

Cambridge Working Papers in Economics

Cambridge Working Papers in Economics: 2054

FORECASTING URBAN RESIDENTIAL STOCK TURNOVER DYNAMICS USING SYSTEM DYNAMICS AND BAYESIAN MODEL AVERAGING

Wei	Eoghan	Alice	David M	Peter
Zhou	O'Neill	Moncaster	Reiner	Guthrie

16 June 2020

Knowing the size of the building stock is perhaps the most basic determinant in assessing energy use in buildings. However, official statistics on urban residential stock for many countries are piecemeal at best. Previous studies estimating stock size and energy use make various debateable methodological assumptions and only produce deterministic results. We present a Bayesian approach to characterise stock turnover dynamics and estimate stock size uncertainties, applied here to the case of China. Firstly, a probabilistic dynamic building stock turnover model is developed to describe the building aging and demolition process, governed by a hazard function specified by a parametric survival model. Secondly, using five candidate parametric survival models, the building stock turnover model is simulated through Markov Chain Monte Carlo to obtain posterior distributions of model-specific parameters, estimate marginal likelihood, and make predictions of stock size. Thirdly, Bayesian Model Averaging is applied to create a model ensemble that combines model-specific posterior predictive distributions of the recent historical stock evolution pathway in proportion to posterior model probabilities. Finally, the Bayesian Model Averaging model ensemble is extended to forecast future trajectories of residential stock development through 2100. The modelling results suggest that the total stock in China will peak around 2065, at between 42.4 and 50.1 billion m². This Bayesian modelling framework produces probability distributions of annual total stock, age-specific substocks, annual new buildings and annual demolition rates. This can support future analysis of policy trade-offs across embodied-versus-operational energy consumption, in the context of sector-wide decarbonisation.

Forecasting Urban Residential Stock Turnover Dynamics using System Dynamics and Bayesian Model Averaging

EPRG Working Paper 2016

Cambridge Working Paper in Economics 2054

Wei Zhou, Eoghan O'Neill, Alice Moncaster, David M Reiner and Peter Guthrie

Abstract Knowing the size of the building stock is perhaps the most basic determinant in assessing energy use in buildings. However, official statistics on urban residential stock for many countries are piecemeal at best. Previous studies estimating stock size and energy use make various debateable methodological assumptions and only produce deterministic results. We present a Bayesian approach to characterise stock turnover dynamics and estimate stock size uncertainties, applied here to the case of China. Firstly, a probabilistic dynamic building stock turnover model is developed to describe the building aging and demolition process, governed by a hazard function specified by a parametric survival model. Secondly, using five candidate parametric survival models, the building stock turnover model is simulated through Markov Chain Monte Carlo to obtain posterior distributions of model-specific parameters, estimate marginal likelihood, and make predictions of stock size. Thirdly, Bayesian Model Averaging is applied to create a model ensemble that combines model-specific posterior predictive distributions of the recent historical stock evolution pathway in proportion to posterior model probabilities. Finally, the Bayesian Model Averaging model ensemble is extended to forecast future trajectories of residential stock development through 2100. The modelling results suggest that the total stock in China will peak around 2065, at between 42.4 and 50.1 billion m². This Bayesian modelling framework produces probability distributions of annual total stock, age-specific substocks, annual new buildings and annual demolition rates. This can support future analysis of policy trade-offs across embodied-versus-operational energy consumption, in the context of sector-wide decarbonisation.

Keywords building stock; lifetime distribution; System Dynamics; Bayesian Model Averaging, Markov Chain Monte Carlo, embodied energy, operational energy; China

JEL Classification C11, O18, R21

Contact wz282@cam.ac.uk
Publication June 2020

Forecasting Urban Residential Stock Turnover Dynamics using System Dynamics and Bayesian Model Averaging

Wei Zhou ^{*1,2}, Eoghan O'Neill ^{1,3}, Alice Moncaster ^{2,4}, David M Reiner ^{1,5} and Peter Guthrie ²

June 2020

Highlights

- A System Dynamics model is developed to characterise building stock turnover
- Five survival models are used to represent Chinese building lifetime distribution
- Bayesian Model Averaging is used to quantify both parameter and model uncertainties
- Posterior predictive distributions of stock and age-specific substocks are obtained
- Chinese urban residential building stock is projected to peak around 2065

Abstract

Knowing the size of the building stock is perhaps the most basic determinant in assessing energy use in buildings. However, official statistics on urban residential stock for many countries are piecemeal at best. Previous studies estimating stock size and energy use make various debateable methodological assumptions and only produce deterministic results. This paper presents a Bayesian approach to characterise stock turnover dynamics and estimate stock size uncertainties, applied here to China. Firstly, a probabilistic dynamic building stock turnover model is developed to describe the building aging and demolition process, governed by a hazard function specified by a parametric survival model. Secondly, using five candidate parametric survival models, the building stock turnover model is simulated through Markov Chain Monte Carlo to obtain posterior distributions of model-specific parameters, estimate marginal likelihood, and make predictions of stock size. Thirdly, Bayesian Model Averaging is applied to create a model ensemble that combines model-specific posterior predictive distributions of the recent historical stock evolution pathway in proportion to posterior model probabilities. Finally, the Bayesian Model Averaging model ensemble is extended to forecast future trajectories of residential stock development through 2100. The modelling results suggest that the total stock in China will peak around 2065, at between 42.4 and 50.1 billion m². This Bayesian modelling framework produces probability distributions of annual total stock, age-specific substocks, annual new buildings and annual demolition rates. This can support future analysis of policy trade-offs across embodied-versus-operational energy consumption, in the context of sector-wide decarbonisation.

Keywords: building stock, lifetime distribution, System Dynamics, Bayesian Model Averaging, Markov Chain Monte Carlo, embodied energy, operational energy, China

* Corresponding author. Email: wz282@cam.ac.uk

¹ Energy Policy Research Group, University of Cambridge

² Department of Engineering, University of Cambridge

³ Faculty of Economics, University of Cambridge

⁴ School of Engineering and Innovation, Open University

⁵ Judge Business School, University of Cambridge

1 Introduction

Decarbonising the building sector is integral to global efforts in delivering the Paris Agreement commitments. Globally, over the period of 2010 to 2018, owing to continued efficiency improvements in building envelopes, systems and equipment, the building sector energy intensity per unit of floor area decreased by 11.9% [1]. However, this progress was not sufficient to offset the combined effect of improved access to energy, increased ownership of energy-consuming devices and rapid growth in building floor area. These factors collectively have been driving up the final energy demand from buildings. In 2018, the final energy consumption in buildings increased by 7% as compared to its 2010 level, amounting to 125 exajoules (EJ) and accounting for 36% of global final energy use [2]. The continuously increasing trend of energy consumption have rendered the building sector off-track from the pathway required to realise the IEA's Sustainable Development Scenario (SDS)[3]. The floor area of the building stock has a direct impact on the acceleration of building energy intensity reduction needed to bring buildings on track with the SDS. The target for building energy intensity will have significant implications for policies, technologies, investment and financing required to support the decarbonisation of both new and existing buildings. Therefore, it is essential to understand the current status of the existing building stock and to model possible trajectories of the stock's expansion and evolution over time. The lack of official statistics on historical floor area, especially in many developing countries [4], presents a critical challenge to this task.

China is a major driving force of the growth of global building sector. The total floor area of the new buildings constructed in China in 2018 was 2.5 billion m² [5], accounting for over a third (33.8%) of the global total of new buildings of 7.4 billion m² [6]. According to the World Energy Outlook 2019 [7], the final energy consumption by buildings in China, as the world's largest building energy consumer, was 504 million tonnes of oil equivalent, or 21.1 EJ, in 2018. In comparison, the buildings in the US, European Union and India respectively consumed 20.8, 18.2 and 9.1 EJ in the same year. China's share in final energy consumption by global buildings increased to 16.3% in 2018, from its 2010 level of 13.1%. Of this total energy consumption in Chinese buildings, urban residential buildings accounted for more than one third [8]. With growing urban population and higher demand for energy services in the built environment, urban residential buildings will have increasing strategic importance in China's efforts in decarbonising its building sector. Whilst the floor area is critical to the stock-level energy consumption, official statistics on total floor area of urban residential buildings in China only exist up to 2006 [9]. As a result, the historical growth trajectory of the Chinese urban residential stock from 2007 onwards is unknown. This creates a key barrier to forecasting possible future trajectories of the overall stock evolution and analysing stock-level energy consumption over the next several decades.

This paper models the evolution of the urban residential stock in China since 2007 and then uses the estimated current status as the starting point for forecasting future evolution of the building stock over the medium to long term. The rest of the paper is organised as follows. Section 2 presents a review of literature closely relating to modelling Chinese building stock, identifies major issues associated with the methodological approaches taken, and justifies the

relevance and novelty of the present study. Section 3 develops the dynamic model for stock turnover and explains the concepts of BMA and its application to building stock turnover dynamics; Section 4 presents key results and discuss potential model applications and policy implications. Finally, section 5 summarises the study and accentuates the model's applicability from an international perspective.

2 Literature review

Key to modelling the stock evolution is a turnover process driven by the dynamic interplay between: new construction, meeting incremental demand growth as a result of economic growth and rising living standards; existing buildings remaining in use but undergoing an ageing process; and old buildings, which are eventually demolished. In this context, building lifetime is a critical factor underlying the dynamic relationship between old and new buildings, as Chinese urban buildings are generally short-lived due to various factors, including quality of building materials, design standards, construction techniques and practices, maintenance and renovation, massive demolition as a result of rapid urbanisation and city rebuilding, etc. [10–14]. The short lifetime suggests a high turnover rate and great complexity and uncertainty associated with building stock characteristics, which have significant implications for stock-wide energy consumption and emissions over the medium to long term. Massive construction and demolition incur significant amounts of embodied energy for building materials production, transportation, construction, demolition and disposal [15]. In 2017, the total embodied energy of buildings in China was 520 million tonnes of coal equivalent, or 15.2 EJ, comparable to the total primary energy of 28.2 EJ used in building operation in China [16]. Meanwhile, a high turnover rate implies lower risk of operational energy and carbon lock-in [17,18], as the building stock is rapidly replenished with energy efficient buildings while suboptimal old buildings are removed. These two arguments suggest a trade-off from the perspective of whole-life building energy, i.e. embodied plus operational energy. This clearly demonstrates the importance of building lifetime and stock turnover, which shall be adequately understood.

Despite their fundamental impacts on the macro-level building energy with significant implications for China's climate targets, building lifetime and stock turnover appear to have been an under-researched area. There are substantially less studies in this area than the wide range of building energy related topics, such as building physics, building materials, energy efficiency, building integrated renewable energy, indoor air quality, thermal comfort, etc. One of the first Chinese building stock models was developed by Yang and Kohler [10]. Their model set the existing stock in 2005 as the initial stock composed of several age cohorts. A cohort-based approach was applied to define the average age of buildings to model the stock evolution on a five-year basis. Whilst no aging process of buildings was modelled, the influence of the probable building lifespans on mass flows and environmental impacts was pointed out. A similarly static approach was taken in the 3CSEP-HEB model [19], which assumed the Chinese building stock had an annual demolition rate of 0.5% and a retrofit rate of 1.4%. Taking a more dynamic perspective, Hu et al. [13] analysed the Chinese building stock by assuming a normal distribution function for building lifetime. Various scenarios of future demand for building materials such as steel and concrete for Chinese residential buildings were explored, both at the national level [20] and at the city level [21]. Shi et al [22] and Huang et al. [23] carried out similar studies investigating materials demand and

environmental impact of buildings and transport infrastructure stock in China, whose lifetimes were assumed to be normally distributed, with the average values in the range of 30 to 40 years. In Hong et al. [24], the same method was applied to develop a building stock model to project the trajectories of demand for building materials and the corresponding embodied energy. Both residential and commercial buildings were assumed to follow normal distributions in terms of their lifetimes, with the standard deviation being set to be one third of the average value. Similarly, investigating the impact of technical progress and the use of renewable energy in the building sector, Shi et al. [25] applied the China TIMES model and represented the lifetime distribution of buildings using a normal distribution. Referring to Hong et al. [24], the model developed by Huo et al. [26] also used a normal distribution for lifetime distribution of residential buildings in China to estimate building floor area, energy consumption and energy intensity.

Methodologically, the above assumptions on building lifetime and stock turnover have inherent limitations. Firstly, the use of a fixed annual demolition rate implies a mixing of the large volume of buildings of different ages in the stock. Referred to as cohort blending [27], this mixing means equal chance of buildings being demolished, regardless of their actual age and significant heterogeneity in their physical characteristics and socio-economic contexts. Secondly, whilst the use of a normal distribution attempts to explicitly account for lifetime uncertainties is comparatively reasonable, the critical issue is the questionable approach of specifying a normal distribution without calibration using empirical data. In previous studies, the mean, representing the average building lifetime, was assumed to take values in the range of 30 to 50 years, based on anecdotal evidence drawn from limited cases. The standard deviation was commonly set to be 30% of the mean, a largely arbitrary assumption. This approach resulted in inadequately substantiated shapes of the probability distributions that were used to describe the lifetime profile of buildings, thereby calling into question the estimated building stock turnover dynamics and the associated stock-level energy consumption and carbon emissions.

A few recent studies have partially addressed the above methodological issue. Cai et al. [14] calibrated the average building lifetime using official statistics on annual new construction and estimated annual demolition and investigated its impact on water withdrawal energy consumption and carbon emissions. Using the same method, Cao et al. [28] estimated the Chinese building lifetime, the size of the building stock, and the associated stock and flow of building materials, such as cement, steel, wood, brick, etc. Zhou et al. [29] used Weibull distribution to represent the lifetime distribution of urban residential buildings. In addition to calibrating parameters and estimating stock size, Zhou's aging chain modelling produced a detailed distribution of age-specific sub-stocks, i.e. a more disaggregated representation of the building stock than those of Cai et al and Cao et al. However, a common salient feature shared by these studies was the frequentist approach taken, which produced single point estimates of distribution parameters leading to a single profile of building lifetime. In this context, model parameters were treated as being deterministic and no uncertainty was considered. Moreover, the calibration of model parameters was conditional upon the model structure as chosen, without considering the uncertainty associated with the model per se, namely the probability of the chosen model being the true model given the observed empirical data. Neglecting the inherent uncertainties at the parameter level and model level, coupled

with having a limited amount of empirical data on historical floor area for calibration, is subject to high risk of undermining the robustness and reliability of the modelling results on building lifetime and stock turnover dynamics.

Evidently, the deficiencies of previous studies in handling uncertainties suggest a critical research gap to be addressed. To this end, this study innovatively develops a building stock turnover model in a Bayesian framework to estimate the stock evolution pathway of Chinese urban residential buildings in a probabilistic manner. The Bayesian approach treats the parameters of the probability distribution models chosen to describe building lifetime profile as random variables and derives posterior distributions of the parameters by taking account of both prior knowledge about parameter values and the likelihood of observing empirical data given certain parameter values. For a given probability distribution model, this presents a full picture of the likely parameter space, thus enabling a good understanding of the global shape of the distribution of the parameter. Such a distribution allows parameter uncertainties to be propagated through to the emergent behaviours of model outputs, such as total building stock size. Moreover, in addition to model-specific parameter uncertainties, a Bayesian approach allows model uncertainty to be estimated. Through Bayesian Model Averaging (BMA), predictions made by individual models are combined in proportion to posterior model probabilities. This means the creation of a BMA model ensemble, which involves a weighted average of the predictions from a number of models, with the weights being equal to the probabilities that the models are the true model given the observed data. BMA can avoid the situation where inferences based on an individual candidate model are overstated, and decision-making based on predictions is subject to higher risk than expected [30,31].

From a policy-making perspective, a probabilistic model offers the ability to generate probability distributions of different potential outcomes of policy scenarios. This is important in the context of analysing the decarbonisation of the generally short-lived Chinese buildings, where there is likely to be a strategic trade-off between operational and embodied energy due to factors such as massive construction and demolition, strengthening design codes for improved energy efficiency, scaled-up energy-related retrofits, technological advances, and so on. Taking a Bayesian approach, a probabilistic model incorporating building stock turnover, energy and carbon will enable future research into the probability that one policy, e.g., extending building lifetime to avoid embodied energy, would yield a more favourable outcome of stock-level decarbonisation as compared to another policy, e.g. accelerating stringency of new building design standards. Improving our understanding of these trade-offs is the overarching objective that motivates this study as an integral part of further research involving a fully-fledged building energy model.

Based on the above considerations, this study, as a first-of-its-kind attempt, applies BMA to develop a probabilistic dynamic model in order to estimate Chinese urban residential stock for the recent historical period of 2006 to 2017 and further forecast the stock development trajectories through 2100. The value of this study goes beyond filling the gap in quantifying the uncertainties in building stock turnover dynamics in Chinese context. It contributes an innovative methodological approach to the general field of building stock modelling. Its generality, flexibility and transparency make it potentially applicable to a wide variety of geographical contexts. It is particularly relevant and useful to countries experiencing rapid

urbanisation and massive construction, such as developing countries in South and Southeast Asia [32–34].

3 Methodology

3.1 Building stock turnover model

Estimating total building stock size requires understanding and modelling the stock turnover, which is characterised by the stock-level dynamics of construction of new buildings as inflow into the stock and demolition of old buildings as outflows from the stock. By the end of a year t , the total volume of demolition that year is the sum of all existing buildings constructed in all previous years that have reached the end of their lifetimes in year t . The building stock is composed of new buildings constructed in year t and those buildings which were previously constructed but have not reached the end of their lifetimes⁶.

Critical to the turn-over dynamics of building stock is building lifetime. Despite design lifetime required by building design regulations, often there is a lack of authoritative statistics relating to actual building lifetime data, particularly in developing countries. At a city or even country level, given the huge volume of buildings and significant heterogeneity in terms of their physical characteristics and socio-economic contexts, it is necessary to consider the uncertainties associated with building lifetime. It is unrealistic to assume that a cohort of buildings, i.e. those constructed in a given year, would be in service for exactly the same period and then demolished simultaneously. In the Chinese urban context, buildings are generally short-lived due to various factors, including quality of building materials, design standards, construction techniques and practices, maintenance and renovation, inappropriately accelerated demolition as a result of rapid urbanisation and city rebuilding, etc. [10,11]. While the degrees to which different factors play out are context-specific and time variant and therefore may differ significantly, the explicit direct outcome is a fast turn-over of building stock and therefore generally short lifetimes of the buildings thereof. Thus, building age can be used as a proxy variable to represent the impact collectively made by these underlying factors on demolition probability.

This paper proposes to apply the concept of survival analysis [35,36] to estimating building lifetime. It uses the probability density function (PDF) of a parametric survival model to approximate the likely lifetime distribution profile of a cohort of buildings built over a twelve-month period, so as to recognise and represent the uncertainties associated with factors collectively influencing lifetime of buildings. Thus, in a given year t , the proportion of demolished buildings in this cohort of buildings is modelled based on a hazard function. Conceptually, the hazard function represents the conditional probability that a building will

⁶ We acknowledge that a building may be disused functionally but still not demolished physically. Since the ultimate interest of modelling building stock turnover is in energy consumed by buildings, a functionally disused building does not consume energy anymore and therefore is considered equivalent to a physically demolished building from an energy perspective. Hence, in the rest of this study, demolition is used to refer to either physical demolition or functional disuse of buildings.

'expire' in that year t , provided that it has successfully survived to the previous year $t-1$. Mathematically, the hazard function is the ratio of the lifetime PDF to the complement of lifetime cumulative distribution function (CDF).

Applying the above concept, the total stock in year t consists of a series of substocks of different ages:

$$Stock_t = \sum_{j=t_0}^t substock_t[t-j] \quad (1)$$

Where $substock_t[t-j]$ represents buildings surviving in year t that are $(t-j)$ years old. For new buildings constructed in year t , they are 0 years old and therefore denoted by $substock_t[0]$.

The aging process undergone by any cohort of buildings is accompanied with annual demolition determined by age-specific hazard rates, $H(\text{age})$. Therefore, the annual total amount of demolition in year t is the sum of age-specific demolition of substocks at all ages.

$$Demolition_t = \sum_{j=t_0}^t H(t-j)substock_t[t-j] \quad (2)$$

For a $(t-j)$ -year-old substock in year t , its volume is determined by the aging process that it has undergone since it was constructed in year j .

$$substock_t[t-j] = \left[\prod_{k=0}^{t-j} (1 - H(k)) \right] substock_j[0] \quad (3)$$

Therefore, equation (1) can be re-written as:

$$\begin{aligned} Stock_t &= \sum_{j=t_0}^t substock_t[t-j] \\ &= \sum_{j=t_0}^t \left\{ \left[\prod_{k=0}^{t-j} (1 - H(k)) \right] substock_j[0] \right\} \end{aligned} \quad (4)$$

In above equation (4), the age-specific hazard rate $H(k)$ is determined by the parametric survival model chosen. Depending upon the specification, the hazard function of a survival model may or may not have a closed form expression.

The aging process described by the above equations can be visually represented by the model developed in Vensim, a commercial software for System Dynamics [37], as shown in Figure 1. As shown, a series of cascading sub-stocks of buildings form an aging chain, with each sub-stock representing a particular age group of buildings. The age group duration represents the time length that buildings in use reside in a sub-stock before shifting to the immediately next sub-stock in the chain. With age group duration set to be 1 year, the chronological aging process is discretized, i.e. each sub-stock represents buildings within a one-year age group. This level of granularity offers a detailed representation of substocks characterised by heterogeneity with respect to age (and energy-related properties, provided that additional layers are added to the model). This therefore enables separately tracking the aging process and experimenting with policy interventions targeting buildings of specific age groups. For example, given a large stock consisting of residential buildings constructed over the past 40 years, a policy-maker may want to know how buildings constructed in 2010 have been performing in terms of energy consumption from 2011 to 2019, what the building stock in 2019 looks like in terms of the composition of buildings of different ages and the corresponding energy performance, what would be the possible trajectories of stock-wide average energy intensity from 2020 to 2030 if an energy efficiency retrofit programme targeting buildings older than 20 years is implemented in 2020, and so on.

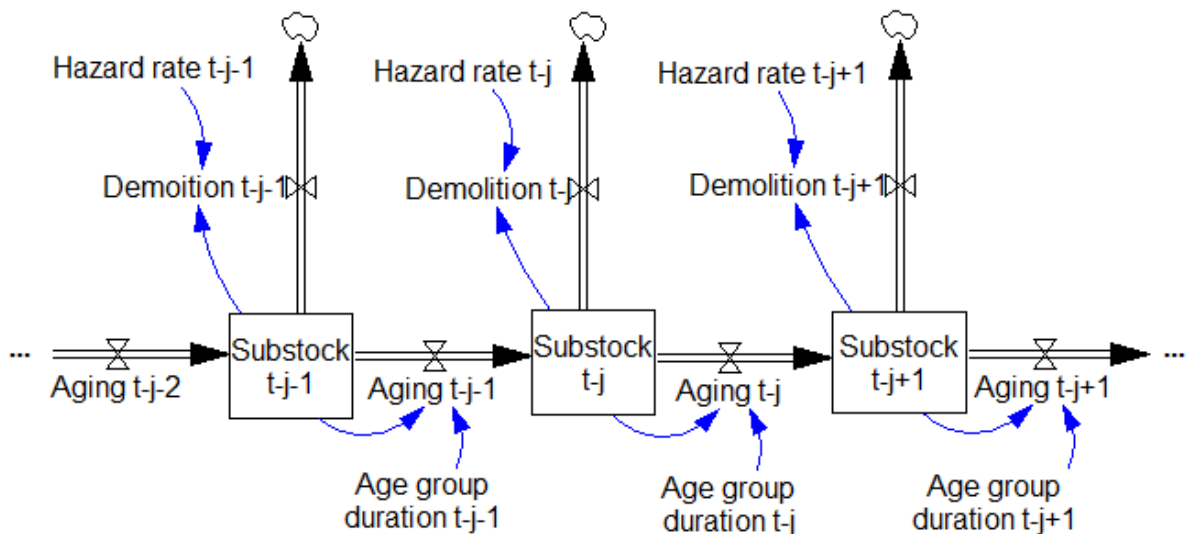


Figure 1: Aging chain with explicit modelling of sub-stock specific demolition

3.2 Bayesian modelling

3.2.1 Statistical model for historical stock

As described by equation (4), the deterministic component of the overall statistical model is the total building stock as the function of unknown parameters θ of a chosen parametric survival model, e.g. Weibull distribution, and the known annual new cohort of buildings

constructed over the historical period. This can be denoted by a function $f(\theta, t)$. The probabilistic component of the model is represented by an error term ε_t , which is assumed to be normally distributed with mean zero and unknown variance σ^2 , i.e. $\varepsilon_t \sim N(0, \sigma^2)$. $f(\theta, t)$ describes the expectation of modelled building stock. Therefore, in the Bayesian framework, the total stock can be described by the overall probabilistic model as follows:

$$Stock_t = f(\theta, t) + \varepsilon_t \quad (5)$$

3.2.2 Bayesian model inference

In the context of the statistical model, let D represent empirically observed data of total stock, y , and annual new buildings, x , for the period of 1978 to 2006, i.e. $D = \{(x_i, y_i), i = 1978, 1979, \dots, 2006\}$. According to Bayes' theorem, the posterior probability density $p(\theta|D)$, given the data D , is calculated as follows:

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} = \frac{p(D|\theta)p(\theta)}{\int p(D|\theta)p(\theta)d\theta} \quad (6)$$

where $p(\theta)$ is the prior distribution of θ , representing subjective prior knowledge about θ . $p(D|\theta)$ is the likelihood function, which can be viewed as a function of θ given the empirically observed data D which is considered fixed. It represents the likelihood that the given set of empirically observed data D is explained by the model with possible parameter values. $p(D)$ is the marginal likelihood, which is an integration of $p(D|\theta)$ over all possible values of θ across its space and therefore is not a function of θ , but a constant. This proportionality constant plays the role of normalizing the posterior density to ensure it integrates to 1. $p(D)$ is also known as model evidence, because it provides evidence for a candidate model, which is critical in selecting and averaging models as discussed later.

The posterior distribution $p(\theta|D)$ fully describes the uncertainty associated with the parameters. Essentially it updates the prior knowledge about the parameters in light of the empirical data. Generally, it is difficult or not possible to analytically express the posterior distribution. The solution is to instead simulate sample draws from the posterior distribution, such that the values of these samples are distributed approximately according to the posterior distribution of the parameters of interest. The samples enable calculation of point estimates of the parameters, such as mean, median, or mode. More importantly, the samples of parameters enable drawing samples from predictive distributions associated with model outputs, e.g. the annual total building stock as the high-level emergent behavior of the dynamic building stock model, thus facilitating policy scenario analysis. Methodologically, this is realised using Markov chain Monte Carlo (MCMC) algorithm, as introduced later.

3.2.3 Posterior predictive distribution

With the posterior distribution $p(\theta|D)$, it is possible to make inferences about the total stock for a given year during the period of 2007 to 2017, an unknown observable denoted as \tilde{y} ,

given the known annual new buildings for the same year, denoted as x^* . This leads to the posterior predictive distribution of \tilde{y} :

$$p(\tilde{y}|x^*, D) = \int p(\tilde{y}|x^*, \theta)p(\theta|D)d\theta \quad (7)$$

This equation suggests that the posterior predictive distribution is derived by marginalising the likelihood function $p(\tilde{y}|x^*, \theta)$ over the entire set of parameters, with each point in the space of parameters weighted according to its posterior probability given the empirically observed data.

3.2.4 Bayesian Model Averaging

The above posterior predictive distribution is conditional upon a choice of model M , i.e. a building stock model employing a particular parametric survival model, e.g. Weibull distribution. The equation can be written more explicitly as:

$$p(\tilde{y}|x^*, M, D) = \int p(\tilde{y}|x^*, \theta, M)p(\theta|D, M)d\theta \quad (8)$$

There are multiple choices of parametric survival model, each of which may characterise the dynamics of building stock turnover. Candidates include Weibull, Lognormal, Gamma, etc. Let M_k denote a building stock turnover model using a plausible survival model k specified by parameter vector θ_k , and let $M = \{M_1, M_2, \dots, M_k\}$ denote the model space under consideration. This creates a model ensemble, which, when making predictions, takes into account the uncertainties associated with not only model-specific parameters but also the models per se. Now, the posterior predictive distribution of total building stock for the period of 2007 to 2017, \tilde{y} , is calculated as:

$$p(\tilde{y}|x^*, D) = \sum_{k=1}^K p(\tilde{y}|x^*, M_k, D)p(M_k|D) \quad (9)$$

Where $p(\tilde{y}|x^*, M_k, D)$ is the posterior predictive distribution under model M_k given data D , and $p(M_k|D)$ is the posterior model probability (*PMP*), which is also referred to as model weight. Hence, the posterior distribution of \tilde{y} predicted by the model ensemble, $p(\tilde{y}|x^*, D)$, is effectively the average of the posterior predictive distribution under each of the candidate models in the model space, weighted by their respective *PMPs*.

The *PMP* of model M_k can be interpreted as the probability of model M_k being the true model predicting \tilde{y} , given the observed data D , thus reflecting the extent to which M_k fits the observations as compared to other candidate models in the model space. *PMP* is given by:

$$p(M_k|D) = \frac{p(D|M_k)p(M_k)}{\sum_{j=1}^K p(D|M_j)p(M_j)} \quad (10)$$

Where $p(M_k)$ is the prior probability of model M_k being the true model, allowing the existing prior knowledge about the plausibility of model M_k to be specified explicitly, and $p(D|M_k)$ is the marginal likelihood (or model evidence) of model M_k , which is given by

$$p(D|M_k) = \int p(D|\theta_k, M_k)p(\theta_k|M_k)d\theta_k \quad (11)$$

Where $p(D|\theta_k, M_k)$ is the likelihood of model M_k given observed data D , and $p(\theta_k|M_k)$ is the prior probability density of the parameters θ_k under model M_k . In fact, $p(D|M_k)$ is the denominator in the above equation (6) for calculating the posterior probability density of parameters θ_k under model M_k , as given by

$$p(\theta_k|D, M_k) = \frac{p(D|\theta_k, M_k)p(\theta_k|M_k)}{\int p(D|\theta_k, M_k)p(\theta_k|M_k)d\theta_k} = \frac{p(D|\theta_k, M_k)p(\theta_k|M_k)}{p(D|M_k)} \quad (12)$$

Compared with equation (6), the above equation (12) explicitly applies subscript k to reflect that both the priors of model-specific parameters θ_k and the likelihood function of the observed data D are conditional on the particular model M_k in the model space.

Based on the above, the posterior mean of \tilde{y} , as predicted by the model ensemble, can be calculated as follows:

$$E[\tilde{y}|x^*, D] = \sum_{k=1}^K E[\tilde{y}|x^*, M_k, D]p(M_k|D) \quad (13)$$

Clearly the BMA model ensemble prediction is essentially the average of individual predictions weighted by the probability that an individual candidate model is true given the observed data. BMA model ensemble leads to a more spread posterior distribution of y than an individual candidate model does. This avoids the situation where inferences made based on an individual candidate model are overstated and decision-making based on predictions is much riskier than expected [30,31,38,39].

3.2.5 Model space

In general, a range of parametric survival distribution functions are available to describe the survival process in various fields [36,40,41]. However, literature on survival analysis or lifetime data analysis on buildings is limited. A survey on buildings in the Netherlands found that empirical survival probabilities of buildings were well approximated by Weibull distribution (OECD 2009). Miatto *et al.* [44] tested various PDFs and found that the lognormal distribution offered the best fit to lifespans of buildings in Nagoya and Wakayama, Japan, where buildings

were short-lived, with average lifespans shorter than 30 years. Zhou *et al.* [29] applied the Weibull distribution to approximate lifetime uncertainties of Chinese urban residential buildings. From an economic perspective, buildings can be regarded as a type of capital asset, and accordingly building stock can be regarded as a type of capital asset, and accordingly building stock can be regarded as capital stock [42,43]. Hence, a range of PDFs that have been used as a proxy for service lives and retirement/discard patterns of capital stocks may be applied to buildings, such as log-normal, Weibull, Gamma, and so on [42,43,45–49].

In this paper, the distribution functions used for approximating the lifetime distribution of Chinese urban residential buildings are Weibull, Lognormal, Loglogistic, Gamma and Gumbel distributions. Each distribution can characterise the turnover dynamics of the building stock, thereby representing a candidate model M_k in the model space M . The PDFs of these distributions are given in Table 1. Specifying the PDF of a distribution allows the CDF, survival function and hazard function of the distribution to be ascertained.

Table 1: Five candidate survival distribution functions

Distribution	Probability density function	Parameters	Priors
Weibull	$f(x) = \left(\frac{\alpha x^{\alpha-1}}{\lambda^\alpha} \right) e^{-(\frac{x}{\lambda})^\alpha}$	Shape $\alpha > 0$ Scale $\lambda > 0$	$\alpha \sim \text{uniform}(1,10)$ $\lambda \sim \text{uniform}(1,100)$
Lognormal	$f(x) = \frac{1}{x\sqrt{2\pi}\sigma'} e^{-\frac{1}{2}\left[\frac{\ln x - \mu'}{\sigma'}\right]^2}$ $\mu' = \ln \left[\frac{\mu^2}{\sqrt{\sigma^2 + \mu^2}} \right],$ $\sigma' = \sqrt{\ln \left[1 + \left(\frac{\sigma}{\mu} \right)^2 \right]}$	Mean $\mu > 0$ Standard deviation $\sigma > 0$	$\mu \sim \text{uniform}(1,100)$ $\sigma \sim \text{uniform}(1,100)$
Loglogistic	$f(x) = \frac{e^{\frac{\ln(x)-\mu}{\sigma}}}{\sigma x (1 + e^{\frac{\ln(x)-\mu}{\sigma}})^2}$	Scale $\mu > 0$ Shape $\sigma > 0$	$\mu \sim \text{uniform}(1,100)$ $\sigma \sim \text{uniform}(1,100)$
Gamma	$f(x) = \frac{1}{\lambda\Gamma(\alpha)} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-\frac{x}{\lambda}}$	Scale $\lambda > 0$ Shape $\alpha > 0$	$\lambda \sim \text{uniform}(1,100)$ $\sigma \sim \text{uniform}(1,100)$
Gumbel	$f(x) = \frac{1}{\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)} e^{-e^{-\left(\frac{x-\mu}{\sigma}\right)}}$	Scale $\mu > 0$ Shape $\sigma > 0$	$\mu \sim \text{uniform}(1,100)$ $\sigma \sim \text{uniform}(1,100)$

3.2.6 Model priors and model parameter priors

Prior probabilities of models reflect the prior knowledge, or belief, that a specific model is the true model in the domain concerned. Eliciting an appropriate prior is a non-trivial task in any Bayesian setting, and such difficulties are compounded in BMA because a probability measure for the model space, which is a more abstract parametric space, is not obvious in principle [31].

Whilst informative priors are expected to benefit model development and improve predictive performance, often non-informative priors have to be used due to little prior knowledge about the relative plausibility of the models considered. As a simple but reasonable neutral choice, it can be assumed that all candidate models in the model space are equally likely *a priori* [30,38]. This means applying a uniform distribution over the model space, so that $p(M_j) = \frac{1}{K}$, for $j = 1, 2, \dots, K$. No model is considered more likely *a priori* than any other one. The consideration is to let the observed data carry all the information. This is the most commonly adopted practice in defining model priors in BMA settings [31]. On this basis, the aforementioned five distributions of this study are assumed to have the equal prior probability equal to 0.2. This leads to the prior model probabilities being cancelled out and the *PMP* of a candidate model being proportional to its model evidence, namely, marginal likelihood.

The same consideration is applied to defining prior distributions of model-specific parameters. For any of the five candidate models, there is little prior information about its model-specific parameters. Hence it is straightforward to specify non-informative priors so as to allow the posteriors to be informed by data. As shown in Table 1, the priors of the model-specific parameters are all assumed to be uniformly distributed over their reasonable ranges in the context of generally short lifetimes of urban residential buildings in China.

3.2.7 MCMC sampling and posterior distribution calculation

MCMC is used to simulate the posterior distribution of a model-specific parameters. The principle is to draw values of a parameter vector θ from approximate distribution and then correct those draws to better approximate the target posterior distribution. Sampling is performed iteratively in such a way that at each step of the process it is expected that draws are made from a distribution that becomes closer to the target posterior distribution [50]. The sampling process is sequential and the draws create an ergodic Markov chain, which, after a large number of iteration steps, evolves through the parameter space, becomes stationary and converges to the target posterior distribution. Subsequent model inference can be made based on samples from this process much as based on samples from the target posterior distribution [51].

This study uses the Metropolis-Hastings algorithm, which is well established amongst available MCMC algorithms. At the start of iteration t , a candidate vector θ^* is generated from $\theta^{(t-1)}$ through a proposal distribution $f(\theta^*|\theta^{(t-1)})$, which is also known as a jumping distribution. The probability of θ^* being accepted to become $\theta^{(t)}$ is:

$$r = \min \left\{ 1, \frac{p(\theta^* | \text{data}) / f(\theta^* | \theta^{(t-1)})}{p(\theta^{(t-1)} | \text{data}) / f(\theta^{(t-1)} | \theta^*)} \right\} \quad (14)$$

The acceptance probability r means that if the result is higher than 1, r is set to 1, the candidate θ^* is accepted and the transition from θ^* to θ^t is made. Otherwise, if the result is lower than 1, the candidate θ^* is accepted with probability equal to r and rejected with probability equal to $1-r$. When accepted, the transition from θ^* to $\theta^{(t)}$ is made. When rejected, no move at iteration t is made, hence $\theta^{(t)} = \theta^{(t-1)}$, meaning that the chain is updated using the current value.

The proposal distribution $f(\cdot)$ is chosen to be a random walk proposal, where θ^* is selected by taking a random perturbation ε around the current value $\theta^{(t)}$, i.e. $\theta^* = \theta^{(t)} + \varepsilon$. The random vector ε is drawn independently of $\theta^{(t)}$ and centered on zero. As a common setting, ε is a normal distribution with mean zero and variance set to obtain efficient jumping algorithm [50,52]. In this regard, this study tunes the algorithm by using adaptive sampling, which generates new candidate parameters with a proposal covariance matrix that is estimated from the covariance matrix of the parameters generated so far, with a scaling factor of $2.4^2/d$, where d is the number of parameters [53,54].

3.2.8 Marginal likelihood calculation

Generally, the marginal likelihood is not analytically tractable and therefore has to be approximated using numerical methods. Typical Monte Carlo sampling methods include naïve Monte Carlo, Importance Sampling (IS), Harmonic Mean (HM), Generalised HM, and Bridge Sampling. The Naïve Monte Carlo is straightforward and in principle unbiased, but numerically inefficient and unstable if the posterior distribution is peaked relative to the prior method [31,55,56]. IS may overcome these issues by having an importance density with fatter tails than the posterior distribution [56,57]. HM uses the posterior distribution as the importance density. This results in the marginal likelihood being equal to the posterior harmonic mean of the likelihood. Despite its convenience and popularity, HM has been criticised extensively due to numerical instabilities and overestimation of the marginal likelihood [55,58,59]. Generalised HM, a more stable version of HM, can be viewed as the reciprocal IS [60]. Thus, for the reason analogous to IS, this method also requires the importance density to be finetuned to avoid unbounded variance. Specifically, it requires importance density to have thinner tails than the posterior distribution [31,56,61]. Bridge Sampling is a general case of the afore-mentioned methods. Compared to IS and Generalised HM, it is more robust to tail behaviours of the proposal distribution (conceptually similar to importance density) relative to posterior distribution and thus avoids large or even infinite variances of estimators [56,60,62,63]. This study uses Bridge Sampling to approximate the marginal likelihood of each of the five candidate models.

3.3 Forecasting future building stock turnover

Conceptually, the possible building stock evolution trajectories in the future are assumed to be subject to the same turnover process as described in Section 2.1, i.e. a dynamic interplay between construction of new buildings, aging of existing buildings in the stock, and demolition of old buildings reaching the end of their lifetimes. The key difference is that annual construction is an exogenous variable in the stock turnover model when predicting the recent stock evolution from 2007 to 2017, whereas it has to be taken as an endogenous variable in the model when forecasting future stock evolution, since future construction is unknown. Methodologically, since the annual construction of new buildings, existing stock and annual demolition are related to each other, there is a need for some exogenous variables to be included in the model to play the role of driving the evolution of the stock. Here the future trajectories of urban population and per capita floor area (PCFA) are used to establish expected demands of total urban residential building stock, which will be met by the net effect of construction and demolition. The mathematical relationships are given by the equations below, which, together with the equations in Section 2.1, formulate the mechanism driving the future stock evolution.

$$ExpectedTotalStock_t = UrbanPopulation_t \cdot PCFA_t \quad (15)$$

$$NewConstruction_t = (ExpectedTotalStock_t - Stock_t) + Demolition_t \quad (16)$$

where the new construction in year t is 0 years old and therefore equivalent to $substock_t[0]$, which is the youngest substock amongst all substocks in $Stock_t$.

The first of the two underlying driving factors, urban population, is determined in turn by total population and urbanisation rate. The historical data and future projection of China's total population are sourced from the official statistics of the Chinese government [64] and the World Population Prospects by the United Nations [65]. As for the urbanisation rate, China has been experiencing consistently rapid urbanisation over the past few decades [64] and such a trend is anticipated to continue in the future [66,67]. The projection by World Urbanization Prospects 2014 indicated that China would reach a urbanisation rate of 68.7% by 2030, 72.8% by 2040 and 75.8% by 2050 [68]. A joint research by the World Bank and the Development Research Centre of China's State Council pointed out that the inflection of China's urbanization rate, namely the highest annual rate of urbanization rate change, already occurred in 2008 and the rate would surpass 62% in 2020, 70% in 2030 and 76% in 2050 respectively [69]. It is therefore considered reasonable to project the future development trend of urbanisation rate based on historical trajectory and the expected saturation level over the long term. The generally S-shaped curve of the overall trend justifies the use of a logistic growth model.

$$Urbanisation_t = \frac{Urbanisation_{Peak}}{1 + e^{a+b(t-BaseYear)}} + \varepsilon_t \quad (17)$$

In the above equation, BaseYear is taken as 1978, the milestone year marking the beginning of nation-wide reform and opening-up policy. Urbanisation_{Peak} is the upper asymptote, representing the expected saturation level of urbanisation over the long term. ε_t is the error term, which is assumed to be normally distributed with mean zero and unknown variance σ^2 , i.e. $\varepsilon_t \sim N(0, \sigma^2)$. In addition, a and b are parameters determining the shape of the Logistic curve. Using urbanisation rate data available from the official statistics of Chinese government [64] and the World Urbanisation Prospects 2014 [68], the vector of unknown parameters of the urbanisation model, including Urbanisation_{Peak}, a, b and σ , can be estimated through MCMC.

The other driving factor, PCFA, is modelled by applying a Gompertz growth model. Both Gompertz and Logistic belong to the Richards family of sigmoidal growth models. Different from Logistic function, the curve of Gompertz function is not symmetric about its point of inflection, which is reached early in the growth trend. The future value asymptote on the right hand is approached much more gradually by the curve than the lower value asymptote on the left side [70–72]. Both Logistic and Gompertz models have been used to model PCFA [4,23,73–75]. Li and Xu [73] and Xu and Liu [75] argued that a Logistic model is more suitable for describing growth at economy take-off stage whereas a Gompertz, showing a more apparent increasing trend at a later stage, can better reflect the development of rising demand for urban housing along with the increase of per capita income. This study uses a Gompertz model, which can be defined as follows:

$$PCFA_t = PCFA_{peak} e^{-e^{a(t-BaseYear)+b}} + \varepsilon_t \quad (18)$$

where PCFA_{peak} is the expected saturation level of PCFA over the long term; a and b are parameters determining the shape of the Gompertz curve; and $\varepsilon_t \sim N(0, \sigma^2)$ is the error term.

In estimating the vector of parameters of the Gompertz model, a challenge is the lack of reliable PCFA data. The historical PCFA data published by various official sources have been found to be over-estimated due to sampling representativeness issues [73,76–79]. This research takes an alternative approach. Samples are drawn from the posterior predictive distribution of the total stock for the period of 2006 to 2017 and each sample, comprising 12 data points, is divided by official data on annual urban population to get a sample of PCFA (12 data points) of the same period. This way, a number of samples of PCFA are obtained. In addition, to increase the number of data points and also to take into account the importance of possible future policy targets, additional data points for future years are added to each sample of PCFA. Literature suggests that there is a considerably large variation in possible future PCFA levels, ranging from 24 to 60 m² per capita. Peng, Yan and Jiang [80] indicated that residential building floor area in urban China should remain at 24m² per capita in order to control total building energy use. The Energy Research Institute (ERI) under China's National Development and Research Commission predicted that total floor area of urban residential buildings in China would reach 43.953 billion m² by 2030, corresponding to 42.7 m² per capita,

a level similar to those in developed countries in 2007 [81]. IEA's Energy Technology Perspectives 2015 forecasted that residential floor area in China would reach 52m² per capita in the timeframe of 2030 to 2050 [82]. A joint report by IEA and THUBERC estimated that average floor area per person could increase to 58m² by 2050, although this level was considered unrealistic given high density levels in China [83]. THUBERC [79] argued that urban residential floor area over the medium to long term should not exceed 35 m², thereby curbing total floor area of urban residential buildings below 35 billion m². According to a recent study on China's deep decarbonization led by China's National Centre for Climate Change Strategy and International Cooperation [84], residential floor area per capita in China should be kept at the level of approximately 36 to 37 m² from 2030 to 2050. Such a large variation suggests high uncertainties associated with PCFA and therefore justifies the forecast of future trajectories in a Bayesian framework. Taking these estimates from the existing literature into consideration, four additional data points reflecting possible future PCFA levels, 35m² for 2030, 40m² for 2040, 45m² for 2050 and 50m² for 2080, are added to the 12 data points of each PCFA sample to estimate the posterior distributions of the four parameters of Gompertz model (PCFA_{Peak}, a, b and σ) through MCMC. Samples are drawn from the results from each MCMC, which is based on each sample of PCFA comprising 16 data points, and then are combined to obtain the posterior distributions of the four parameters.

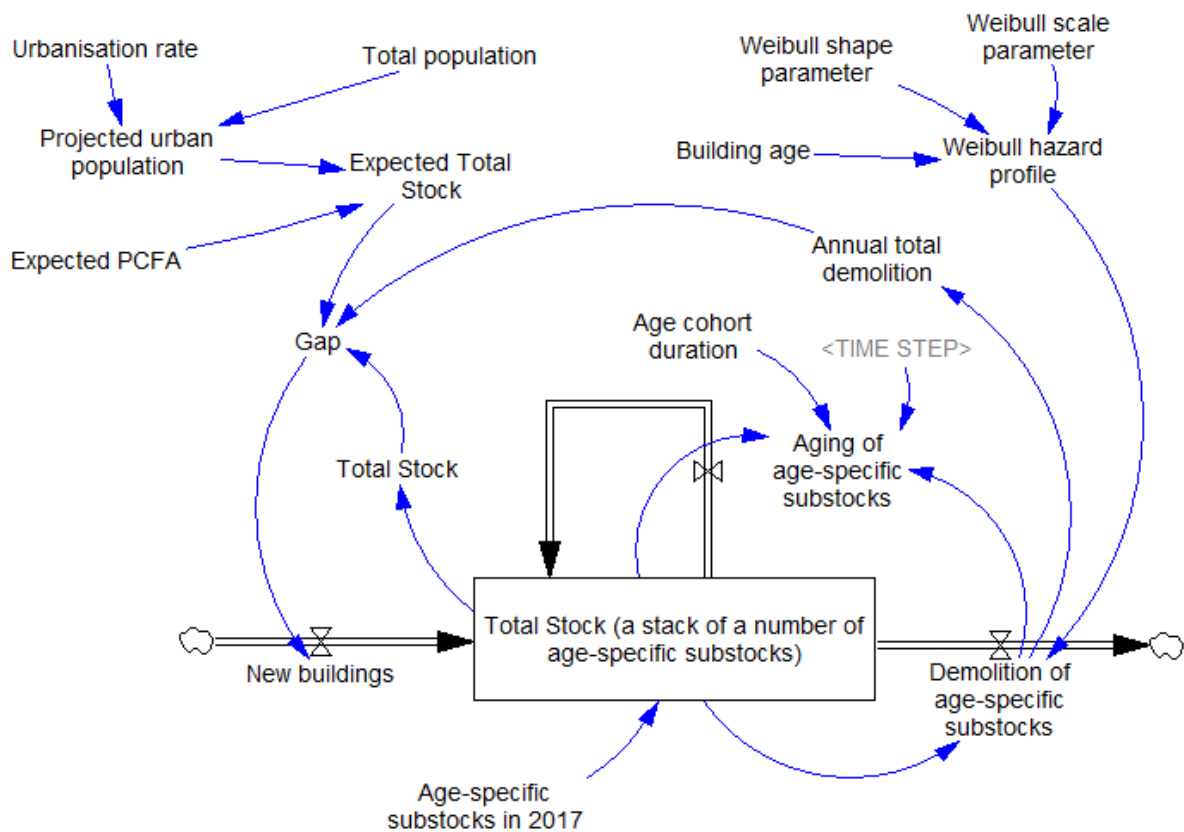


Figure 2: Extended building stock turnover model using Weibull distribution

A graphical representation of the above described future stock dynamic turnover model is shown in Figure 2. It should be noted that the model structure in Figure 1 is reformulated to improve representation and analytical convenience as shown in Figure 2. The building stock

is a stack of a number of age-specific substocks that are implicitly represented. Each substock is subject to an age-specific demolition rate determined by the hazard profile derived from each of the five candidate parametric survival models. In this figure, a Weibull distribution is used as the survival model for illustration purpose. It should also be noted that each run of the model is based on a particular combined set of samples drawn from the posterior distributions of survival model parameters, the posterior distributions of the parameters of the Gompertz model for expected PCFA, and the posterior distributions of the parameters of the Logistic model for urbanisation rate.

4 Results and discussion

4.1 Posterior model probabilities (PMPs)

Based on the methodology elucidated above, the posterior distributions of model-specific parameters of each candidate model, $p(\theta_k|D, M_k)$, were obtained using official statistics on total stock of urban residential buildings up to 2006. The primary data sources included China Statistical Yearbook and MOHURD's Statistical Communique on Urban Housing. Then, the evidence of each candidate model, i.e. the marginal likelihood, was numerically estimated using bridge sampling technique, and the PMP was calculated (Table 2).

Table 2: Prior and posterior probabilities of the five candidate models

Model	Prior	PMP
Weibull	0.2	0.219
Lognormal	0.2	0.25
Loglogistic	0.2	0.096
Gumbel	0.2	0.42
Gamma	0.2	0.015

4.2 Prediction of historical stock

With each candidate model, the posterior predictive distribution of total stock over the period of 2007 to 2017, \tilde{y} , was obtained through running the probabilistic stock turnover model using the posterior distributions of model-specific parameters, i.e. $p(\theta_k|D, M_k)$, and official statistics on annual new construction from 2007 to 2017. The posterior distribution of \tilde{y} predicted by the BMA model ensemble is the PMP-weighted average of the posterior predictive distribution of \tilde{y} under each candidate model in the model space. Operationally this was obtained by drawing samples from model-specific predictions with probabilities equal to the PMPs and then combining the samples. Figure 3 shows the 95% credible interval of posterior prediction of total stock by the BMA model ensemble. As expected, the total stock size was characterised by a continuously ascending pattern over time. The mean of the credible interval increased by 33% over eight years from 17.7 billion m² in 2010 to 23.6 billion m² in 2017. Clearly the line representing the mean of credible interval exhibits a good fit with the estimate by the Annual Report on China Building Energy Efficiency [16], which was developed by Tsinghua University

Building Energy Research Centre (THUBERC) and is widely recognised as an authoritative report on the overall situation of building energy in China.

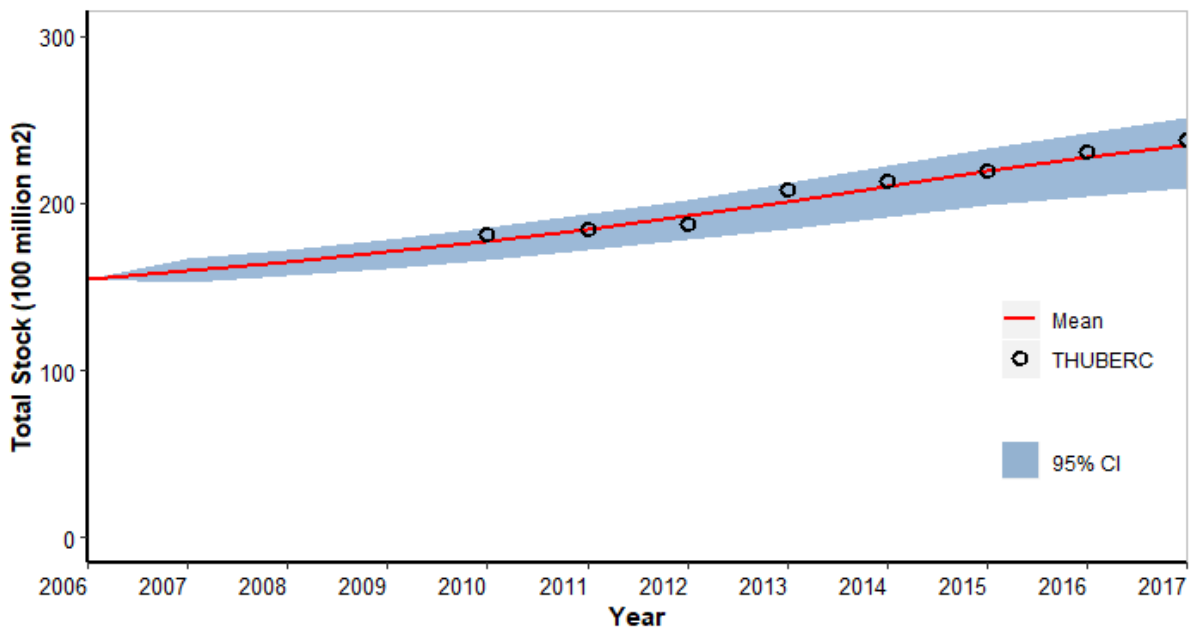


Figure 3: 95% Credible interval of BMA ensemble’s posterior prediction of total building stock

Due to the lack of official statistics on annual demolition, it is not possible to directly cross-validate the modelling result using historical annual demolition data. In an indirect way, though, comparing the demolition estimated in this study with previous studies provides an alternative basis for evaluating the robustness of the modelling approach and results. According to THUBERC [85], the ratio of aggregated demolished buildings to aggregated newly constructed buildings over China’s 11th Five-Year Plan Period (2006 to 2010) was approximately 34%. In this study, using the mean of the posterior predictive distribution of the aggregated demolished buildings modelled over this period, this ratio is calculated to be 32%, very close to the THUBERC [85] estimate. In absolute terms, the mean of posterior predictive distribution of annual demolition of this study is of the same order of magnitude as previous studies. For example, for 2010, the annual demolition was estimated by this study to be 1.47 billion m², approximately 1.3 billion m² by [13], and approximately 1.7 billion m² by [23].

Compared with a single point, deterministic estimate of annual stock size, the BMA approach taken by this study produces a profile for annual stock size, i.e. the posterior predictive distribution (Figure 4). This probabilistic estimate of annual stock size captures both models’ and the model-specific parameters’ uncertainties. Having depicted all possible pathways of stock evolution, it provides a full distribution of existing stock size per year and therefore helps to improve the reliability and robustness of not only the estimate of existing stock, but also the forecasting of future total stock which, as explained in previous sections, is a function of the existing stock, the underlying survival models and parameters, and expected future demand for housing area. The results of forecasting future stock are given in the next section.

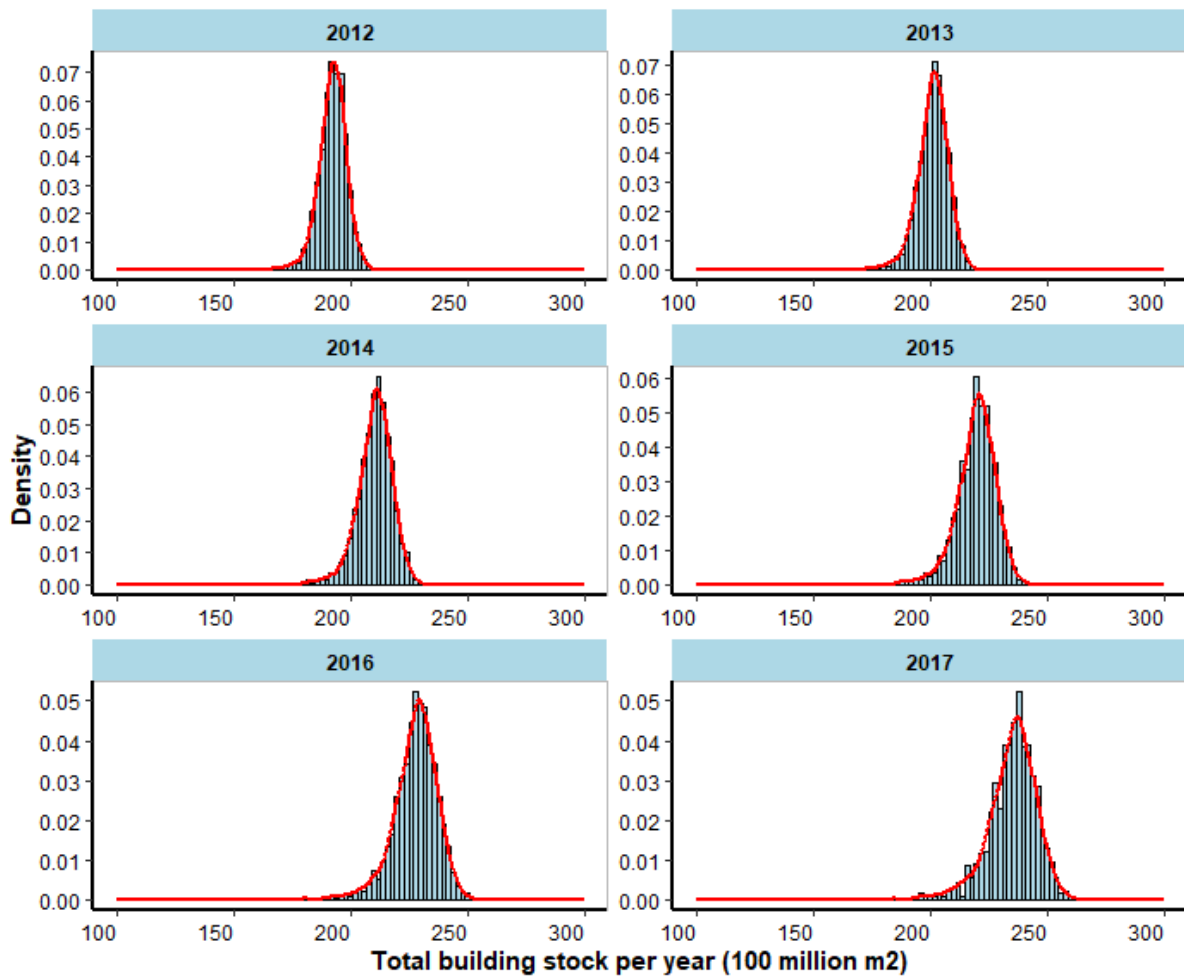


Figure 4: Posterior predictive distribution of total stock (2012-2017)

4.3 Forecast of future stock

Our forecasts of future stock turnover dynamics are based on posterior predictive distribution of total stock in 2017 and the interplay between new construction, demolition, existing stock and expected stock. Operationally, the System-Dynamics-based future stock turnover model has five versions, each of which is based on one of the five candidate survival models (Weibull, Lognormal, Loglogistic, Gumbel, Gamma). For each version, such as Weibull, a large number of combined samples are drawn from the posterior distributions of survival model-specific parameters, posterior distributions of vector of parameters of PCFA model, and posterior distributions of vector of parameters of urbanisation model. Using each sample of combined parameters, that particular version of stock turnover model is run to obtain a distinct trajectory of future stock evolution up to 2100. The trajectory reflects the stock's underlying composition of age-specific substocks in each year. Upon completing running that particular version of stock turnover model using all the samples of combined parameters, the family of trajectories of future stock evolution together produce the distribution of the total stock size per year

corresponding to that particular version of the stock turnover model. By repeating this process for each version of the stock turnover model and then combining the version-specific family of trajectories in proportion to posterior model probabilities, the distribution of the total stock size corresponding to the BMA model ensemble is obtained.

As shown in Figure 5, driven by rising urbanisation rate and expected demand for PCFA, the total stock steadily increases over the next two decades. From around 2040 onwards, while the total stock remains the general ascending trend, its growth starts to decelerate as the inflection point is being approached. The peak of the curve occurs around 2065, when the mean of the distribution of total stock size reaches 46.3 billion m² and the 95% credible interval is in the range of 42.4 to 50.1 billion m². After 2065, the total stock starts to shrink slowly. This is due to the projected decrease of total population around the same period, the effect of which outweighs the continuous (but decelerating) increase in urbanisation rate and PCFA. By 2100, the total stock is estimated to have a 95% credible interval of 39.1 to 47.3 billion m². The probability distribution of annual total stock size in various years is presented in Figure 6. This figure offers a full picture of the probability distribution of the total stock size per year, as the uncertainties associated with not only model-specific parameter vector, but also the models in the model space are propagated through the emergent behaviours of the model outputs, namely, the dynamic evolution of the stock and the associated substocks.

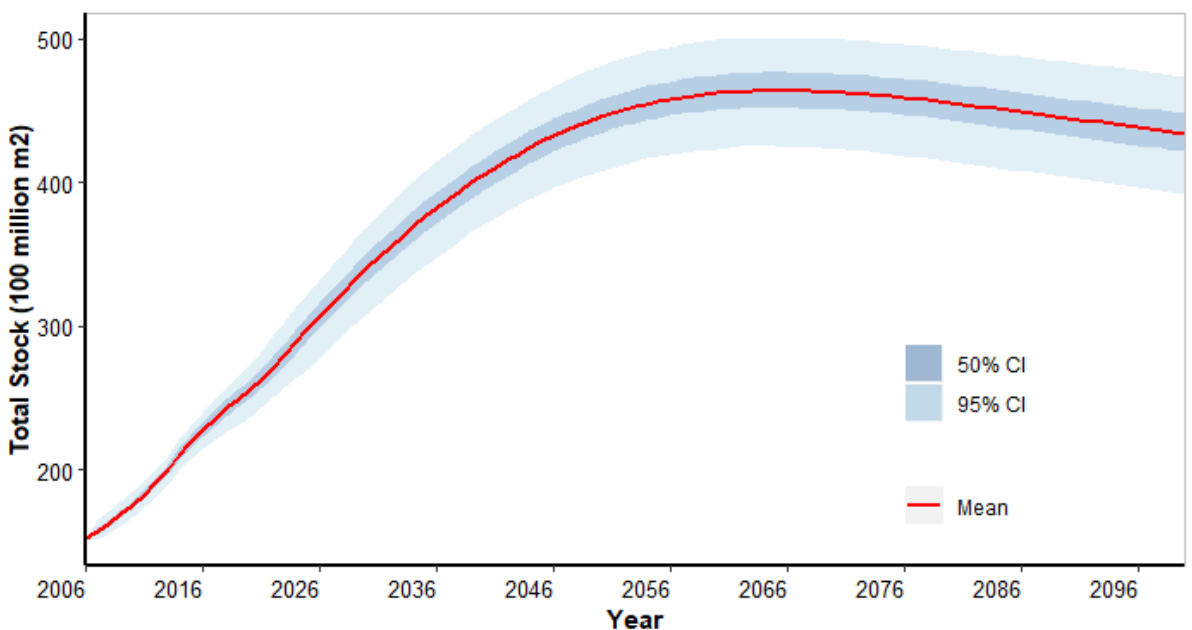


Figure 5: Posterior predictive distribution of stock trajectories from 2006 to 2100

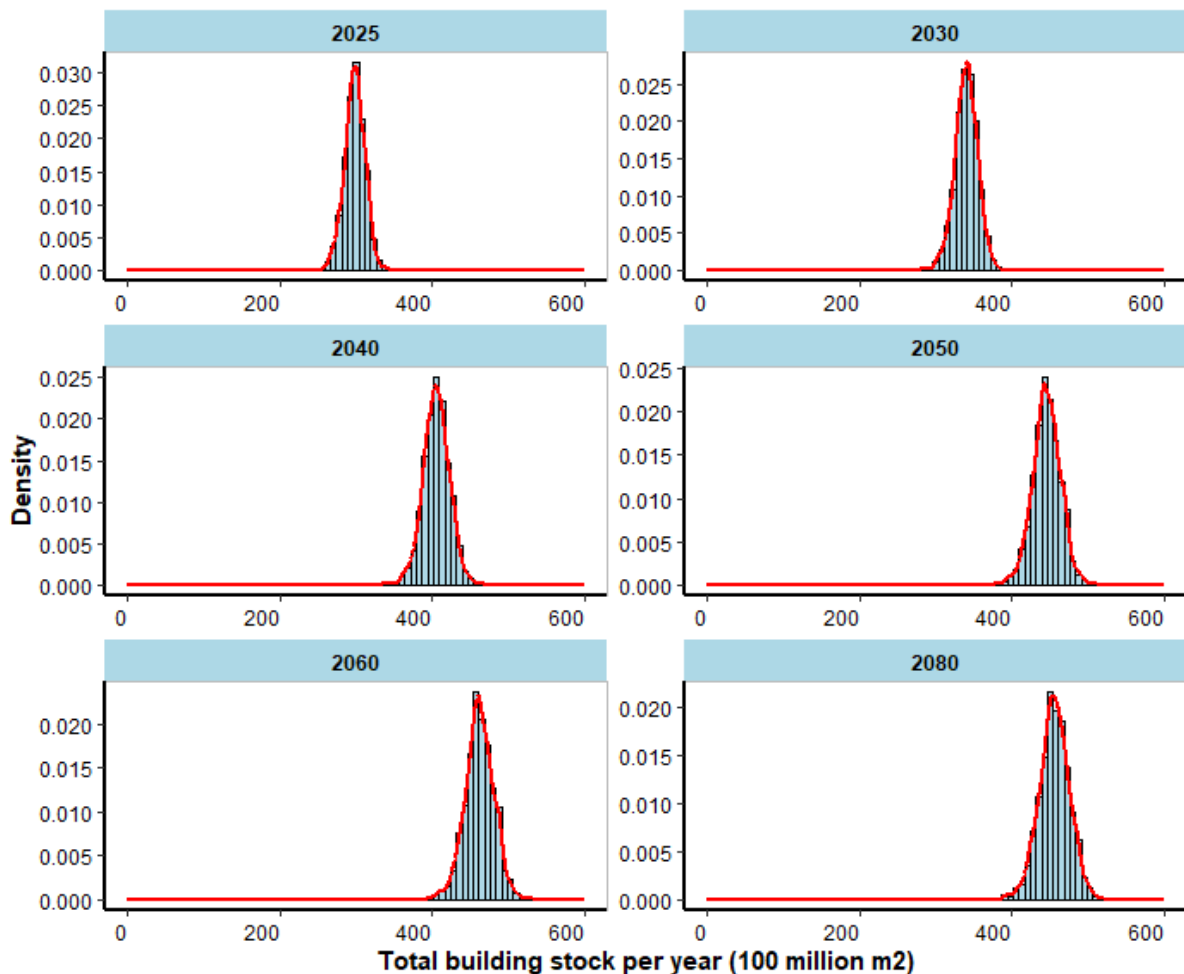


Figure 6: Posterior predictive distribution of total stock in various years

Underlying total stock is the dynamic evolution of the underlying age-specific substocks, which collectively form the total stock. The highly granular building stock turnover model developed through System Dynamics modelling offers additional insights into the composition of building stock through explicitly modelled building aging process. For each parameter vector in the parameter space of a candidate model in the mode space, the annual total stock is disaggregated into age-specific substocks, each of which goes through an aging process subject to age-specific demolition rate determined by the hazard function specified by this particular parameter vector of this particular candidate model. For each year, the substock of new buildings constructed in this year and the substocks of existing buildings at various ages that remain in use collectively create the age profile of the entire stock. For the year after, the stock's age profile is updated due to new construction, aging and demolition. These on-going dynamics, which result in the turnover of the overall stock and detailed representation of age-specific substocks, are fully captured in the dynamic model and, more importantly, are further characterised probabilistically by the BMA model ensemble through the posterior distributions of model-specific parameters and PMPs of candidate models. This allows us to obtain the full distribution of each age-specific substock in any given year and makes possible tracking and analysis of specific substocks. Figure 7 shows the posterior predictive distributions of substocks aged 10, 20, 30, 40, 50 and 60 years within the total stock in 2060.

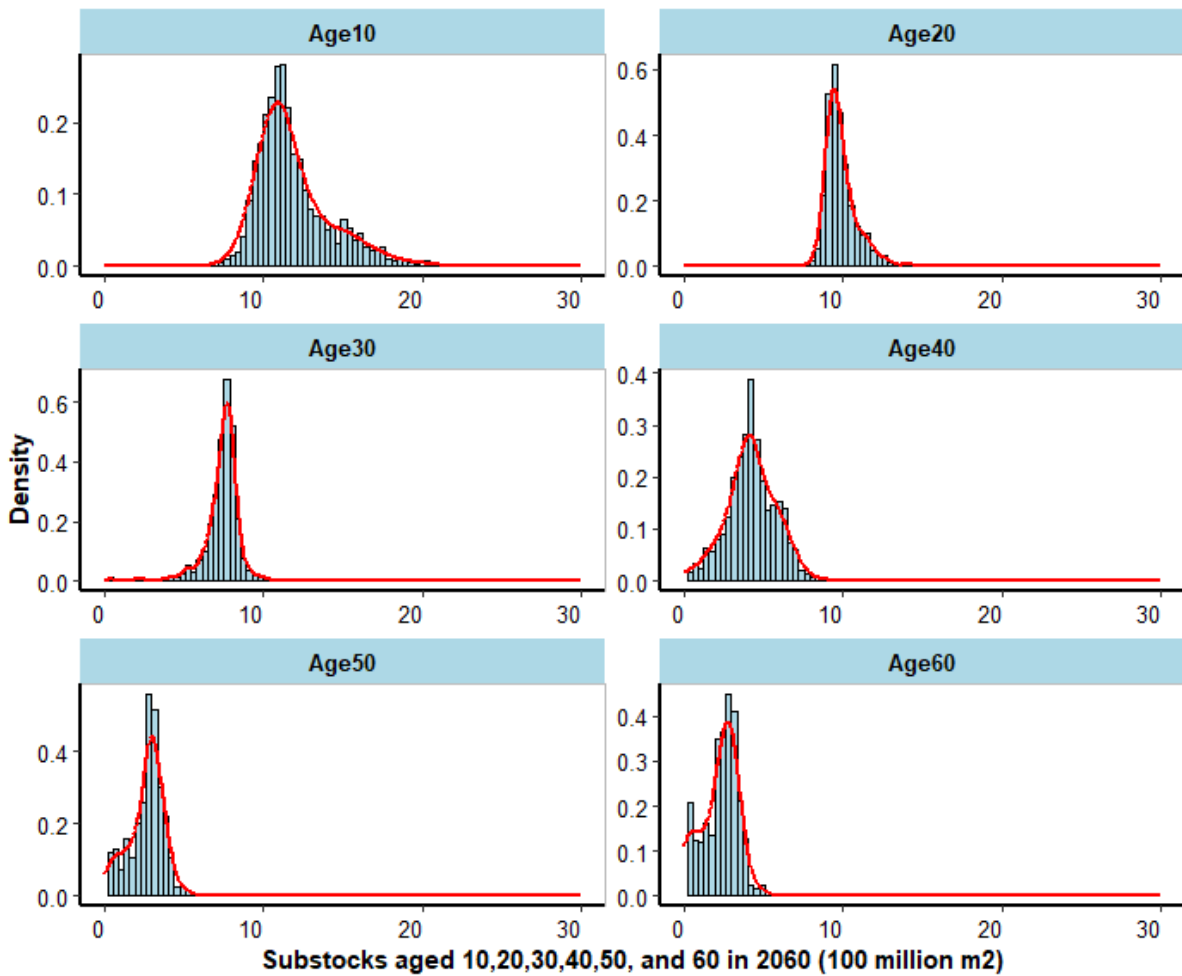


Figure 7: Posterior predictive distribution of substocks at various ages in 2060

Figure 8 and Figure 9 respectively show the distributions of annual construction of new buildings and annual demolition of old buildings obtained from the BMA model ensemble. Overall, future new buildings constructed per year are not expected to vary significantly from the recent historic levels. After experiencing a gradual increase from 2018 to 2025, the mean of annual new buildings maintains a generally stable level at approximately 1.5 billion m^2 for two decades and then starts to progressively decrease and eventually approach 1 billion m^2 by 2100. Such a general trend is as expected, because annual new buildings, serving as the incremental part of the total stock, is predominantly influenced by the expected demand of total stock whose increase rate slows down over time. As for annual demolition, its curve shows a smoothly slowly ascending trend because a stock that builds up over time comes with an increasing aggregate of various amounts of demolished buildings at various ages. From 2053 to 2100, the mean of annual demolition remains being in a relatively narrow range of 1.1 to 1.2 billion m^2 for nearly 50 years. The peak of annual demolition occurs around 2073 and 2074, several years later than that of the total stock. The reason is that, from a stock-and-flow perspective, the demolition is an outflow from the stock with a time lag relative to the inflow. Whilst the annual demolition is made up of by a large number of age-specific demolition quantities corresponding to age-specific substocks, such a mathematical complexity does not

change the fundamental mechanism but only the extent of time lag, which is dependent upon building lifetime distribution. For both annual new buildings and annual demolition, their predictive distributions, as characterised by 50% and 95% credible intervals, reflect the uncertainties associated with the parameters of the Logistic model for urbanisation, the parameters of the Gompertz model for PCFA, the parameters of each of the five survival models for lifetime distribution and demolition probability, and the survival models themselves.

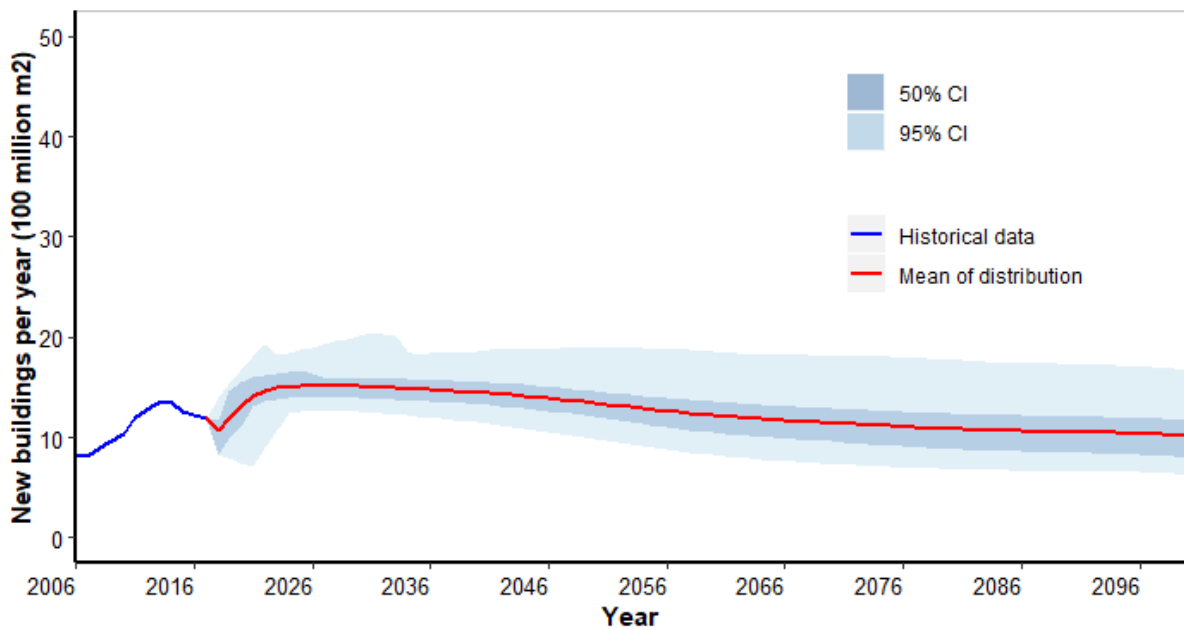


Figure 8: Posterior predictive distribution of annual new buildings

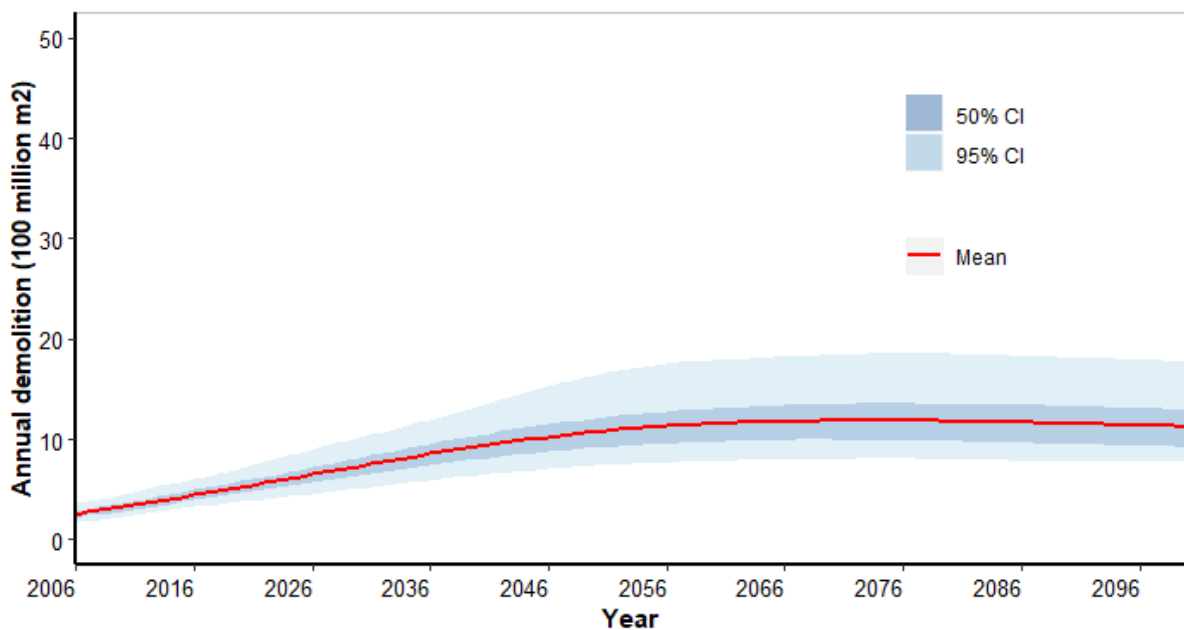


Figure 9: Posterior predictive distribution of annual demolition

4.4 Implications for building stock energy modelling

The value of establishing the BMA model ensemble and obtaining posterior predictive distributions of stock, sub-stocks, annual construction and annual demolition goes beyond forecasting the turnover dynamics of the residential stock per se. It has significant implications for further modelling and analysis of building energy consumption and carbon emissions at stock level.

Firstly, the possible lifetime distribution profile specified by a parameter vector in the parameter space of a candidate model enables explicit estimate of annual construction of new buildings and demolition of old buildings, which are fundamental to quantifying the initial and demolition embodied energy and carbon incurred every year. The lifetime distribution enables a closer look at annualised embodied energy and carbon converted from the capital cost in embodied energy and carbon associated with future cohorts of new buildings, in the context of changing building materials and construction techniques and the resultant changing embodied energy and carbon intensities. The impact of potentially varying lifetime distribution on embodied energy and carbon, such as longer average lifetime and smaller variance, via planning policy or as a result of economic and environmental factors, can be examined.

Secondly, model granularity at the level of age-specific building sub-stocks offers a detailed representation of the building stocks heterogeneity with respect to operational energy performance. With buildings identified as a key sector for energy savings, it is reasonable to expect that new buildings, particularly those in urban areas, will be built to higher standards of operational energy performance due to increasingly stringent design codes, on-going technological advances and improving operation and management practices. Separately tracking the aging process of different cohorts of buildings enables a holistic and in-depth understanding the dynamics of the stock composition of buildings with different operational performance and a detailed evaluation of the trajectories of stock-wide average operational energy intensity per m². Explicitly modelling the aging process also provides analytical convenience to enable detailed policy experimentation, such as by targeting old buildings at different ages for different depths of energy-related retrofits, which shall make not only technical but also economic sense.

Thirdly, more importantly, the ability to model the temporal stock dynamics enables the integration of embodied and operational dimensions of building energy and carbon. By simultaneously investigating both dimensions, it is possible to explore their relative importance in the context of future building sectoral developments in increasingly extensive production and use of green building materials, improving construction practice and building quality, strengthening design codes for new buildings, scaling up energy-related retrofits of existing buildings, possible policy considerations to manage aged buildings, etc. In so doing, a fuller understanding of stock-level lifecycle energy and carbon of urban residential buildings can be reached so as to better assist policy-makers in formulating policies aiming to promote energy savings and decarbonisation of buildings.

Across the three dimensions, the uncertainties associated with model-specific parameter vectors and candidate models, as fully captured by the BMA model ensemble, along with uncertainties of other parameters and input variables needed for modelling energy and carbon, can be propagated into the emergent stock-level outputs, such as annual total embodied energy and annual total operational energy of total stock. The full Bayesian approach and the resultant probabilistic distributions of stock-level outputs can mitigate the risk of potential over- or under-estimate that would otherwise be more likely to be produced by deterministic models. From a policy-making perspective, probability distributions of different potential outcomes of policy scenarios can be generated. This is particularly important in the context of analysing the decarbonisation of the generally short-lived Chinese buildings, where there is likely to be a strategic trade-off between operational and embodied energy due to the above-mentioned factors. For example, it would be useful to explore the probability that one policy, e.g. extending building lifetime to avoid embodied energy, would yield a more favourable outcome of stock-level whole-life building energy consumption as compared to another policy, e.g. accelerating stringency of new building design codes to reduce operational energy. In this sense, this study creates a powerful modelling framework with enhanced robustness and reliability, thereby preparing a solid ground for more effectively experimenting and analysing policies aiming to decarbonise buildings in the broader context of curbing China's economy-wide emissions.

4.5 Wider applicability of the modelling approach

It is useful to accentuate the value of this study in a broader international context. This study contributes an innovative methodological approach to the general field of building stock modelling. The generality, flexibility and transparency of the approach enables its application in a wide variety of geographical contexts. It is particularly relevant and useful to countries experiencing rapid urbanisation and massive construction, such as developing countries in South and Southeast Asia [32–34]. Addressing building energy and carbon emissions has been emphasized as a key climate change mitigation strategy in the Nationally Determined Contributions (NDCs) of these countries [86,87]. Correspondingly, building energy is also amongst the focus areas on the development agenda of multilateral and bilateral donor agencies providing loans, grants and technical assistance to these countries. The modelling approach and its application in the Chinese context as elaborated in this paper is well placed to provide a useful reference for these countries and donor agencies to take stock of their existing buildings, forecast possible stock evolution pathways and evaluate stock-level energy consumption and carbon emissions under different policy and intervention scenarios.

5 Conclusions

A good understanding of building stock turnover dynamics is a fundamental prerequisite to sensible modelling of stock-level building energy consumption to better inform policy-making. We present a novel modelling approach to estimating recent historical total stock of urban residential buildings in China and also to forecasting future trajectories of the stock evolution. A disaggregated build stock model is developed using System Dynamics to characterise the building aging process and stock turnover dynamics. This model is then operationalised by

separately investigating five candidate parametric survival models to represent the uncertainties associated with building lifetime. With each survival model, the stock turnover is simulated through Markov Chain Monte Carlo methods to obtain the posterior predictive distribution of total historical stock and the marginal likelihood used to estimate the posterior model probability. Bayesian Model Averaging is applied to create a model ensemble to combine model-specific predictions of the historical stock evolution pathway based on model probabilities. By extending the model structure and incorporating variables relating to possible trends in urbanisation and demand for per capita floor area, future stock turnover dynamics through 2100 are forecasted and then combined through model averaging. In so doing, we can obtain not only forecasts of total stock, age-specific substocks, annual new construction and annual demolition, but also their posterior predictive distributions which fully characterise their uncertainties.

In summary, our study offers a first-of-a-kind analysis that employs a full Bayesian approach to investigate the uncertainties associated with modelling Chinese building stock, which is a policy relevant but under-researched area. The modelling approach adopted here is well suited to carry out studies of stock-level energy and carbon impacts. In particular, the model's ability to explicitly track the aging process of substocks and fully represent probability distributions at both the stock and substock level is critical to analysing policy trade-offs facing Chinese residential buildings regarding embodied versus operational energy consumption and carbon emissions in the context of sector-wide decarbonisation. Beyond the present study aimed at assessing the Chinese building stock, the generality and flexibility of the modelling approach suggests its wider applicability in other geographical contexts.

6 References

- [1] IEA, Buildings sector energy intensity in selected regions in the Sustainable Development Scenario, 2000-2030, (2019). <https://www.iea.org/data-and-statistics/charts/buildings-sector-energy-intensity-in-selected-regions-in-the-sustainable-development-scenario-2000-2030> (accessed May 21, 2020).
- [2] Global Alliance for Buildings and Construction, 2019 Global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector, 2019. <https://globalabc.org/uploads/media/default/0001/02/dbe116e4b6799486d69b2157bdcb30532b3c9a0c.pdf>.
- [3] IEA, Tracking Buildings, (2019). <https://www.iea.org/reports/tracking-buildings> (accessed May 21, 2020).
- [4] IEA, Energy Technology Perspectives 2017 - Catalysing Energy Technology Transformations, 2017. <http://www.iea.org/etp2017/>.
- [5] National Bureau of Statistics, China Statistical Yearbook 2019, China Statistics Press, 2019.
- [6] IEA, Estimated global buildings construction in the Sustainable Development Scenario, 2018-2030, (2020).
- [7] IEA, World Energy Outlook 2019, 2019. https://www-oecd-ilibrary-org.ezp.lib.cam.ac.uk/energy/world-energy-outlook-2019_caf32f3b-en.
- [8] THUBERC, 2020 Annual Report on China Building Energy Efficiency, China Architecture & Building Press, Beijing, 2020.
- [9] THUBERC, 2015 Annual Report on China Building Energy Efficiency, China Architecture & Building Press, 2015.
- [10] W. Yang, N. Kohler, Simulation of the evolution of the Chinese building and infrastructure stock, *Build. Res. Inf.* 36 (2008) 1–19. <https://doi.org/10.1080/09613210701702883>.
- [11] B.W. Fawley, Y. Wen, The Great Chinese Housing Boom, *Econ. Synopses.* (2013). <http://research.stlouisfed.org/publications/es/article/9774>.
- [12] W. Huang, Speech at the meeting of “She Hui Zhu Yi Xin Nong Cun Jian She” [“Building new countryside of the socialist society”], (2006). http://www.mohurd.gov.cn/jsbffd/200612/t20061225_165486.html (accessed March 16, 2019).
- [13] M. Hu, H. Bergsdal, E. Van Der Voet, G. Huppel, D.B. Muller, Dynamics of urban and rural housing stocks in China, *Build. Res. Inf.* 38 (2010) 301–317. <https://doi.org/10.1080/09613211003729988>.
- [14] W. Cai, L. Wan, Y. Jiang, C. Wang, L. Lin, Short-Lived Buildings in China: Impacts on Water, Energy, and Carbon Emissions, *Environ. Sci. Technol.* 49 (2015) 13921–13928. <https://doi.org/10.1021/acs.est.5b02333>.

- [15] Institution of Civil Engineers, Embodied Energy and Carbon, (2015). <https://www.ice.org.uk/knowledge-and-resources/briefing-sheet/embodied-energy-and-carbon>.
- [16] THUBERC, 2019 Annual Report on China Building Energy Efficiency, China Architecture & Building Press, Beijing, 2019.
- [17] IEA, Transition to Sustainable Buildings - Strategies and opportunities to 2050, 2013. <https://doi.org/10.1787/9789264202955-en>.
- [18] IPCC, Buildings. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, 2014. https://www.ipcc.ch/pdf/assessment-report/ar5/wg3/ipcc_wg3_ar5_chapter9.pdf.
- [19] D. Urge-Vorsatz, K. Petrichenko, M. Antal, M. Staniec, M. Labelle, E. Ozden, et al., Best Practice Policies for Low Energy and Carbon Buildings: A Scenario Analysis, 2012. [http://www.gbpn.org/sites/default/files/08.CEU Technical Report copy_0.pdf](http://www.gbpn.org/sites/default/files/08.CEU%20Technical%20Report%20copy_0.pdf).
- [20] M. Hu, S. Pauliuk, T. Wang, G. Huppes, E. van der Voet, D.B. Müller, Iron and steel in Chinese residential buildings: A dynamic analysis, *Resour. Conserv. Recycl.* 54 (2010) 591–600. <https://www.sciencedirect.com/science/article/pii/S0921344909002407>.
- [21] M. Hu, E. van der Voet, G. Huppes, Dynamic Material Flow Analysis for Strategic Construction and Demolition Waste Management in Beijing, *J. Ind. Ecol.* 14 (2010) 440–456. <https://doi.org/10.1111/j.1530-9290.2010.00245.x>.
- [22] F. Shi, T. Huang, H. Tanikawa, J. Han, S. Hashimoto, Y. Moriguchi, Toward a low carbon-dematerialization society: Measuring the materials demand and co₂ emissions of building and transport infrastructure construction in China, *J. Ind. Ecol.* 16 (2012) 493–505. <https://doi.org/10.1111/j.1530-9290.2012.00523.x>.
- [23] T. Huang, F. Shi, H. Tanikawa, J. Fei, J. Han, Materials demand and environmental impact of buildings construction and demolition in China based on dynamic material flow analysis, *Resour. Conserv. Recycl.* 72 (2013) 91–101. <https://doi.org/10.1016/j.resconrec.2012.12.013>.
- [24] L. Hong, N. Zhou, W. Feng, N. Khanna, D. Fridley, Y. Zhao, et al., Building stock dynamics and its impacts on materials and energy demand in China, *Energy Policy*. 94 (2016) 47–55. <https://doi.org/10.1016/j.enpol.2016.03.024>.
- [25] J. Shi, W. Chen, X. Yin, Modelling building's decarbonization with application of China TIMES model, *Appl. Energy*. 162 (2016) 1303–1312. <https://doi.org/10.1016/j.apenergy.2015.06.056>.
- [26] T. Huo, H. Ren, W. Cai, Estimating urban residential building-related energy consumption and energy intensity in China based on improved building stock turnover model, *Sci. Total Environ.* 650 (2019) 427–437. <https://doi.org/10.1016/j.scitotenv.2018.09.008>.
- [27] R.L. Eberlein, J.P. Thompson, Precise modeling of aging populations, *Syst. Dyn. Rev.* 29 (2013) 87–101. <https://doi.org/10.1002/sdr.1497>.
- [28] Z. Cao, L. Shen, S. Zhong, L. Liu, H. Kong, Y. Sun, A Probabilistic Dynamic Material Flow Analysis Model for Chinese Urban Housing Stock, *J. Ind. Ecol.* 22 (2018) 377–

391. <https://doi.org/10.1111/jiec.12579>.
- [29] W. Zhou, A. Moncaster, D.M. Reiner, P. Guthrie, Estimating lifetimes and stock turnover dynamics of urban residential buildings in China, *Sustain.* 11 (2019). <https://doi.org/10.3390/su11133720>.
- [30] G. Li, J. Shi, Application of Bayesian model averaging in modeling long-term wind speed distributions, *Renew. Energy.* 35 (2010) 1192–1202. <https://doi.org/10.1016/j.renene.2009.09.003>.
- [31] T.M. Fragoso, W. Bertoli, F. Louzada, Bayesian Model Averaging: A Systematic Review and Conceptual Classification, *Int. Stat. Rev.* 86 (2018) 1–28. <https://doi.org/10.1111/insr.12243>.
- [32] World Bank, East Asia's Changing Urban Landscape: Measuring a Decade of Spatial Growth. Urban Development Series, Washington, DC, 2015. <https://openknowledge.worldbank.org/handle/10986/21159>.
- [33] P. Ellis, M. Roberts, Leveraging Urbanization in South Asia: Managing Spatial Transformation for Prosperity and Livability, World Bank Publications, Washington, DC, 2016. <https://doi.org/10.1596/978-1-4648-0662-9>.
- [34] ASEAN Secretariat, ASEAN Sustainable Urbanisation Strategy, Jakarta, Indonesia, 2018. <https://asean.org/storage/2018/11/ASEAN-Sustainable-Urbanisation-Strategy-ASUS.pdf>.
- [35] P. Allison, *Survival Analysis Using SAS: A Practical Guide*, SAS Institute, 2010.
- [36] X. Liu, *Survival Analysis: Models and Applications*, Wiley, 2012. <https://doi.org/10.1002/9781118307656>.
- [37] Ventana Systems Inc., Vensim Software, (2019). <https://vensim.com/vensim-software/>.
- [38] J.A. Hoeting, D. Madigan, A.E. Raftery, C.T. Volinsky, Bayesian Model Averaging: A Tutorial, *Stat. Sci.* 14 (1999) 382–417. <https://projecteuclid.org/euclid.ss/1009212519>.
- [39] J.A. Hoeting, Methodology for Bayesian model averaging: an update, in: *Proc. Invit. Pap. Present. Int. Biometric Conf.*, Freiburg, Germany, 2002: pp. 231–240. [https://doi.org/10.1016/s0093-691x\(99\)00089-8](https://doi.org/10.1016/s0093-691x(99)00089-8).
- [40] C. Forbes, M. Evans, N. Hastings, B. Peacock, *Statistical Distributions (Fourth Edition)*, John Wiley & Sons, Inc., Hoboken, New Jersey, 2011.
- [41] Palisade, *Risk Analysis and Simulation Add-In for Microsoft® Excel*, 2016. https://www.palisade.com/downloads/documentation/75/EN/RISK7_EN.pdf.
- [42] OECD, *Measuring capital: OECD manual: measurement of capital stocks, consumption of fixed capital and capital services*, 2001. <https://www.oecd.org/std/na/1876369.pdf>.
- [43] OECD, *Measuring Capital: OECD manual 2009*, 2009. <https://www.oecd.org/std/productivity-stats/43734711.pdf>.
- [44] A. Miatto, H. Schandl, H. Tanikawa, How important are realistic building lifespan assumptions for material stock and demolition waste accounts?, *Resour. Conserv. Recycl.* 122 (2017) 143–154. <https://doi.org/10.1016/j.resconrec.2017.01.015>.

- [45] I.M. Johnstone, Energy and mass flows of housing: Estimating mortality, *Build. Environ.* 36 (2001) 43–51. [https://doi.org/10.1016/S0360-1323\(99\)00066-9](https://doi.org/10.1016/S0360-1323(99)00066-9).
- [46] R.A. Bohne, H. Brattebø, H. Bergsdal, P.J. Hovde, Estimation of the Service Life of Residential Buildings, and Building Components, in Norway., in: *Toward. City Surf. Tomorrow*, 2006: pp. 29–33. https://www.academia.edu/27632589/Estimation_of_the_Service_Life_of_Residential_Buildings_and_Building_Components_in_Norway.
- [47] D.B. Müller, Stock dynamics for forecasting material flows — Case study for housing in The Netherlands, *Ecol. Econ.* 59 (2006). <https://doi.org/10.1016/j.eco>.
- [48] M. Aksözen, U. Hassler, M. Rivallain, N. Kohler, Mortality analysis of an urban building stock, *Build. Res. Inf.* 45 (2016) 1–19. <https://doi.org/10.1080/09613218.2016.1152531>.
- [49] C. McLaren, C. Stapenhurst, ONS Methodology Working Paper Series No 3: A note on distributions used when calculating estimates of consumption of fixed capital, 2015. <http://webarchive.nationalarchives.gov.uk/20160111030849/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/gss-methodology-series/ons-working-paper-series/index.html>.
- [50] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, D.B. Rubin, *Bayesian Data Analysis, Third Edition*, Chapman and Hall/CRC, 2014. <https://doi.org/10.1007/s13398-014-0173-7.2>.
- [51] M.S. Hamada, A.G. Wilson, C.S. Reese, H.F. Martz, *Bayesian Reliability*, Springer-Verlag New York, 2008. <https://doi.org/10.1007/978-0-387-77950-8>.
- [52] S. Jackman, *Bayesian Analysis for the Social Sciences*, Wiley, 2009. <https://doi.org/10.1002/9780470686621>.
- [53] K. Soetaert, T. Petzoldt, Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME, *J. Stat. Softw.* 33 (2010) 1–28. <https://www.jstatsoft.org/article/view/v033i03>.
- [54] F. Hartig, F. Minunno, S. Paul, *BayesianTools: General-Purpose MCMC and SMC Samplers and Tools for Bayesian Statistics*, (2018). <https://github.com/florianhartig/BayesianTools>.
- [55] A. Schöniger, T. Wöhling, L. Samaniego, W. Nowak, Model selection on solid ground: Rigorous comparison of nine ways to evaluate Bayesian model evidence, *Water Resour. Res.* 50 (2014) 9484–9513. <https://doi.org/10.1002/2014WR016062>.
- [56] Q.F. Gronau, A. Sarafoglou, D. Matzke, A. Ly, U. Boehm, M. Marsman, et al., A tutorial on bridge sampling, *J. Math. Psychol.* 81 (2017) 80–97. <https://doi.org/10.1016/j.jmp.2017.09.005>.
- [57] E.L. Ionides, Truncated importance sampling, *J. Comput. Graph. Stat.* 17 (2008) 295–311. <https://doi.org/10.1198/106186008X320456>.
- [58] W. Xie, P.O. Lewis, Y. Fan, L. Kuo, M.H. Chen, Improving marginal likelihood estimation for bayesian phylogenetic model selection, *Syst. Biol.* 60 (2011) 150–160. <https://doi.org/10.1093/sysbio/syq085>.
- [59] A. Pajor, Estimating the marginal likelihood using the arithmetic mean identity, *Bayesian*

- Anal. 12 (2017) 261–287. <https://doi.org/10.1007/s11406-013-9487-5>.
- [60] S. Frühwirth-Schnatter, Estimating marginal likelihoods for mixture and Markov switching models using bridge sampling techniques, *Econom. J.* 7 (2004) 143–167. <https://doi.org/10.1111/j.1368-423X.2004.00125.x>.
- [61] Y.B. Wang, M.H. Chen, L. Kuo, P.O. Lewis, A new Monte Carlo method for estimating marginal likelihoods, *Bayesian Anal.* 13 (2018) 311–333. <https://doi.org/10.1214/17-BA1049>.
- [62] X.-L. Meng, H.W. Wong, Simulating ratios of normalizing constants via a simple identity: A theoretical exploration, *Stat. Sin.* 6 (1996) 831–860. <https://pdfs.semanticscholar.org/6a40/18c9c2927d702d85257c34130b1204fa7584.pdf>.
- [63] Q.F. Gronau, H. Singmann, E.-J. Wagenmakers, bridgesampling: An R Package for Estimating Normalizing Constants, (2018). <https://doi.org/10.1108/nfs.2003.01733cab.009>.
- [64] National Bureau of Statistics, *China Statistical Yearbook 2018*, China Statistics Press, 2018.
- [65] United Nations Population Division, *World Population Prospects: The 2017 Revision, 2017*. https://population.un.org/wpp/Publications/Files/WPP2017_Volume-I_Comprehensive-Tables.pdf.
- [66] Asian Development Bank, *Strategic Options for Urbanization in the People's Republic of China - Key Findings*, Manila, Philippines, 2013. <https://www.adb.org/sites/default/files/publication/30397/options-urbanization-prc-findings.pdf>.
- [67] UNDP, *Sustainable Urbanisation Strategy: UNDP's Support to Sustainable, Inclusive and Resilient Cities in the Developing World*, 2016. [http://www.undp.org/content/dam/undp/library/Sustainable Development/Urbanization/UNDP_Urban-Strategy.pdf](http://www.undp.org/content/dam/undp/library/Sustainable%20Development/Urbanization/UNDP_Urban-Strategy.pdf).
- [68] United Nations Population Division, *World Urbanization Prospects: The 2014 Revision, 2014*. <https://esa.un.org/unpd/wup/CD-ROM/>.
- [69] World Bank, *Urban China: Toward Efficient, Inclusive and Sustainable Urbanization*, Washington, DC, 2014. <https://openknowledge.worldbank.org/handle/10986/18865>.
- [70] C. V. Trappey, H.Y. Wu, An evaluation of the time-varying extended logistic, simple logistic, and Gompertz models for forecasting short product lifecycles, *Adv. Eng. Informatics.* 22 (2008) 421–430. <https://doi.org/10.1016/j.aei.2008.05.007>.
- [71] A.K. Shukla, S.K. Yadav, V. Tiwari, Linear Models for S-Shaped Growth Curves, *J. Stat. Appl. Probab.* 117 (2015) 113–117. <http://www.naturalspublishing.com/files/published/96sp961dj2687j.pdf>.
- [72] K.M.C. Tjørve, E. Tjørve, The use of Gompertz models in growth analyses, and new Gompertz-model approach: An addition to the Unified-Richards family, *PLoS One.* 12 (2017). <https://doi.org/10.1371/journal.pone.0178691>.
- [73] X. Li, D. Xu, Clarification of misled statistic data: overestimated per-capita housing area, *China Econ. J.* 6 (2013) 134–151. <https://doi.org/10.1080/17538963.2013.860681>.

- [74] L. Hong, N. Zhou, D. Fridley, W. Feng, N. Khanna, Modeling China's Building Floor-Area Growth and the Implications for Building Materials and Energy Demand, in: 2014 ACEEE Summer Study Energy Effic. Build., 2014: pp. 146–157. <https://aceee.org/files/proceedings/2014/data/papers/10-230.pdf>.
- [75] D. Xu, Y. Liu, Understanding China's Overcapacity, Springer Singapore, Singapore, 2018. <https://doi.org/10.1007/978-981-13-0881-9>.
- [76] X. Shen, Has the per capita urban residential floor area in China exceeded 32 m²? The biggest lie of Chinese livelihood, Tencent Hous. (2013). <https://house.qq.com/bside/31.htm> (accessed May 22, 2019).
- [77] W. Liang, Is the per capita urban residential floor area in China 33 m² or 18 m²?, (2014). http://www.cssn.cn/preview/zt/19364/19369/201403/t20140313_1027648.shtml (accessed May 22, 2019).
- [78] Z. Ren, C. Xiong, X. Bai, Urban housing in China: surplus or shortage? [in Chinese], (2019). <http://m.zqrb.cn/house/hangyedongtai/2019-02-18/A1550447505753.html> (accessed May 22, 2019).
- [79] THUBERC, 2017 Annual Report on China Building Energy Efficiency, China Architecture & Building Press, Beijing, 2017.
- [80] C. Peng, D. Yan, Y. Jiang, China Road Map for Building Energy Conservation, in: D. Stolten, V. Scherer (Eds.), Transit. to Renew. Energy Syst., Wiley-VCH Verlag GmbH & Co. KGaA, 2013: pp. 891–911. <https://doi.org/10.1002/9783527673872.ch42>.
- [81] ERI, A study on low-carbon building scenarios and policy roadmaps in China [in Chinese], 2014. <http://www.efchina.org/Reports-zh/reports-20140706-zh>.
- [82] IEA, Energy Technology Perspectives 2015 - Mobilising Innovation to Accelerate Climate Action, 2015. <http://www.iea.org/etp/etp2015/>.
- [83] M. LaFrance, Y. Jiang, D. Yan, J. Dulac, S. Guo, S. Hu, et al., Building Energy Use in China: Transforming Construction and Influencing Consumption to 2050, 2015. <https://www.iea.org/reports/partner-country-series-energy-use-in-the-chinese-building-sector>.
- [84] National Center for Climate Change Strategy and International Cooperation, Pathway and policy analysis to China's deep decarbonization [in Chinese], 2018. <http://www.ncsc.org.cn/yjcg/zlyj/201804/P020180920508767376089.docx>.
- [85] THUBERC, 2012 Annual Report on China Building Energy Efficiency, China Architecture & Building Press, 2012.
- [86] IEA and UNEP, 2018 Global Status Report - Towards a zero-emission, efficient and resilient buildings and construction sector, 2018. <https://www.unenvironment.org/resources/report/global-status-report-2018>.
- [87] UNFCCC, NDC Registry (Interim), (2020). <https://www4.unfccc.int/sites/NDCStaging/Pages/All.aspx> (accessed May 25, 2020).

