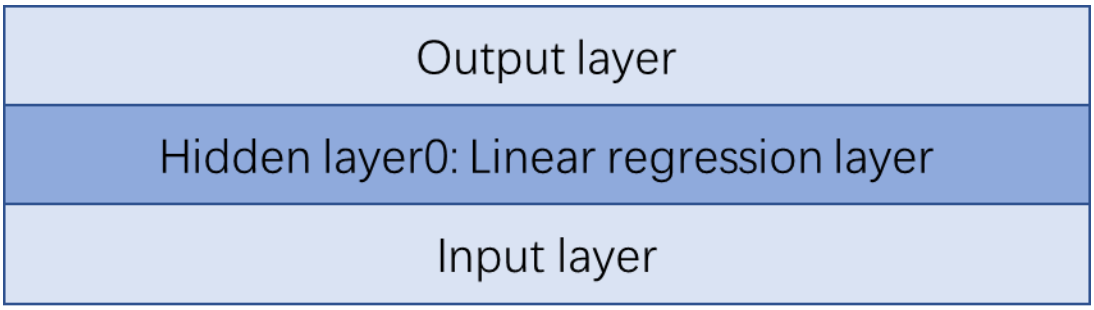

Figure 6.5 Linear Regression Model components

The epoch number for each train stage is set to be 500. Accordingly, learning rate is set to be 0.001.

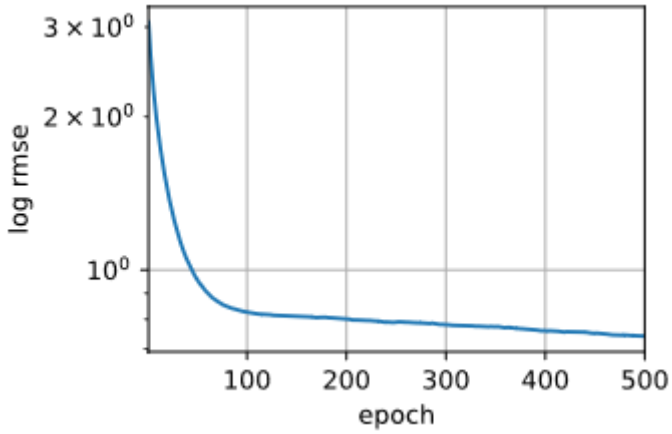

Figure 6.6 The bias between test data and predicted data

The bias decreases sharply before 100 times training and then the decrease become pretty slow. Finally, the bias maintains as 0.742313.

The model parameters are given as:

Layer 0: linear layer
Weight: 1.3199, 0.3430, -0.1607
Bias: 3.2314

The components of MLP are given as

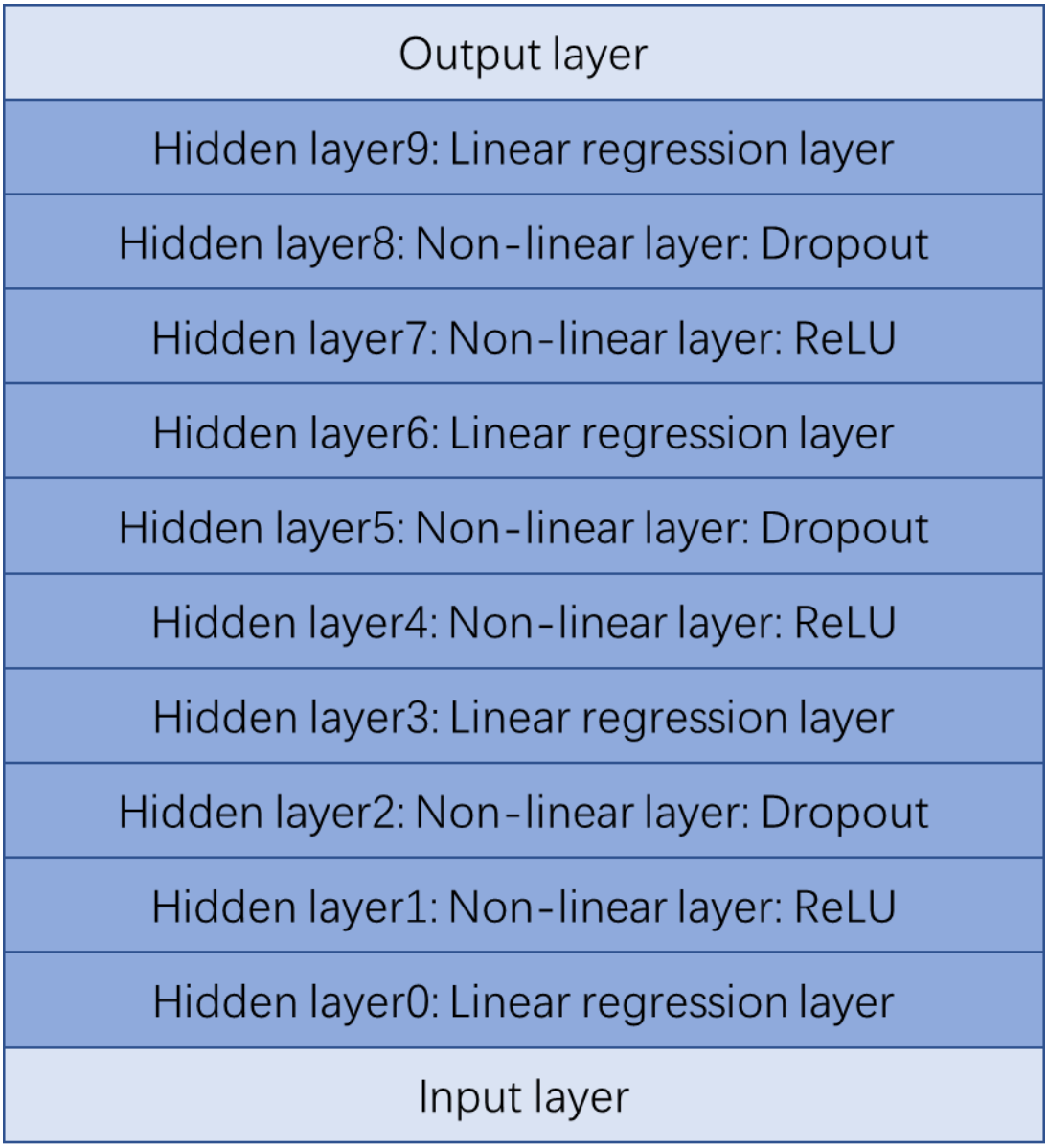

Figure 6.7 MLP components

The learning rate maintains 0.001. However, the model complexity is increased due to 4 linear regression layers applied. As a result, 400 epochs are sufficient to reduce bias. Also, the introduction of dropout layers has some noise injected to layers. As a result, the robustness is enhanced. Finally, the bias maintains as 0.562255.

The model parameters are given as:

Layer 0: linear layer

Weight 0:

[[3.1883e-01, 1.1630e-01, -5.8901e-01], [-5.3194e-01, -3.5612e-02, -5.0157e-01],
 [3.9762e-01, 2.5054e-01, -6.1829e-01], [3.8939e-02, -1.1781e-01, -2.2913e-01],
 [1.7379e-01, 1.7430e-01, 2.2754e-01], [-3.3412e-01, -4.0481e-01, -1.0282e-01],
 [3.0190e-01, -4.6410e-01, -2.8157e-01], [-1.6452e-01, -1.1443e-01, 5.5956e-01],
 [-3.0243e-01, -2.2352e-01, -2.8019e-01], [-7.6986e-02, 6.7840e-02, -3.2474e-01],
 [-5.4202e-01, -2.2381e-01, 1.7839e-01], [5.0463e-01, -4.2084e-01, -5.3314e-01],
 [-7.5232e-01, 2.1721e-01, 4.0491e-01], [3.4983e-01, -2.9771e-01, 4.1514e-01],
 [1.3381e-01, 2.6990e-01, -4.1475e-01], [-2.0922e-01, -4.9653e-04, 4.1777e-01],
 [2.9823e-01, 2.3948e-01, 1.6057e-01], [-1.1743e-02, 3.5408e-01, -6.4086e-01],
 [-7.8270e-02, -1.5995e-01, -3.8884e-01], [8.3452e-02, 6.2681e-03, -2.2459e-01],
 [-1.4837e-02, -3.2675e-01, 4.4351e-03], [3.1927e-01, -2.6191e-01, -1.8118e-01],
 [3.0355e-01, -3.6933e-02, 1.5892e-01], [2.1200e-01, 8.2536e-02, -6.8460e-02],
 [2.2817e-01, 8.8609e-02, -6.5848e-02], [1.7914e-01, -4.6667e-01, 2.0291e-01],
 [-7.5970e-02, -8.9831e-02, 2.0126e-01], [4.8089e-01, -1.6736e-02, -3.5889e-01],
 [9.7811e-02, -2.7037e-02, 4.4085e-01], [2.5879e-01, -1.4638e-01, -2.4663e-01],
 [1.7419e-01, -2.6131e-01, 4.7237e-01], [1.6298e-01, 1.8511e-03, -3.1336e-01],
 [2.4219e-01, -1.2651e-02, -2.9405e-01], [-4.7738e-01, 1.8357e-01, -2.8179e-02],
 [-5.5151e-01, 2.3074e-01, -2.9715e-01], [-3.6551e-01, 4.5260e-03, 4.8313e-01]]

Bias 0:

[0.0594, 0.4296, 1.1348, -0.0910, 0.8861, -0.4874, -0.1911, -0.5425,
  0.2088, -1.7427, 0.5117, 0.1694, -0.9521, -0.0516, 1.0688, 0.4992,
  0.3804, 1.5019, -0.3914, 1.6235, -0.5072, 0.0469, 1.7512, 0.3994,
  0.0840, 0.3211, 0.1509, 0.8237, 0.4443, 0.2007, 0.3675, 1.7152,
  1.8500, 1.4030, 0.4135, -1.2858]

Layer 1: ReLU

Layer 2: Dropout

Layer 3: Linear layer

Weight3 :

[[0.0610, -0.0092, 0.1431, ..., 0.0180, -0.0219, 0.1357],
[0.0421, -0.0843, 0.0652, ..., 0.0933, 0.0302, -0.1632],
[-0.1478, 0.0363, -0.1360, ..., 0.1154, 0.0117, 0.0623],
...,
[0.0527, -0.0708, -0.1409, ..., -0.0501, 0.0349, -0.1473],
[0.0429, 0.0832, 0.0497, ..., 0.1353, -0.0791, -0.2811],
[0.0086, 0.0455, -0.0145, ..., 0.1798, 0.1608, 0.0726]]

Bias 3:

[1.2437, 1.2949, -0.0999, 1.3931, 0.6267, 1.1190, 1.3892, 0.0576,
1.4471, 1.5556, 1.5903, 1.1809, 0.2126, 1.2505, 1.0710, 1.4447,
0.0451, 1.4975, 1.0408, 0.0412, 1.4850, -0.0869, 1.5063, 1.5969,
1.7023, 1.5098, 1.4677, 0.0068, -0.1442, 0.1959, 0.1560, 1.5150,
1.3991, 1.0086, 1.1120, 1.4943, 1.2248, 0.1094, 1.6768, -0.1497,
0.1436, -0.1731, 1.5654, 0.8960, 1.3053, 1.6458, 1.3493, -0.1969,
1.2725, -0.0195, -0.2857, 1.6460, 1.0070, 1.3184, 1.5878, -0.3357,
0.1544, -0.0864, 0.0705, 0.1230, 1.1906, 0.9752, 1.6593, 0.2868,
0.1110, 1.5692, 1.2533, 1.3392, 0.6902, 0.1308, 0.0594, 1.5452,
-0.1676, 1.1706, -0.3209, -0.0288, -0.0917, 1.2923, 1.3014, 1.6133,
-0.5448, 0.7286, 0.0355, -0.4072, 1.6396, 0.2267, 1.3661, 1.5300,
0.4674, -0.0686, -0.0615, 1.3360, 1.4873, 1.2682, 0.2598, -0.0475,
-0.4176, 0.0977, 1.1994, 1.2815, 1.4772, 0.9800, -0.0473, 0.6060,
-1.1143, 1.1060, 0.1224, 0.9529, -0.0862, 1.7523, 0.0214, 0.1130,
1.4465, 0.0383, 1.2736, 0.7910, 1.3099, 1.4951, 1.3730, 1.6808,
0.9495, 1.0245, 1.3160, 0.0083, 0.1279, 1.0819, 1.3888, 1.0815]

Layer 4: ReLU

Layer 5: Dropout

Layer 6: Linear layer

Weight 6:

[[0.1258, 0.2024, 0.0418, ..., 0.0749, 0.0482, 0.0729],
[-0.0030, 0.1573, 0.0166, ..., 0.1400, 0.0472, 0.1268],
[0.0854, 0.1644, -0.0354, ..., -0.0965, 0.0994, -0.0114],
...,
[-0.0209, 0.0480, 0.0160, ..., 0.0208, -0.0086, 0.0365],
[0.1363, 0.1306, 0.0354, ..., 0.0304, 0.1019, 0.0963],
[-0.0791, 0.0600, -0.0402, ..., -0.0022, -0.0454, -0.0657]]

Bias 6:

[1.1033e+00, 1.1929e+00, -1.7325e-01, 1.3990e+00, 1.2059e+00,
  1.2203e+00, -1.9923e-04, 1.1198e+00, 1.1477e+00, 4.3964e-02,
  1.6162e-02, 1.1650e+00, 1.0324e+00, 1.0832e+00, -2.0320e-01,
  1.1120e+00, -9.6822e-02, -1.0106e-01, 1.2024e+00, 1.1946e+00,
  5.3448e-02, 1.2836e+00, 1.1909e+00, 6.4903e-02, 1.1289e+00,
  9.6680e-01, -8.3870e-02, 1.6582e+00, 2.2928e-02, 1.2209e+00,
 -5.7140e-03, 1.3135e+00, 1.0366e+00, -9.1442e-02, 1.3685e+00,
 -5.8625e-02]

Layer7: ReLU

Layer8: Dropout

Layer9: Linear layer

Weight9:

[0.1554, 0.0947, -0.1193, 0.2280, 0.1241, 0.1027, -0.0052, 0.2126,
  0.2325, 0.1347, 0.0042, 0.1136, 0.1471, 0.2050, -0.0618, 0.2237,
 -0.0226, -0.0758, 0.1343, 0.1418, -0.0025, 0.1381, 0.1162, -0.0879,
  0.1628, 0.1821, -0.1443, 0.1645, -0.1151, 0.1128, -0.0727, 0.1592,
  0.1578, -0.1057, 0.1122, -0.0002]

Bias 9:

[1.0839]

There are 85 sets of LES-urban are used as training dataset and 45 sets of LES-urban are used for validation. The observed shape factor and predicted shape factor are summarized in Table 6.2 as

Table 6.2 The comparison between observed $H_{uw}$and simulated H_{uw}, containing the bulk geometric variables

ID (m)	H_{ave} (m)	H_{max} (m)	σ_H (m)	H_{uw} (observed)	H_{uw} (Linear)	H_{uw} (MLP)
86	18.3	41.9	11	36	39.99049	60.907
87	22.5	44.4	11.7	38	46.27919	63.01762
88	9.7	40.1	5.7	38	28.87361	50.87933
89	10	44.4	6.2	40	30.66422	50.21308
90	8.7	46.6	6	40	29.73511	64.29086
91	13.8	53.3	11.8	48	37.83272	62.61169
92	14.3	66.1	11.2	58	42.97977	51.76644
93	9.3	39.9	5.1	32	28.37348	59.45463
94	14.3	65.9	9.6	44	43.16834	42.16148
95	16.8	55.4	11	44	42.6414	67.00761
96	29.1	188.5	21.3	118	102.8769	76.36995
97	37.4	154.1	41.3	146	98.81784	76.26255
98	29.9	247	35	235	121.7976	90.75718
99	26.3	115	19.4	110	74.27454	84.14059
100	30.9	140	16.9	100	89.32349	94.71799
101	13.8	52	10.6	47.5	37.57966	51.99029
102	9.8	82	11.6	80	42.42989	51.9672
103	8.7	100	6.4	45	47.98812	67.06039
104	11.3	85	8.9	45	45.87278	57.8982
105	10.8	93	9.7	87.5	47.8284	69.75746
106	16.5	70	14.5	50	46.69098	55.59009
107	8.3	27	4.5	25	22.72504	53.46703
108	8	8	0	10	16.53493	70.54001
109	12	12	0	14	23.18669	80.37418
110	16	16	0	18	29.83844	101.6739
111	24	24	0	26	43.14194	125.4774
112	36	36	0	38	63.0972	74.11012
113	56	56	0	58	96.35596	175.8824
114	10	10	0	12	19.86081	64.75251
115	8	8	0	10	16.53493	60.66944
116	12	12	0	14	23.18669	83.66654
117	16	16	0	18	29.83844	102.0347
118	24	24	0	26	43.14194	109.846
119	36	36	0	38	63.0972	159.9987
120	56	56	0	58	96.35596	209.5878
121	10	10	0	14	19.86081	90.83289
122	10	10	0	14	19.86081	77.05594
123	10	10	0	14	19.86081	70.85958
124	11.8	17	3.1	22	24.13956	73.7025
125	14.6	28	7.8	32	30.85314	64.38362
126	11.5	16	2.6	20	23.48092	80.54179
127	14.5	28	7.8	32	30.72115	57.84172
128	11.5	16	2.6	20	23.48092	63.57747
129	14.5	28	7.8	32	30.72115	56.83765
130	14.6	28	7.8	32	30.85314	73.94659

Compared to the traditional linear regression method, with log RMSE 0.742313, MLP improved predicting performance and the log RMSE becomes 0.562255. In absence of the parameters capable of showing the interior construction, employing MLP is an impressive method to analyze data and the prediction bias is reduced by 32%.

6.2 Simulation Validation

As Tokyo tower is a landmark for ID63 district, it contributes much to the local canopy's geometric features. The Tokyo tower is a quadrangular pyramid 333 m tall and with an 80 m length at the basement, while the majority of other buildings are with relatively low vertical length scale. In other words, it significantly increased maximum building height and building height deviation.

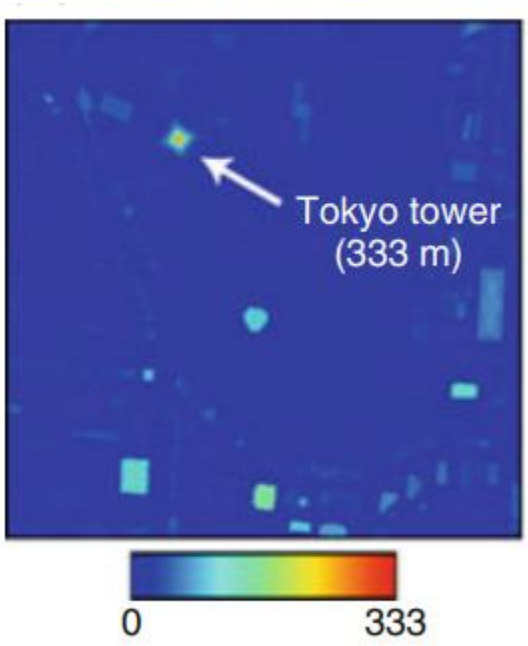

Figure 6.8 Map of building height for ID 63 [26]

As the CM-BEM is utilized to analyze energy consumption, Tokyo tower district, a prosperous commercial district, is the last to leave behind. The size of this canopy is $500\ \mathrm{m} * 500\ \mathrm{m}$. As it is shown in Figure 6.9, This particular canopy has a large green land included, which means Tokyo tower significantly influenced the local flux transfer process. As a conclusion, I can evaluate mixing length approximation performance by comparing simulation results derived from CM-BEM and new parametrization respectively.

The CM-BEM simulation period was July 28th, 2002 to August 12th, 2002, during which time period there was no rainfall. Thus, the atmospheric layers are in a relatively stable situation. Unfortunately, there was no other sounding spot except for Tokyo tower which was

able to collect meteorological data far from the ground. However, it will do no harm for us to research this district to validate simulation performance.

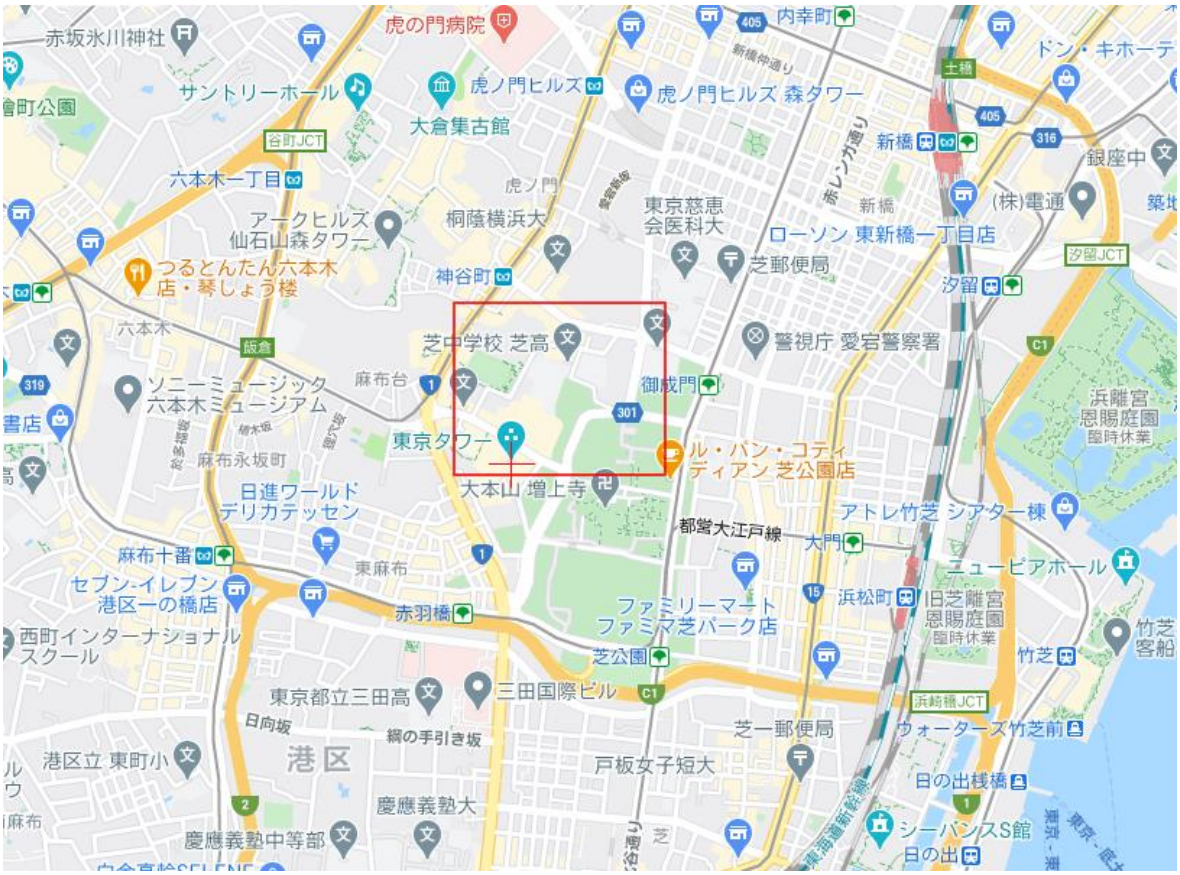

Figure 6.9 the specific CM-BEM canopy having Tokyo tower involved

6.2.1 Mixing Length Validation

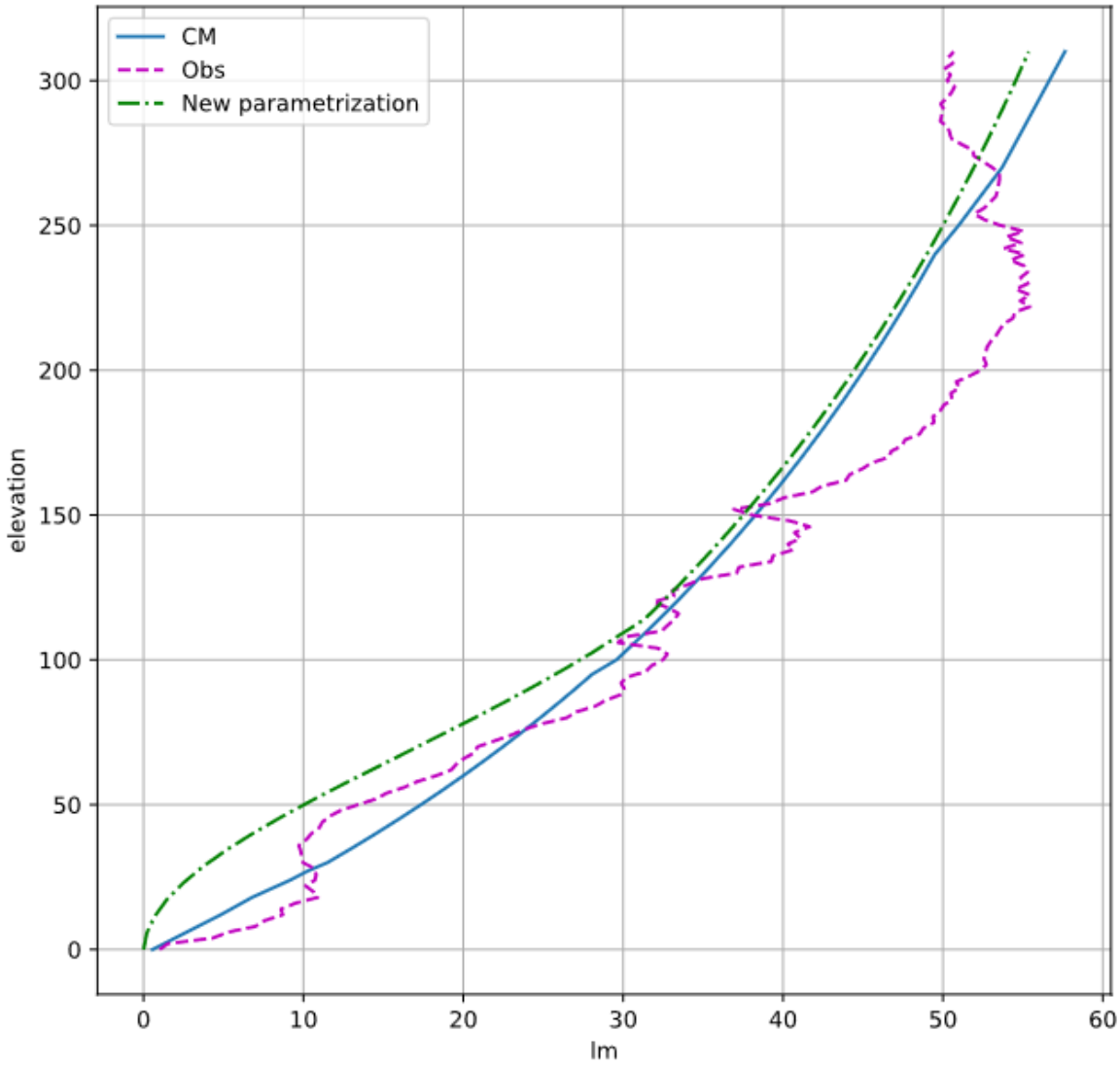

Figure 6.10 Profile of Mixing length in Tokyo tower district

According to LES-urban, the canopy height H_{uw} and shape factor Λ is given as $114m$ and -8.55911 respectively. Referring to chapter 4.4, I can approximate the profile of mixing length by new solution as *New Parametrization.*

Obs is the mixing length profile simulated by vertical momentum flux and wind speed gradient acquired from LES-urban as

$$l_m = \frac{\sqrt{-\langle u'w' \rangle}}{\left| \frac{d\sqrt{u^2 + v^2}}{dz} \right|}.$$

CM is the mixing length profile obtained from the CM-BEM simulation result. Given the fact that the noise involved in *Obs* is rather hard to ignore, we can draw the conclusion safely that both new parametrization and CM-BEM can successfully calculate mixing length.

As Tokyo tower district is a typical commercial district, Figure 6.10 strongly demonstrated that new parametrization is a rather competitive candidate to simulate mixing length.

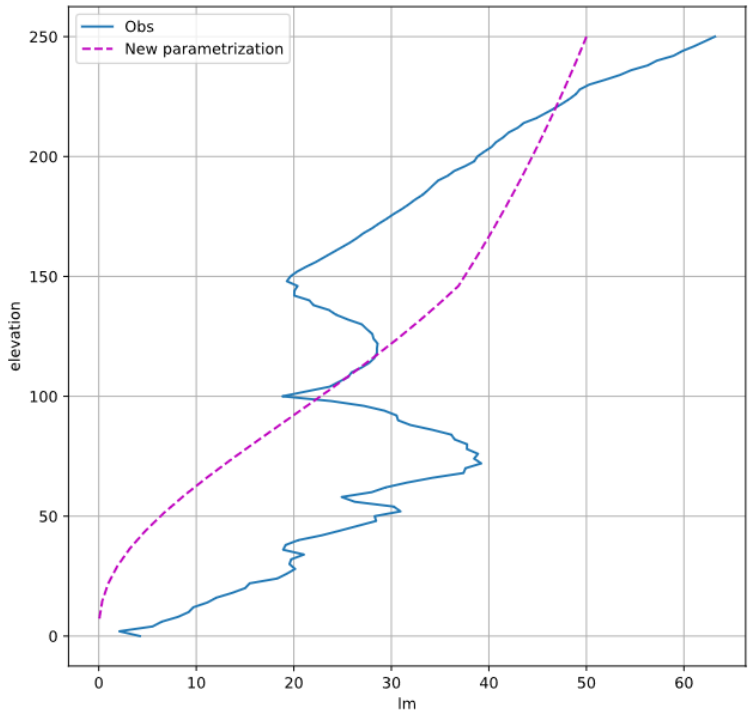

Figure 6.11 Profile of Mixing length in ID 97

ID97 is with a cluster of skyscrapers with large maximum building height and building height deviation produced a large momentum flux, which means that numerous amounts of kinetic energy should be generated and that usually influenced flux transfer process in the upper atmospheric layers. However, as it is shown in Figure 6.11, the new parametrization performs pretty well under skyscraper circumstances. Notably, Kanda (2013) [26] gave the conclusion

that the towers in the local district will generate a large amount of kinetic energy in the upper atmospheric layers, which usually results in bad simulation performance. Thus, it is rather impressive that the new parametrization performs pretty well in these two districts distinguished with several skyscrapers. As long as commercial districts generally with the majority of the air-conditioning load, the new parametrization of mixing length will be a great addition to CM-BEM.

6.2.2 Wind Speed Validation

To improve simulation performance, the integration time for CM-BEM simulation is set to be one day. Thus, I started to compare the simulation result from July 29th. The velocity sounding spots are 24 m, 110 m, and 250 m respectively from the ground surface.

Firstly, let's refer to Table 6.4 to see the wind profile in the Tokyo tower district on July 29th. [40]

Table 6.3 wind profile on July 29th

July 29th	Velocity observed at Tokyo tower			Velocity simulated by CM-BEM			Velocity acquired from WRF			Velocity from WRF- velocity observed		
Time	25m (m/s)	107m (m/s)	250m (m/s)	25m (m/s)	107m (m/s)	250m (m/s)	25m (m/s)	107m (m/s)	250m (m/s)	25m (m/s)	107m (m/s)	250m (m/s)
1:00	1.6	0.7	0.6	1.2	2.8	4.3	0.6	1.3	4.1	-0.97	0.56	3.48
2:00	1.5	1.1	0.9	1.1	2.4	3.7	0.2	1.1	3.7	-1.26	-0.02	2.75
3:00	1.2	0.8	1.4	1.1	2.3	3.3	0.6	1.7	3.4	-0.58	0.86	2.04
4:00	1.8	1.2	0.9	1.1	2.1	3.1	1.4	2.1	3.1	-0.38	0.87	2.22
5:00	1.1	0.8	1.4	1.1	2.2	3.4	2.2	2.5	3.3	1.06	1.71	1.88
6:00	1.2	0.8	0.9	1.3	2.6	3.8	2.7	3.0	3.6	1.51	2.21	2.68
7:00	1.5	0.9	0.5	1.4	2.6	3.7	3.1	3.5	3.9	1.62	2.60	3.41
8:00	1.7	1.3	0.7	1.6	3.0	4.3	4.1	4.4	4.8	2.40	3.14	4.11
9:00	1.4	1.1	0.9	1.9	3.5	4.9	4.5	4.9	5.3	3.10	3.78	4.41
10:00	1.7	0.5	0.4	1.9	3.5	4.8	4.1	4.6	5.2	2.42	4.05	4.77
11:00	1.7	1.8	1.5	1.8	3.2	4.3	3.5	3.9	4.4	1.79	2.06	2.95
12:00	2.5	1.1	0.3	1.7	2.9	3.9	4.1	4.3	4.4	1.64	3.23	4.07
13:00	4	1.7	3.4	1.8	3.2	4.4	4.4	4.6	4.8	0.40	2.94	1.43
14:00	3.8	3.1	3.6	2.0	3.6	4.9	4.3	4.7	5.2	0.54	1.59	1.59
15:00	3.4	3.7	3.3	1.3	2.2	2.7	1.7	2.0	2.7	-1.66	-1.69	-0.61
16:00	1.2	2.2	3	1.0	1.6	2.2	2.5	2.8	2.4	1.29	0.55	-0.57
17:00	2.1	3	3.6	1.2	1.9	2.3	2.4	2.7	2.5	0.27	-0.31	-1.09
18:00	1.4	4.4	4.6	0.8	1.4	2.3	1.4	1.7	2.4	0.00	-2.73	-2.24
19:00	1.4	2.6	3.1	1.7	3.4	5.0	2.6	3.4	4.8	1.19	0.79	1.71
20:00	1.8	2.4	3.2	1.6	3.4	5.2	3.9	4.4	5.2	2.10	2.04	1.96
21:00	2.4	2.7	4.1	1.7	3.8	5.8	4.9	5.4	6.1	2.47	2.72	2.01
22:00	2.1	3.3	4.1	1.7	3.7	5.5	5.1	5.6	6.0	3.02	2.32	1.94
23:00	1.6	2.1	2.1	1.6	3.6	5.3	4.8	5.3	5.7	3.25	3.20	3.58
0:00	1.2	2.8	2.9	1.6	3.6	5.5	4.5	5.0	5.6	3.28	2.19	2.75
Average	1.9	1.9	2.1	1.5	2.8	4.1	3.1	3.5	4.3	1.2	1.6	2.1

Given the fact that 250 m is pretty far from the ground, there should be not that much canopy elements' density at that altitude. In other words, velocity observed at that altitude should be very close to the velocity acquired from WRF, which is considered to be close enough to geostrophic wind and capable of initializing atmospheric components. However, as we can

see in Table 6.5, the velocity from WRF is generally bigger than the velocity observed at 250 m high, we can assume that WRF overestimated velocity and that can result in over-valuation of flux in the early simulation process.

Table 6.4 daily average wind speed for the whole simulation period

Time		Daily average windspeed			Velocity simulated by CM-BEM			Velocity acquired by WRF		
MONTH	DAY	25m (m/s)	107m (m/s)	250m (m/s)	25m (m/s)	107m (m/s)	250m (m/s)	25m (m/s)	107m (m/s)	250m (m/s)
JUL	29th	1.9	1.9	2.1	1.47	2.85	4.10	3.08	3.53	4.28
JUL	30th	2	2.4	2.7	1.76	3.71	5.61	4.75	5.34	6.11
JUL	31th	2.3	2.6	3	1.57	3.16	4.69	3.51	4.19	4.99
AUG	1st	2	2.3	2.6	2.01	4.24	6.48	3.86	4.91	6.57
AUG	2nd	1.4	1.8	2.5	2.20	4.75	7.30	4.24	5.33	7.49
AUG	3rd	2.1	1.6	2.8	2.99	6.94	11.08	6.40	8.59	11.75
AUG	4th	1.8	2.3	2.4	3.02	7.01	11.16	6.19	8.64	12.01
AUG	5th	1.7	1.7	2	1.55	2.93	4.18	2.86	3.60	4.75
AUG	6th	1.9	2.2	2.5	1.57	3.03	4.38	2.81	3.41	4.59
AUG	7th	2.5	3.5	4.4	2.05	4.31	6.56	4.07	5.33	7.15
AUG	8th	2.5	5.6	8.3	1.72	3.47	5.12	3.51	4.41	5.56
AUG	9th	3.4	6.6	10	1.87	3.86	5.83	3.31	4.25	6.27
AUG	10th	3.3	6.1	9.4	1.19	2.08	2.83	1.98	2.37	3.05
AUG	11th	2.3	4.3	6.2	1.14	1.95	2.61	2.11	2.34	2.98
AUG	12th	2.2	1.3	3	1.35	2.50	3.54	3.02	3.51	3.94
	Average	2.22	3.08	4.26	1.83	3.79	5.70	3.71	4.65	6.10

As introduced in Table 6.5, the average velocity simulated by CM-BEM is very close to the velocity observed, while WRF tends to overestimate velocity at all altitudes. This result is not surprising: after a series of vertical momentum flux transfer processes, the velocity bias is weakened due to the robustness of CM-BEM.

In the absence of energy consumption data, I am not able to approximate local energy consumption during that period of time. However, the simulation result proves that the new parametrization of mixing length together with CM-BEM is a strong enough numeric simulation method to simulate the momentum transferring process. As momentum diffusivity is the base of heat diffusivity, we can draw the conclusion safely that the simulation performance is reliable for a relatively longer simulation period.

H_{uw} is proved to be a great addition to mixing length approximation. For the district beyond H_{uw}, the atmospheric layers are slightly influenced by canopy elements and thus the diffusivity at high altitudes is similar to the area without any buildings. That is why the general solution given to the upper atmospheric layers is precise enough to have a great simulation result.

In the lower areas, shape factor Λ , derived from basic canopy geometric features, is capable of describing the influence of canopy elements. The effects of canopy elements are pressure-gradient force, in essence. Thus, the positive and negatives of Λ can describe whether the local canopy elements have a drag effect or accelerating effects on atmospheric layers inside the canopy area. The equation to calculate Λ is rooted in strict mathematical deduction and is based on N-S equations. Thus, Λ can be capable of explaining the physical phenomenon inside a one-dimension canopy and successfully simulate mixing length under both drag and accelerating effect.

Since H_{uw} and Λ are, actually, observed from LES-urban, I cannot utilize elementary mathematics to obtain analytical solutions. It is natural to turn to linear regressions to reveal the intrinsic relationship between these two factors and basic geometric features. To improve approximation performance, I employed MLP and have the bias between data observed and data predicted reduced.

CM-BEM simulation is conducted to evaluate simulation accuracy. Due to the fact that the meteorology simulation results are not accurate enough to refer to. There was a velocity bias generated in the early simulation process. However, as the mixing length is properly set at every altitude, the bias between velocity simulated and velocity observed diminished gradually. That proved the robustness of the new parametrization together with CM-BEM: even the noise exists, the simulation model can express momentum diffusivity and have a good result in the long run.

However, there are some topics left as future work. Kanda et al (2013) [26] put forward a proposal to parametrize competitive canopy features: roughness length and displacement height. As the wind goes through the rough ground and paves, numerous kinetic energies will be generated, and that these two factors are competitive factors to describe the flux near the ground. That can make a better solution to H_{uw} and Λ simulation. Also, since WRF is not that much accurate, finding a better meteorological data source may significantly improve simulation performance.

[1] National Aeronautics and Space Administration. Goddard Institute for Space Studies: GISS Surface Temperature Analysis (v4).
https://data.giss.nasa.gov/gistemp/graphs_v4/

[2] Malte Meinshausen, Nicolai Meinshausen, William Hare, Sarah C. B. Raper, Katja Frieler, Reto Knutti, David J. Frame & Myles R. Allen. Greenhouse-gas emission targets for limiting global warming to 2 ℃[J]. Nature,2009,458,30 : 1158–1162.

[3] Cohen B. Urbanization, City Growth, and the New United Nations Development Agenda[J]. The Official Journal of the World Coal Industry, 2015:4-7.

[4] (Japan Meteorological Agency) 日本の年平均気温偏差の経年変化（1898～2020 年）
https://www.data.jma.go.jp/cpdinfo/temp/an_jpn.html

[5] (Japan Meteorological Agency, 2013) ヒートアイランド現象の要因について http://www.data.jma.go.jp/cpdinfo/himr_faq/02/qa.html

[6] Office-and-ministries liaison conference related to countermeasures against urban heat islands. Outline of countermeasures against urban heat islands. 2004 [in Japanese].

[7] Ihara et al (2008): Changes in year-round air temperature and annual energy consumption in office building areas by urban heat-island countermeasures and energy-saving measures, Applied Energy, 85 (1), 25.

[8] Kikegawa Y, Genchi Y, Kondo H, Hanaki K. Impacts of city-block-scale countermeasures against urban heat-island phenomena upon a building's energy-consumption for air-conditioning. Appl Energ, 2006;83:649–68.

[9] Kondo H, Liu FH. A study on the urban thermal-environment obtained through one-dimensional urban canopy model. J Jpn Soc Atmos Environ, 1998;33(3):179–92 [in Japanese].

[10] Schubauer G B. Turbulent Processes as Observed in Boundary Layer and Pipe. J Appl Phys, 1954,35:188

[11] Kondo, J., Akashi, S. Numerical studies on the two-dimensional flow in horizontally homogeneous canopy layers. Boundary-Layer Meteorol 10, 255–272 (1976). https://doi.org/10.1007/BF00919389

[12] Monin AS, Yaglom AM (1971) Statistical fluid mechanics: mechanics of turbulence, vol I. The MIT Press,Cambridge, 769 pp

[13] Kondo, H., Inagaki, A. & Kanda, M. A New Parametrization of Mixing Length in an Urban Canopy Derived from a Large-Eddy Simulation Database for Tokyo. Boundary-Layer Meteorol 156, 131–144 (2015).

[14] Kikegawa Y, Genchi Y, Yoshikado H, & Kondo H. (2003). Development

of a numerical simulation system toward comprehensive assessments of urban warming countermeasures including their impacts upon the urban buildings' energy-demands. Applied Energy, 76(4), 449-466.

[15] Kondo H, Genchi Y, Kikegawa Y, Ohashi Y, Yoshikado H, Komiyama H (2005) Development of a multi-layer urban canopy model for the analysis of energy consumption in a big city: structure of the urban canopy model and its basic performance. Boundary-Layer Meteorol 116:395–421

[16] Maruyama, T.: 1991, 'Numerical Simulation of Turbulent Boundary Layer over Complicated Surfaces Such as Urban Areas', J. Wind Eng. 47, 81–81(in Japanese).

[17] Hagishima, A., Tanimoto, J. and Katayama, T.: 2001, 'Experiment Study on the Validity of an Urban Canopy Model in and Above the Canopy Layer', J. Archit. Plan. Environ. Eng. 548, 31–37(in Japanese).

[18] Uno, I., Ueda, H. and Wakamatsu, S.: 1989, 'Numerical Modeling of the Nocturnal Urban Boundary Layer', Boundary-Layer Meteorol. 49, 77–98.

[19] Gambo, K.: 1978, 'Notes on the Turbulence Closure Model for Atmospheric Boundary Layers', J. Meteorol. Soc. Jpn. 56, 466–480.

[20] Heinz Lettau: 1952, A Re-examination of the "Leipzig Wind Profile" Considering some Relations between Wind and Turbulence in the Frictional Layer, Tellus, 2, 125-129, 1950

[21] Watanabe T, Kondo J (1990) The influence of canopy structure and density upon the mixing length within and above vegetation. J Meteorol Soc Jpn 68:227–235

[22] Coceal O, Belcher SE (2004) A canopy model of mean winds through urban areas. Q J R Meteorol Soc 130:1349–1372

[23] Maruyama T (1993) Optimization of roughness parameters for staggered arrayed cubic blocks using experimental data. J Wind Eng Ind Aerodyn 46–47:165–171

[24] Blackadar AK (1962) The vertical distribution of wind and turbulent exchange in neutral atmosphere. J Geophys Res 67:3095–3102

[25] Mellor GL, Yamada T (1974) A hierarchy of turbulence closure models for planetary boundary layers. J Atmos Sci 31:1791–1806

[26] Kanda M, Inagaki A, Miyamoto T, Gryschka M, Raasch S (2013) A new aerodynamic parametrization for real urban surfaces. Boundary-Layer Metetorol 148:357–377

[27] Inagaki A, Castillo MC, Yamashita Y, Kanda M, Takimoto H (2012) Large eddy simulation study of coherent flow structures within a cubical canopy. Boundary-Layer Meteorol 142:207–222

[28] Raasch S, Schröter S (2001) A large-eddy simulation model performing on massively parallel computers. Meteorol Z 10:363–372

[29] Kanda M, Moriwaki R, Kasamatsu F (2004) Large eddy simulation of turbulent organized structure within and above explicitly resolved cube arrays. Boundary-Layer Meteorol 112:343–368

[30] Letzel MO, Krane M, Raasch S (2008) High resolution urban large-eddy simulation studies from street canyon to neighborhood scale. Atmos Environ 42:8770–

8784

[31] Letzel MO (2007) High resolution LES of turbulent flow around buildings. PhD dissertation, University of Hannover, Hannover, Germany, 126 pp

[32] Hattori Y, Moeng CH, Suto H, Tanaka N, Hirakuchi H (2010) Wind-tunnel experiment on logarithmic-layer turbulence under the influence of overlying detached eddies. Boundary-Layer Meteorol 134:269–283

[33] Inagaki A, Kanda M (2008) Turbulent flow similarity over an array of cubes in near-neutrally stratified atmospheric flow. J Fluid Mech 615:101–120

[34] Inagaki A, Kanda M (2010) Organized structure of active turbulence developed over an array of cube within the logarithmic layer of atmospheric flow. Boundary-Layer Meteorol 135:209–228

[35] Castillo MC, Inagaki A, Kanda M (2011) The effects of inner and outer layer turbulence of a convective boundary layer in the near-neutral inertial sublayer over an urban-like surface. Boundary-Layer Meteorol 140:453–469

[36] Cheng H, Castro IP (2002) Near wall flow over urban-like roughness. Boundary-Layer Meteorol104:229–259

[37] Cheng H, Hayden P, Robins AG, Castro IP (2007) Flow over cube arrays of different packing densities. J Wind Eng Ind Aerodyn 95:715–740

[38] Aston Zhang, Zachary C. Lipton, Mu Li, Alexander J. Smola(2007) Dive into Deep Learning

[39] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. The Journal of Machine Learning Research, 15(1), 1929–1958.

[40] (Bureau of Environment) 大気汚染測定結果ダウンロード
https://www.kankyo.metro.tokyo.lg.jp/air/air_pollution/torikumi/result_measurement.html?fbclid=IwAR31Bb0STHS-mPIVoXEozAkHR9-tmtUc8wE77EZraYb4gt2PzveixGp-Sxw

[41] Allen T. Chwang(2011): Viscous Fluid Mechanics(Second Edition)

Acknowledgments

I would like to express my deepest gratitude first and foremost to my supervisor, Associate Professor Ihara, who has walked me through the whole stage of my thesis writing and has offered me valuable suggestions in my academic studies. The thesis would not have been completed without his patient instruction and expert guidance.

Secondly, I would like to thank my sub-supervisor Professor Tonokura. He provides me excellent guidance and a great idea of my thesis that has complete my thesis inadequate.

I also would like to owe gratitude to all members of my laboratory and my friends, who also help me a lot in my study and daily life in Japan, which making me a great time for studying in Japan.

I would finally like to express my gratitude to my beloved parents who have always been supporting my decision without a word of complaint and helping me to walk through difficulties.