1 Intro

1) The effects of public transport systems on urban development patterns, and 2) Evaluating public transport accessibility and its relationship with urban development.

On the first area, prior work has generally established that public transport investments can shape urban form and land use in several ways. Some studies highlight how high-capacity transit enables more compact, high-density urban development by reducing reliance on automobiles (Author 1 et al., 2004; Author 2, 1989). Others examine public transport's role in achieving sustainable urbanization goals like reducing energy use, air pollution, and traffic congestion (Author 3 et al., 2016; Author 4, 1986). Additionally, research has identified symbiotic relationships where transportation infrastructure influences surrounding land uses, which in turn generate new travel demand patterns influencing future transport investments (Author 5, 1999; Author 6 et al., 2013).

Within this broad area, the transit-oriented development (TOD) concept has emerged as an intensive, mixed-use urban model centering growth around transit nodes and promoting public transport use (Author 7, 1993). While early TOD research focused on impacts to travel behavior (Author 8 et al., 1996; Author 9 et al., 2014), more recent work has analyzed effects on urban development intensity, land values, job accessibility and economic development (Author 10 et al., 2001; Author 11, 2015; Author 12 et al., 2013).

The second area covers different approaches to defining, quantifying and evaluating public transport accessibility. Scholars have proposed various metrics focusing on spatial distribution of opportunities, transport network performance, potential economic benefits, and individual mobility constraints (Author 13 & Author 14, 2007; Author 15 et al., 2002; Author 16 et al., 2021). When studying TOD specifically, measures of access to transit stations and surrounding land use characteristics become highly relevant (Author 17 et al., 2011; Author 18 et al., 2012).

Despite this extensive theoretical and empirical research, gaps remain in understanding the precise nature of the relationship between public transport accessibility and urban development levels. Most prior studies assume linear associations and do not account for potential nonlinear effects or thresholds, especially in the context of topographically-constrained mountainous cities. This study addresses this gap through a case analysis combining big data, machine learning, and interpreting nonlinear effects in a major Chinese mountainous city.

2 Q

What is the relationship between public transport accessibility and urban development in mountainous cities like Chongqing?

What are the relative importance of different public transport accessibility factors in influencing urban development in Chongqing?

3 method

To investigate the nonlinear relationship between public transport accessibility and urban development levels in the mountainous city of Chongqing, this study employs a data-driven analytical approach combining big data sources, machine learning algorithms, and interpretable modeling techniques.

The study employs a powerful machine learning technique called Gradient Boosting Decision Trees (GBDT) to model the potentially nonlinear relationships between the collected variables. GBDT is an ensemble method that incrementally combines weak regression tree models to optimize predictive performance. Crucial advantages include high accuracy, ability to automatically handle non-linearities and interactions, and inbuilt prevention of overfitting.

To interpret the resulting GBDT model's nonlinear effects and each variable's relative importance, a state-of-the-art method called SHapley Additive exPlanations (SHAP) is used. SHAP computes Shapley values, a game theory-based approach to fairly distribute the prediction among feature contributions. Visual techniques like dependency plots are then employed to scrutinize variable influence across their data ranges.

The complete analytical workflow spans data collection and cleaning, GBDT model training on Chongqing's metro service area samples using the curated variables, interpreting the GBDT's nonlinear effects via SHAP values and dependence plots, and finally deriving planning implications based on the nonlinear effects revealed for different types of stations/neighborhoods.

4 data

The level of urban development is quantified using point of interest (POI) data obtained from a major online mapping platform's API. POIs represent geographic locations and entities crucial for daily urban activities like restaurants, shopping centers, businesses etc. Their density and concentration can serve as a proxy for development intensity. For Chongqing, over 50,000 filtered POIs falling within 600m radius metro service areas were collected.

Public transport accessibility measures were derived from several spatial data sources. Road network data from OpenStreetMaps enabled calculating street-level betweenness centrality. Official metro system data including station locations and rail routes were used to compute various rail accessibility metrics like average distances to stations, network betweenness, and flow/demand measures using built-in network analysis tools.

Other relevant variables extracted include the number of bus stops in each metro service area, walking times to the central business district, station characteristics (interchange status, years operating), and average residential property prices as a control variable.

5 ex

6 analysis

7 dis

Q

The study found nonlinear and threshold effects of public transport accessibility factors like metro betweenness, street accessibility, and number of bus stops on urban development levels measured by number of POIs in metro service areas of Chongqing.

Metro betweenness and street accessibility measures had the strongest positive influence, followed by some other metro accessibility metrics. Bus stop density had a relatively weaker effect.