Cities 155 (2024) 105486

Contents lists available at ScienceDirect

Cities

journal homepage: www.elsevier.com/locate/cities

Assessing and modeling the impact of urbanization on infrastructure development in Africa: A data-driven approach

Rachida El-Bouayady [a,*], Hassan Radoine [a], Nour-Eddin El Faouzi [a,b], Soukaina Tayi [a], Hakki Can Ozkan [c]

[a] *School of Architecture, Planning and Design (SAP+D), University Mohammed VI Polytechnic (UM6P), Ben Guerir 43150, Morocco*
[b] *LICIT-ECO7, Gustave Eiffel University, 25, avenue F. Mitterrand, F-69675 Lyon, France*
[c] *Faculty of Architecture and Design, Fatih Sultan Mehmet Vakıf University, Sütlüce, Karaağaç Cd 12/A, 34445 Beyoğlu, İstanbul, Turkey*

ARTICLE INFO

Keywords:
Urban infrastructure development
Urbanization
2-Step GMM
Data-driven approach
Simulation
Africa

ABSTRACT

Africa has recently experienced significant infrastructure endowments alongside unprecedented urbanization, leading to increased demand for infrastructure services. This paper empirically examines the relationship between urban infrastructure and urbanization across 42 African countries from 2005 to 2021. Utilizing panel data, we employ a 2-step system Generalized Method of Moments (GMM) approach, with six specifications based on infrastructure type. The findings indicate that (i) the impact of urbanization on infrastructure development varies by infrastructure type, generally showing a positive effect; (ii) economic growth enhances all infrastructure types, with a 10 % increase in GDP leading to a 4.2 % improvement in broadband infrastructure; (iii) industrial production's effects are mixed across infrastructure types; and (iv) trade openness has a dual influence, positively affecting overall infrastructure and most ICT components, while negatively impacting fixed broadband subscriptions and electricity infrastructure. Additionally, a cluster-based analysis is conducted to simulate infrastructure development, categorizing variations in development stages and configurations. The study discusses policy implications derived from these findings.

1. Introduction

The crucial role of infrastructure has gained worldwide attention in the literature, particularly after the pioneering study by (Aschauer, 1989). Urban infrastructure—referring here to transportation, energy, Information and Communication Technology (ICT), water, and sanitation—directly affects the livelihood, mobility, health, and quality of life of people. Thus, it is essential for ensuring the sustainability and resilience of communities. The UN's SDG9 underscores the importance of developing sustainable infrastructure to enhance economic growth and well-being, focusing on affordability and accessibility.

In an increasingly urbanized Africa, it will be crucial to ensure that urban services are delivered to a wider population through adequate infrastructure. Recent empirical studies in the context of Africa show that infrastructure development significantly reduces poverty and enhances the positive effects of education, employment, and environmental quality while mitigating the negative impacts of income inequality and urbanization (Calderón & Servén, 2004; Calderón & Servén, 2010; Saadaoui Mallek et al., 2024). Urbanization is necessary for cities' growth and development, especially in developing countries. No country has ever reached middle-income status without a significant population shift into cities (Spence et al., 2009). National governments need to understand how urbanization interacts with infrastructure development to improve service provision in developing urban areas.

Urbanization in Africa presents both challenges and opportunities. Rapid population growth in cities requires significant investments in infrastructure to avoid the pitfalls of unplanned urban expansion, such as slums and inadequate services (UN-Habitat, 2022). At the same time, if managed properly, this urban growth presents opportunities for economic development and improved quality of life. Urbanization has been examined from multiple perspectives. The prevalent metrics employed to gauge this phenomenon include the percentage of the population residing in urban areas (a demographic measure) and the extent of built-up areas (a spatial measure) (Tayi & Radoine, 2023). Urbanization has been extensively investigated in the literature (Njoh, 2003; Svirejeva-Hopkins, 2012; Güneralp et al., 2018). However, this interest mainly

---

* Corresponding author.
E-mail address: rachida.el-bouayady@um6p.ma (R. El-Bouayady).

https://doi.org/10.1016/j.cities.2024.105486
Received 15 April 2024; Received in revised form 25 September 2024; Accepted 1 October 2024

focused on the interrelationship between urbanization and economic growth. (Turok & McGranahan, 2013), pointed out that the relationship between urbanization and economic growth is variable and cannot simply be reduced to a linear relationship. This varying impact of types of urbanization on growth is also highlighted by (Gross & Ouyang, 2021), who found that economic growth is associated with increases in residual urbanization and not with natural urbanization.

Most studies on infrastructure have mainly focused on productivity, investment, and economic growth (Collier & Venables, 2016; Srinivasu & Srinivasa Rao, 2013; Owusu-Manu et al., 2019). An empirical study on the sub-Saharan region, (Calderón & Servén, 2010) found strong evidence that infrastructure development has a negative effect on income inequality and a favorable influence on long-term growth, as indicated by an increase in infrastructure stocks and improvement in the quality of infrastructure services. It is rather unexpected, in this aspect, that infrastructure research has grown independently of the wider literature on urbanization and urban growth (El-Bouayady & Radoine, 2023).

Despite the recognized importance of infrastructure, there is a lack of comprehensive understanding of how urbanization influence infrastructure development in African cities. Studies exploring the relationship between urbanization and infrastructure either focus on individual components of infrastructure (Kasraian et al., 2016; Pradhan et al., 2021; Wang et al., 2021a) or specific urban challenge (Saadaoui Mallek et al., 2024). However, few provide a holistic view that encompasses the multi-dimensional nature of urban infrastructure and the potential influence that urbanization might have.

This paper aims to address this gap by developing a methodological framework to analyze how urbanization affects infrastructure development in African cities. Specifically, it investigates the influence of urbanization on infrastructure development across 42 African countries from 2005 to 2021 using a data-driven approach. By examining the infrastructure-urbanization nexus through two-step GMM estimation and exploring development patterns via a clustering-based probability distribution model, this study provides a comprehensive empirical foundation for understanding these dynamics. The findings aim to guide policymakers in fostering sustainable infrastructure amidst rapid urban growth, contributing to broader goals of economic development and poverty reduction. The results will provide valuable insights for policymakers, helping them allocate resources more effectively and design strategies that leverage urbanization for sustainable development.

2. Literature review

2.1. Urbanization-infrastructure-economic growth nexus

Urbanization is a key driver of sustained economic growth and industrialization across regions and countries (Henderson, 2003). As cities grow, they become hubs of consumption and economic activity due to the concentration of populations and industrial enterprises (McFarlane, 2011). This urban concentration fosters external economies of scale, reduces transaction costs, and promotes enterprise specialization, ultimately enhancing economic efficiency (Krugman, 1991; Kumar & Kober, 2012).

However, the relationship between urbanization and economic development is complex, particularly in low-income African countries. Here, economic growth can positively influence urbanization, while the reverse can have nuanced and sometimes adverse effects on growth (Liddle & Messinis, 2015). Furthermore, this relationship can be nonlinear; for example, in ASEAN countries, urbanization initially supports economic growth but may hinder it beyond a threshold of approximately 68–70 % (Nguyen & Nguyen, 2016).

The rapid urbanization witnessed in Africa has significantly transformed the built environment, shifting rural areas into urban settlements and redistributing populations. This trend is projected to result in an increase of Africa's urban population from 395 million in 2010 to 1.339 billion by 2050, making the continent home to 22 % of the world's urban population (United Nations, 2019). Such unprecedented growth intensifies the strain on already insufficient infrastructure, which is essential for economic development. As noted in (Baker & Judy, 2012), the process of urbanization in Africa has not led to the expected levels of economic growth and prosperity. Severe deficiencies in infrastructure and services are defining features of cities in developing countries, compounded by a lack of economic dynamism, governance failures, inadequate land management, and social disintegration (Rasiah et al., 2022).

Adequate infrastructure—encompassing transportation, energy, and water services—plays a crucial role in enhancing productivity and national competitiveness (Palei, 2015). However, in many African countries, the provision of infrastructure lags behind population growth, leading to inefficiencies that hinder economic development. The limited infrastructure is often a reflection of low GDP per capita, creating a cycle of underdevelopment (World Bank, 1994).

Moreover, research has shown that mobile phone penetration can significantly drive economic growth and enhance financial inclusion in African countries, promoting urbanization and infrastructure-led growth (Andrianaivo & Kpodar, 2012; Chinoda & Kwenda, 2019) Similarly, investments in paved roads yield substantial economic returns, demonstrating the critical role of infrastructure in supporting economic growth (Ng et al., 2019).

As urbanization continues to reshape infrastructure demands, it is vital to recognize that inadequate infrastructure can exacerbate social challenges, limiting productivity and industrial development. The growing mismatch between urban population expansion and infrastructure capacity presents significant challenges for African cities striving for sustainable urban development (Baker & Judy, 2012; Krugman, 1991). Addressing these infrastructural deficits is essential for fostering economic resilience and improving the quality of life for urban residents.

2.2. Urban infrastructure: definition, theoretical framework, and status quo in Africa

There remains a lack of consensus on the definition of urban infrastructure, with interpretations varying across fields of study and academic perspectives. The foremost reference to the concept of infrastructure is (Hirschman, 1958), who describes it as encompassing “those basic services without which primary, secondary, and tertiary productive activities cannot function.” In the economic field, early studies refer to infrastructure as public capital as established by (Aschauer, 1989). Urban infrastructure is viewed as a combination of physical, economic, social, and technological systems, as noted in (Journal of Infrastructure Systems, 2024), which includes energy production, transportation, and waste management. Consequently, there have been attempts to distinguish between “hard” infrastructures—physical networks essential for a modern society—and “soft” infrastructure, which encompasses the institutional and governance systems that support these networks. In this article, urban infrastructure is set to include transportation, energy, Information and Communication Technology (ICT), water, and sanitation—that directly affects the livelihood, mobility, health, and quality of life of people. These elements are frequently discussed in the literature (Srinivasu & Srinivasa Rao, 2013), yet they are often studied in isolation or in conjunction with other phenomena. Notably, transportation and ICT are more commonly associated with infrastructure discussions compared to other components (Kasraian et al., 2016; Li, Song, & Chen, 2017; Wang et al., 2019; Pradhan et al., 2021; Wang et al., 2021b; Maparu & Mazumder, 2021). On the other hand, other studies adopt diverse thematic measures of infrastructure. (Heshmati & Rashidghalam, 2020) define infrastructure as a composite index comprising 15 components. In contrast, (Li, Zheng, et al., 2017) employed five thematic areas of infrastructure, which include energy efficiency, sustainable urban transport, water and wastewater, and urban ecosystem management.

Despite the rapid growth of urban populations in numerous African countries over the past few decades, key challenges related to urban infrastructure remain pervasive. Issues such as underdeveloped infrastructure, pollution, congestion, inadequate housing, and persistent poverty continue to hamper sustainable urban development (Momoh, 2016). Energy access is critically low, with nearly 600 million people in sub-Saharan Africa lacking reliable electricity, resulting in an access rate of just over 40 %, the lowest globally (IEA, 2019; AfDB, 2018). This energy deficit severely affects manufacturing, as 80 % of sub-Saharan companies experience frequent disruptions, leading to substantial economic losses. Furthermore, over 900 million individuals rely on traditional fuels for cooking, which contributes to approximately 500,000 premature deaths annually and exacerbates environmental degradation (IEA, 2019). In terms of transportation, road access is only 34 %, significantly lower than the 50 % average in other developing regions, and transportation costs are 100 % higher, hindering economic efficiency (AfDB, 2018; Union Africaine, 2014). The density of paved roads in Africa is less than a quarter of that in other low-income regions, further impeding connectivity (Foster & Briceño-Garmendia, 2010). Water and sanitation infrastructure are also underdeveloped; as of 2015, only 36 % of Africans had access to improved sanitation, and just 63 % had access to basic drinking water services, lagging far behind other regions (AfDB, 2018). In sub-Saharan Africa, more than half of urban population lack access to basic sanitation while close to 20 % does not have access to safe water, proper drainage and waste management (Rasiah et al., 2022). Lastly, the telecommunications sector demonstrates substantial shortcomings, with only a 6 % internet penetration rate in 2012 compared to 40 % in other developing areas, and a mobile penetration rate of 73 % in sub-Saharan Africa (Bank & Flagship, 2016; Union Africaine, 2014).

The urbanization–infrastructure nexus can be understood using three main theoretical frameworks. The political economy of urbanization emphasizes how expanding urban populations exert pressure on governments to improve infrastructure. As cities grow the demand for essential services intensifies, necessitating policy responses that reflect the political dynamics shaping infrastructure investments. The economies of scale theory posits that urbanization enables more efficient infrastructure provision. Higher population density allows for reduced costs per capita in delivering services, enhancing productivity and living standards. This perspective highlights the economic benefits of urban growth, underscoring the potential for well-planned urbanization to foster sustainable infrastructure development. Conversely, urban-bias theory critiques the preferential treatment urban areas receive regarding infrastructure investment, particularly in developing countries. This bias leads to rapid urban growth as rural populations migrate to cities seeking better opportunities, often resulting in infrastructural strains and increased urban poverty.

2.3. Gaps in the literature review

While existing studies often explore individual components of infrastructure, there is a noticeable lack of comprehensive integration with broader urbanization literature, particularly in African contexts. For example (Doc et al., 2007) revealed that urban sprawl significantly impacts water fluxes and the urban water balance. Similarly, (Kasraian et al., 2016) conducted a long-term study on the relationship between railway infrastructure and urbanization in the Randstad region of the Netherlands, showing how new railway stations followed existing urbanization patterns and later spurred further development. This, however, reflects a broader issue in the literature, where studies often focus on specific infrastructure components—such as transport or water—rather than adopting a comprehensive definition of infrastructure that includes water and sanitation, ICT, transport, and energy. To our knowledge, no research has investigated the relationship between urbanization and infrastructure holistically, especially in the African context, which is a significant gap this study aims to address.

Furthermore, (Maparu & Mazumder, 2021) explored the causal direction between transport infrastructure and urbanization during the post-liberalization era in India, indicating that the expansion and development of transport infrastructure contributed to urban growth, aligning with modernization theory. While these studies are valuable, they reflect a broader trend in the literature where the focus is typically limited to individual components, such as transport. Most studies on transport infrastructure and urbanization also focus on energy and environmental impacts (Wang et al., 2019). For instance, (Miatto et al., 2021) highlighted the environmental concerns related to the material demands of rapid urbanization in Hanoi.

Other studies also predominantly investigate infrastructure development through the lens of transport infrastructure. (Li, Song, & Chen, 2017) found that infrastructure development influences urban expansion, land prices, and housing prices in Chinese cities, while (Heshmati & Rashidghalam, 2020) emphasized the effects of various infrastructure components on urbanization in China.

The above discussion highlights that existing studies often fail to comprehensively cover all infrastructure components. Many adopt an urbanization-centric perspective, viewing infrastructure primarily as a catalyst for urban growth rather than exploring how urbanization influences infrastructure development. The reverse relationship, where urbanization affects infrastructure development, remains understudied. Our theoretical framework is inspired by (Boserup, 1981) whose examination of preindustrial development posited that larger populations facilitated infrastructure development. A basic political-economy perspective on infrastructure provision suggests that growing urban populations necessitate improved infrastructure.

This study aims to address this gap by investigating the impact of urbanization on various infrastructure elements using an infrastructure composite index. Furthermore, this research establishes a contextual framework specifically within African countries, which, to our knowledge, has not been explored comprehensively. We adopt a mixed methodology with the following specific objectives:

1. Investigate the relationship between infrastructure development and urbanization in African countries using a 2-steps GMM model.
2. Identify typical development patterns of urban infrastructure development with the selected sample using clustering-based model.
3. Provide representative set of synthetic data for simulation purposes.

3. Methodology

The proposed methodology is a two-step, mixed approach that first aims to study the relationship between infrastructure development and urbanization at the African country level. The secondary approach aims to simulate the development of urban infrastructures in African countries using a clustering-based probability distribution model.

3.1. African urbanization-infrastructure nexus

3.1.1. Data and methods

To assess the impact of urbanization on infrastructure development, we used balanced panel data from 42 African countries[1] for 2005–2021. Data were sourced from the World Development Indicators (WDI) and the African Development Bank (AfDB). The African Infrastructure Development Index served as the dependent variable, encompassing

---

[1] Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cape Verde, Cameroon, Central African Republic, Chad, Congo, Democratic Republic of the Congo, Ivory Coast, Djibouti, Egypt, Equatorial Guinea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Madagascar, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Rwanda, Senegal, South Africa, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe.

sub-indexes for ICT, transport, water and sanitation, and electricity. While recent studies, such as (Nchofoung et al., 2022), have utilized this index, most research on infrastructure has focused on roads and ICT (Z. Wang et al., 2019; Maparu & Mazumder, 2021; Pradhan et al., 2021). In this study, we also considered additional infrastructure indicators, specifically ICT, electricity, and water infrastructure, to provide a more comprehensive analysis. Our independent variables include the degree of urbanization and economic growth, with industrial production and trade openness included as control variables. Tables 1 and 2 provide the main variables and descriptive statistics for the period 2005–2021. Table 3 shows the correlation matrix, indicating no multi-collinearity issues, even though such issues are minimal with interactive estimators.

The estimation method to explain the links between infrastructure and urbanization, in addition to other defined variables is performed using the 2-step system Generalized Method of Moments (GMM). The GMM has recently been used by many researchers with infrastructure related studies (Asongu & Odhiambo, 2021; Nchofoung & Asongu, 2022; Saadaoui Mallek et al., 2024). Moreover, it offers several advantages over the traditional dynamic panel-data method, which has been noted to have certain limitations by its developers (Arellano & Bover, 1995; Blundell & Bond, 1998). The choice of this method developed by (Roodman, 2009) is motivated by several factors. First, our dependent variable is highly correlated with its first-period lag $LnInfras_{i,t-1}$, with a correlation coefficient of 0.9980, confirming the dynamic character of the model with persistence over all specifications. Second, our dataset has a time dimension of 17 years and a cross-sectional dimension of 42 countries. According to (Roodman, 2009), for GMM to be appropriately applied, the cross-sectional dimension should be greater than the time dimension, which is true for our data. Third, this technique helps address endogeneity issues arising from potential reverse causality between urbanization and infrastructure development by using lagged values of the variables as instruments. Fourth, the two-step system GMM performs well in terms of limiting the proliferation of instruments and controlling for cross-sectional dependence (Boateng et al., 2018).

The following equations in level (Eq. (1)) and first difference (Eq. (2)) summarize the standard system GMM estimation:

$$LnInfras_{it} = \beta_0 + \beta_1 LnInfras_{i,t-r} + \beta_2 LnUrb_{i,t} + \beta_2 LnGDP_{i,t} + \sum_{n=1}^{2} \delta_n S_{n,i,t-r} + \gamma_i + \mu_t + \varepsilon_{i,t} \tag{1}$$

$$LnInfras_{it} - LnInfras_{i,t-r} = \beta_1 \left(LnInfras_{i,t-r} - LnInfras_{i,t-2r} \right) + \beta_2 \left(LnUrb_{it} - LnUrb_{i,t-r} \right) + \beta_2 \left(LnGDP_{i,t} - LnGDP_{i,t-r} \right) + \sum_{n=1}^{2} \delta_n \left(S_{n,i,t-r} - S_{n,i,t-2r} \right) + \left(\mu_t - \mu_{t-r} \right) + \varepsilon_{i,t-r} \tag{2}$$

where $LnInfras_{it}$ is the composite infrastructure development index for country i at year t, β_0 is a constant, $LnUrb_{i,t}$ is the degree of urbanization for each country i at year t, $LnGDP_{i,t}$ entails per capita GDP growth rate for country i at year t, S is the vector of control variables (trade openness and industrial production), r equals the unity, μ_t is the time-specific constant, γ_i is the country effect, and $\varepsilon_{i,t}$ is the error term.

Table 1
Variable definitions.

Variable name	Definitions	Source
Infras	African Infrastructure Development Index (AIDI)	AfDB
Urb	Degree of urbanization	WDI
GDP	Per capita GDP growth rate	WDI
Trd	International Trade	WDI
IndProd	Industrial productivity	WDI
Inrnt	Internet penetration	WDI
MbSub	Cellphone penetration	WDI
FBSub	Telephone penetration	WDI
Sanit	Sanitation services	WDI
Wtr	Drinking water services	WDI
Elect	Electric power consumption	WDI

Note: WDI. World Development Indicators; AfDI. African Infrastructure Development Bank.

Table 2
Summary statistics for total panel (2005–2021).

	Obs.	Mean	S.D.	Min.	Max.
Infras	712	22.19	17.75	1.12	88.74
Urb	712	43.83	18.21	9.38	90.42
IndProd	712	$6.61*10^9$	$1.37*10^{10}$	$1.54*10^7$	$8.37*10^{10}$
Inrnt	712	17.83	19.18	0.22	88.13
MbSub	712	66.49	39.96	0.53	168.92
Wtr	712	85.29	9.68	48.06	100.00
Elect	712	72.00	21.22	11.80	1000.00
GDP	712	2447.94	2765.14	151.19	19,849.72
FBSub	712	72.00	21.22	11.80	100.00
Trd	712	73.13	37.81	22.24	348.00

Note. S.D. Standard deviation. Min. Minimum. Max. Maximum.

We use the Hansen-Sargan test to check the validity of the instruments. Additionally, we perform the AR(2) test to detect second-order serial correlation. If the null hypothesis is accepted in the AR(2) test, it indicates no evidence of autocorrelation. Similarly, a non-significant result in the Hansen J-test confirms that the instruments are appropriate.

3.1.2. Results and discussion

Table 4 presents the outcomes of the 2-step GMM estimation. Results show that urbanization has a positive and significant impact on composite infrastructure development index, indicating a substantial and positive role. This result is replicated when the infrastructure indicators of electricity, ICT and water are used. However, for the findings to be valid, there needed to be no residual autocorrelation of either the first or second order. Therefore, AR(2) test for second-order autocorrelation should have a P-value >10 %, indicating no evidence of autocorrelation. The null hypotheses of the Sargan and Hansen over-identification restriction tests for instrument validity should not be rejected if the *P*-value is higher than 10 %. Additionally, the null hypothesis of the Wald test for the model's overall significance should be rejected if the P-value is <10 %. Finally, the number of instruments is kept to be less than the number of cross-sections as recommended by (Roodman, 2009). Our results meet the above-cited criteria. For instrument tests, the Hansen test is preferred due to its robustness to heteroscedasticity and suitability for dynamic panel data models with many instruments.

Our study underscores the positive impact of urbanization on various aspects of infrastructure development, as evidenced by the composite infrastructure index, ICT, and electricity indicators. This finding is consistent with the broader literature, which often highlights the interconnectedness of urbanization and infrastructure improvements (Pradhan et al., 2021). The positive impact of urbanization on infrastructure development observed in our study is consistent with the findings of (Weinhold & Reis, 2001). Their research highlighted a strong positive correlation between urban population growth and infrastructure development. This suggests that as urban areas expand, they often enhance their infrastructure to support the growing population. This is particularly true as the growth of urban populations leads to increased demands for infrastructure development. This is the case for Africa given that the continent has shown one of the fastest trends in terms of population growth in the world (UN DESA, 2019).

Our results indicate that increases in urbanization are linked to improvements in infrastructure, as evidenced by the composite infrastructure index, ICT, and electricity indicators. This finding aligns well with recent trends in Africa's infrastructure development, where significant progress is largely attributed to the rapid expansion of the ICT sector (Kengdo et al., 2020). At the same time urban populations in

Table 3
Correlation matrix.

	Infras	*Urb*	*IndProd*	*Inrnt*	*MbSub*	*Water*	*Elect*	*GDP*	*FBSub*	*Trd*
Infras	1.000									
Urb	0.429***	1.000								
	(0.000)									
IndProd	0.539***	0.264*	1.000							
	(0.000)	(0.013)								
Inrnt	0.730***	0.482***	0.267***	1.000						
	(0.000)	(0.000)	(0.000)							
MbSub	0.642***	0.438***	0.228***	0.733***	1.000					
	(0.000)	(0.000)	(0.000)	(0.000)						
Wtr	0.675***	0.245***	0.299***	0.532***	0.598***	1.000				
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)					
Elect	0.609***	0.4546***	0.275***	0.591***	0.600***	0.696***	1.000			
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)				
GDP	0.209***	0.263***	0.087***	0.254***	0.108*	0.1443***	0.308***	1.000		
	(0.000)	(0.000)	(0.000)	(0.000)	(0.036)	(0.000)	(0.000)			
FBSub	0.714***	0.218***	0.214***	0.590***	0.440***	0.424***	0.398*	0.101**	1.000	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.0065)		
Trd	0.077*	0.406***	−0.118*	0.161***	0.022	0.129***	0.096**	0.1412***	0.122***	1.000
	(0.039)	(0.000)	(0.001)	(0.000)	(0.554)	(0.000)	(0.010)	(0.000)	(0.001)	

Note: Standard errors in brackets.

* $p < 10$ %.

** $p < 5$ %.

*** $p < 1$ %.

Table 4
GMM regression results for the period (2005–2021).

Dependent variable	*Infras*					
	AIDI	*ICT*			*Elect*	*Wtr*
	–	*Inrnt*	*MbSub*	*FBSub*	–	–
$LnInfras_{i,t-r}$	0.922***(0.007)	0.957***(0.005)	0.784***(0.006)	0.509***(0.236)	0.585***(0.235)	1.047***(0.003)
LnUrb	0.023**(0.022)	0.0341**(0.039)	−0.077***(0.023)	1.915***(0.207)	0.392***(0.029)	−0.038***(0.03)
LnGDP	0.081***(0.006)	0. 099***(0.013)	0.036**(0.014)	0.422***(0.126)	0.049***(0.004)	0.009***(0.001)
LnIndProd	−0.002**(0.001)	0.022***(0.003)	0.010***(0.001)	0.220***(0.0463)	−0.004(0.002)	−0.0002(0.0009)
LnTrade	0.043***(0.007)	0.0008*(0.003)	0.092***(0.012)	−1.383***(0.954)	−0.145***(0.012)	0.003***(0.003)
constant	−0.555***(0.084)	−1.122***(0.197)	0.397***(0.107)	−7.574***(0.954)	0.675***(0.129)	−0.146***(0.010)
AR(2)	[0.179]	[0.956]	[0.124]	[0.347]	[0.189]	[0.352]
Sargan	[0.092]	[0.215]	[0.379]	[0.378]	[0.114]	[0.305]
Hansen	[0.183]	[0.187]	[0.060]	[0.511]	[0.413]	[0.256]
Instruments	22	35	35	22	35	35
Wald test	101,294.11***	103,731.30***	65,224.61***	1625.05***	2755.11***	399,657.51***
Obs.	712					
Groups	42					

Notes: Figures between parenthesis () signify to the standard error, whereas figure between brackets [] denote to the probability value. *, **, ***: significance levels of 10 %, 5 %, and 1 %, respectively.

Africa continue to rise, reaching 29 % between 2018 and 2030 and 44 % between 2030 and 2050 (United Nations, 2019). Investments in infrastructures in Africa have contributed between 14 % and 48 % to the current economic growth and have the potential to do even more (AfDB, 2018).

Similarly, urbanization positively influence electricity infrastructure. A 10 % increase in the degree of urbanization can increase electricity consumption by 3.92 %. This indicate that higher levels of urbanization are associated with substantial improvements in electricity infrastructure. This aligns with the expectation that as urban areas expand, there is a greater demand for reliable and enhanced electricity services to support the growing population and economic activities.

Despite this positive trend, it is important to note that per capita electricity demand across the continent remains very low. Currently, it accounts for <200 kWh per capita annually and is projected to be just over 430 kWh by 2040 (IEA, 2019). This still represents <15 % of the world average. The challenges of inadequate electricity provision can be traced back to historical factors, including colonial institutional arrangements and ongoing political instability (Güneralp et al., 2018). This historical context underscores the significant gap in electricity provision that needs to be addressed to meet the increasing demands driven by urbanization. Interestingly, there is a scarcity of studies focusing specifically on electricity infrastructure within the context of urban or city planning. Often, electricity consumption is analyzed alongside other infrastructure components, such as building use or electrified transportation modes, in demand-centered analyses (Howard et al., 2012).

On the other hand, GMM results indicate a strong positive relationship between economic growth and various infrastructure developments, which is consistent with (Palei, 2015). These results suggest that as the economy grows, there are significant improvements in infrastructure, particularly in ICT (internet and mobile subscriptions) and broadband (fixed broadband subscriptions). Electricity infrastructure also benefits notably from economic growth, though the effect on water infrastructure is comparatively smaller. Given that the relationship between infrastructure and economic growth is context-dependent (Holtz-Eakin & Schwartz, 1995), this study provides clear evidence in the African context, where infrastructure is often under-supplied. Economic growth consistently promotes infrastructure development across all categories. These results underscore the importance of economic growth as a driver of infrastructure development in Africa.

With regards to control variables, industrial production supports ICT infrastructure but has limited effects on electricity and water infrastructure. Particularly, a 10 % increase in industrial production is associated with a 2.2 % increase in mobile subscriptions, indicating that industrial growth supports the expansion of ICT infrastructure. This is likely due to the increased demand for communication technologies to support industrial activities. On another side, trade openness seems to have complex relationship with different infrastructure components. Trade openness positively influences the composite infrastructure index and ICT development, which is supported by existing literature (Nchofoung & Asongu, 2022). However, our results also indicate that, in the African context, trade openness has a negative impact on electricity infrastructure, which appears inconsistent with the same study's findings. This discrepancy underscores the need for further investigation into the nuanced effects of trade openness on different infrastructure types.

3.2. Simulation and clustering analysis in African urban infrastructures development

After conducting a comprehensive analysis to understand the factors influencing urban infrastructure in African countries, the significance of variables, such as urbanization, economic growth, trade openness and industrial production, has become evident. While the GMM estimation provides valuable insights into the relationships between these factors and infrastructure development, extending our analysis to include simulation of urban infrastructure development is equally important.

Simulation allows us to model and explore existing patterns, offering a forward-looking perspective on how various factors may interact and shape infrastructure development over time. Clustering analysis will be employed to identify distinct patterns and groupings among the studied countries studies, yielding a more nuanced understanding of regional disparities. Subsequent to this, distribution fitting will be instrumental in creating synthetic curves that reflect the intrinsic dynamics of infrastructure development. In summary, the clustering-based probability distribution model offers added value by capturing heterogeneity, pinpointing non-linear relationships, generating synthetic datasets, offering detailed insights, and facilitating spatial and temporal explorations, which traditional regression-based methods may not fully encompass.

3.2.1. Modeling framework and building blocks

The establishment of the model entails three distinct stages: feature extraction, two-step cluster analysis, and distribution fitting.

3.2.1.1. Features definition. Data of urban infrastructure were collected over a specified 16-year timeframe for each of the 45 African countries. To characterize urban infrastructure development in these countries, the following feature are extracted:

- μ: Represents the average infrastructure development index, offering a measure of the country's overall infrastructure progress within the study period.
- CV $= \sigma/\mu$: The coefficient of variation of the annual infrastructure development index. This quantifies the variability of infrastructure development across a consistent 17-year span within a single country.
- σ: The standard deviation of the annual infrastructure development index.
- $(I_1, I_2, \ldots, I_{17})$ A sequence representing the infrastructure development patterns, reflecting the behavior of a single country's infrastructure development trajectory over time.

For normalization and standardization of each data point within the dataset, the z-score normalization is used. The z-score standardization (Mustoe & Walker, 1977) is given by:

$$I_i = (x_i - \mu)/\sigma \tag{3}$$

3.2.1.2. Clustering analysis. In this analysis, a two-step clustering procedure is introduced, based on the infrastructure development index. The initial step involves clustering all countries based on the yearly average urban infrastructure development index μ and the standard deviation σ of the yearly infrastructure index. Subsequently, the second step of clustering is conducted within each cluster formed in the first step, with a focus on infrastructure development patterns $(I_1, I_2, \ldots, I_{17})$.

The present study uses the most established category of unsupervised learning, called clustering, which aims at grouping instances based on their degree of similarity (Xu & Wunsch, 2005). Several clustering algorithms have been developed over the past decades with K-means being the most well-known and widely used algorithms for clustering (Hartigan & Wong, 1979), given its ease of application, efficiency and simplicity. K-means is a portioning clustering algorithm, that is, it aims to find all the clusters simultaneously as a partition of the data and does not impose a hierarchical structure. K-means necessitates a pre-defined number of clusters.

The within-cluster-sum-of-squares (WCSS) is first used to determine the optimum number of clusters, where the highest drop in the value of WCSS corresponds to the optimum K value. The quality of clustering is then evaluated using the Silhouette score, graphical aid that evaluates how similar or dissimilar data points are within one another and from other clusters to determine the quality of a cluster. Higher numbers indicate better cluster separation; values range from −1 to 1 (Rousseeuw, 1987). The Caliński-Harabasz (CH) index, introduced in 1974, is a widely used metric for evaluating and comparing clustering methods on the same dataset (Calinski & Harabasz, 1974). It quantifies the ratio of within-cluster to between-cluster dispersion, providing a concise measure of clustering effectiveness. Finally, The Davies-Bouldin (BD) Index (Davies & Bouldin, 1979) is also used to evaluate the quality of clustering, as seen in Table 5.

3.2.1.3. Probability distribution fitting. To generate synthetic data that will be used to simulate infrastructure development in African countries, we perform distribution fitting after the clustering analysis. Within each sub-cluster, samples consist of nineteen feature parameters, which include the annual average infrastructure development (μ), the coefficient of variation (CV) of the annual infrastructure development index, and seventeen infrastructure development patterns. To fit all 19 parameters, we employ distinct probability distributions for each

Table 5
Metrics used for evaluating clusters quality.

Formula	Parameters	Interpretation
$CH(K) = \frac{Tr(S_B) \times (N-K)}{Tr(S_W) \times (K-1)}$ (1)	N: total number of samples K: number of clusters $Tr(S_B), Tr(S_W)$: trace of the between and within class scatter matrix, respectively. $CH(K)$: Caliński-Harabasz Index	The K that maximizes the value of CH is selected as the optimal. Higher values indicate better clusters.
$BD(K) = \frac{1}{K}\sum_{i=1}^{k} \underset{i \neq j}{max} \left(\frac{\delta_i + \delta_j}{d(c_i + c_j)}\right)$ (2)	K: number of clusters c_x: the center of cluster x δ_x: the average distance between any data in cluster x and c_x $d(c_i, c_j)$: the distance between c_i and c_j	Lower values indicate better clusters. lower values of the Davies-Bouldin Index indicate better-defined clusters with good separation and minimal overlap

parameter, allowing the distributions to correspond to each sub-cluster.

The $(I_1, I_2, \ldots, I_{17})$ infrastructure development patterns data are influenced by a variety of factors and exhibit a symmetric balanced distribution. The Gaussian distribution, which is a type of distribution that is symmetric about the mean, is used (DeGroot & Schervish, 2011). The probability density function is shown in Eq. (4), in which σ is the standard deviation of the distribution and μ is the mean of the distribution:

$$f(x|\mu, \sigma_d) = \frac{1}{(2\pi\sigma_d)^{1/2}} exp\left(-\frac{(x-\mu_d)^2}{2\sigma_d^2}\right) \tag{4}$$

As for fitting the mean μ and the coefficient of variation CV of infrastructure development index, log-normality is plausible (Devoire, 2011). The logarithmic normal distribution equation is shown below:

$$f(x|\mu, \sigma_d) = \left\{ \begin{matrix} 0 \; x > 0 \\ \frac{1}{(2\pi\sigma_d)^{1/2}} exp\left(-\frac{(x-\mu_d)^2}{2\sigma_d^2}\right) \; x \leq 0 \end{matrix} \right\} \tag{5}$$

To represent the data from each sub-cluster, nineteen distributions are employed. Given the multitude of sub-clusters, the model for the entire sample is comprehensive, encompassing the total number of sub-clusters and the respective distributions for all parameters.

3.2.2. *Model regeneration: a statistical perspective*

The regeneration process involves generating random feature parameters based on the fitted probability distributions for each sub-cluster. These random feature parameters are then used to reconstruct the infrastructure development patterns, allowing for the generation of synthetic infrastructure development patterns that reflect the statistical characteristics of the original data. The reconstruction process involves using the feature parameters to revert z-score normalization and recreate the infrastructure development patterns from the feature parameters. This synthetic data will be utilized for simulating urban infrastructures development in African countries.

This simulation approach serves as a valuable complement to the regression analysis, enabling us to not only understand the current determinants but also project potential trajectories and trends. By leveraging clustering and distribution fitting techniques, we aim to create synthetic representations of urban infrastructure development, providing a tool for scenario planning, policy evaluation, and strategic decision-making in the dynamic landscape of African urban development.

3.2.3. *Study case*

3.2.3.1. Modeling and clustering analysis. The predefined features are used to describe each urban infrastructure development in each country. The clustering is performed using the k-means clustering algorithm in two steps.

A sample of 45 countries is first grouped based on infrastructure development. Two clusters are formed in this first step. Then a second-step clustering is performed based on infrastructure development patterns. Four sub-clusters are formed in this step. Results of the two steps are shown in Fig. 1.

The quality of a cluster formation is evaluated using Caliński-Harabasz (C-H) index and Davies-Bouldin index, see Table 6.

Statistical characteristics of the resulting clusters are represented in Table 7. The findings indicate that 46.66 % of the analyzed African countries exhibit an average infrastructure index below 11, with 37.77 % showing an index below 22. Among the countries analyzed, just 6.66 % have a mean value of 66.85, while 4, or 8.88 % of the total, have an index of 46.24.

Among 45 countries' infrastructure development tracks, four typical patterns have emerged, as shown in Fig. 2.

3.2.3.2. Distribution fitting. For each of the four sub-clusters, a lognormal was used to fit both the coefficient of variation CV and the mean μ of the yearly infrastructure development. The fitting results of

Table 6
Cluster quality metrics.

Evaluation metric	First-step clustering	Second-step clustering
Caliński-Harabasz (C-H) index	162.69	0.398
Davies-Bouldin Index	202.71	0.53

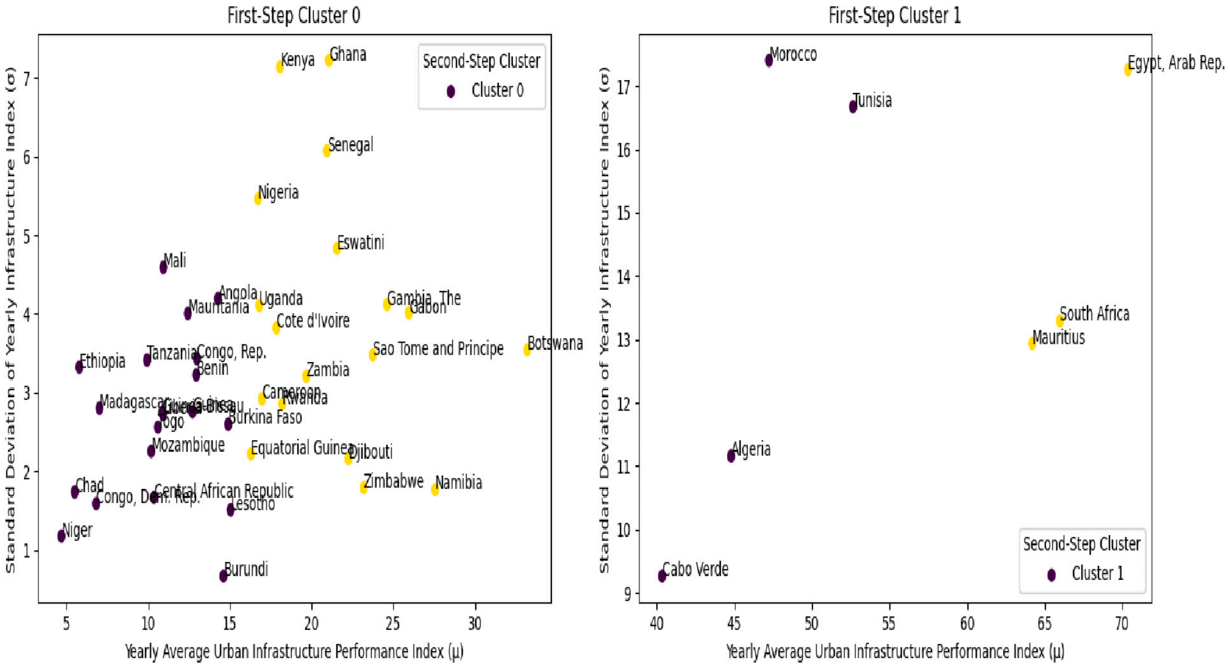

Fig. 1. Countries per cluster distribution.

Table 7
First and second step clusters' statistics.

Clusters N°	Yearly infrastructure development index	Standard deviation of yearly development infrastructure index	Percentage & count of countries
0	15.74	6.62	84.44 % (38)
1	55.08	11.73	15.56 % (7)
0–2	10.94	3.38	46.66 % (21)
0–3	21.67	4.44	37.77 % (17)
1–0	46.24	5.12	8.88 % (4)
1–1	66.85	3.17	6.66 % (3)

CV and μ in cluster 0–2 are displayed in Fig. 3, where the x-axis shows the value of the fitted variable, the bars show the frequency of the samples, and the curve shows the probability density function of the fitted distribution.

Normal probability distributions were used for infrastructure patters $(I_1, I_2, \ldots, I_{17})$ fitting. The results of fitting pattern parameters to a normal distribution for each year in the sub-cluster 0–2 are shown in Fig. 4. This model includes 17 probability distributions for each of the four sub-clusters from the original sample sets. The composition of this sample set is represented by the ratio of the number of countries in each sub-cluster to the overall number of countries.

This model can statistically replicate infrastructure development

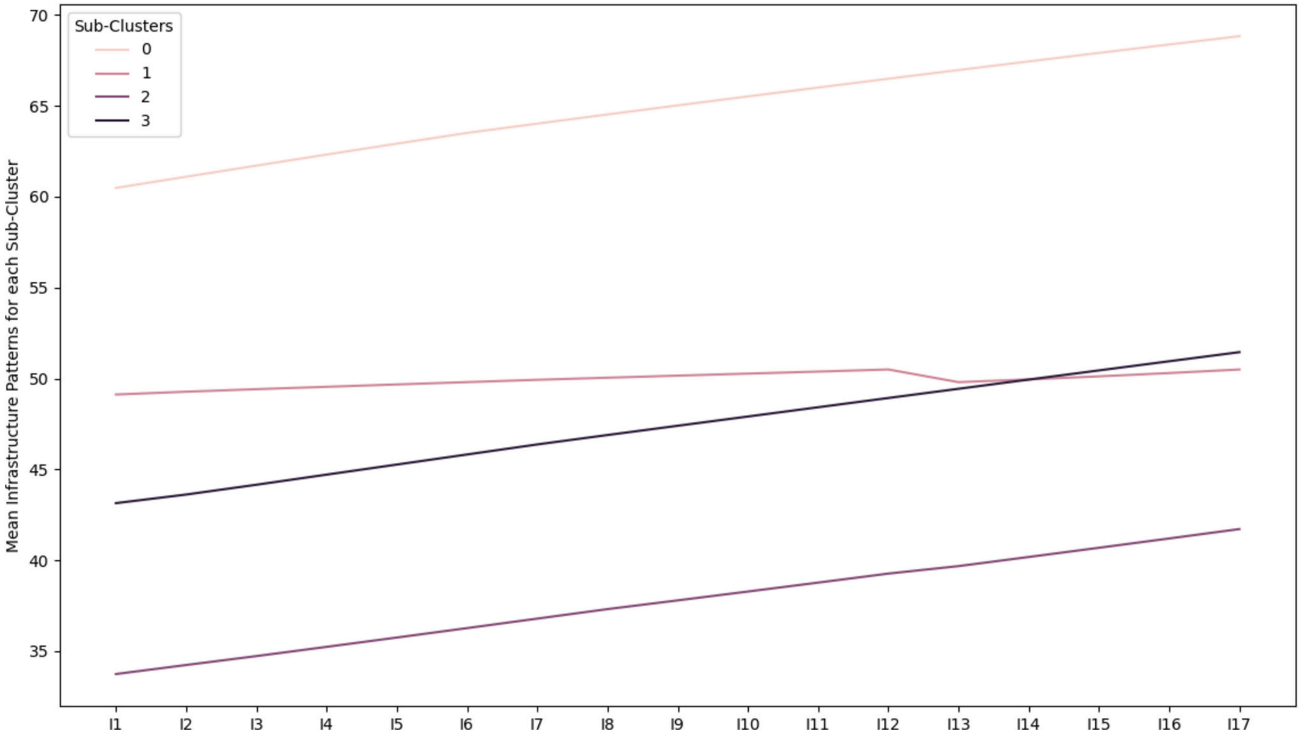

Fig. 2. Infrastructure development patterns for each resulted cluster.

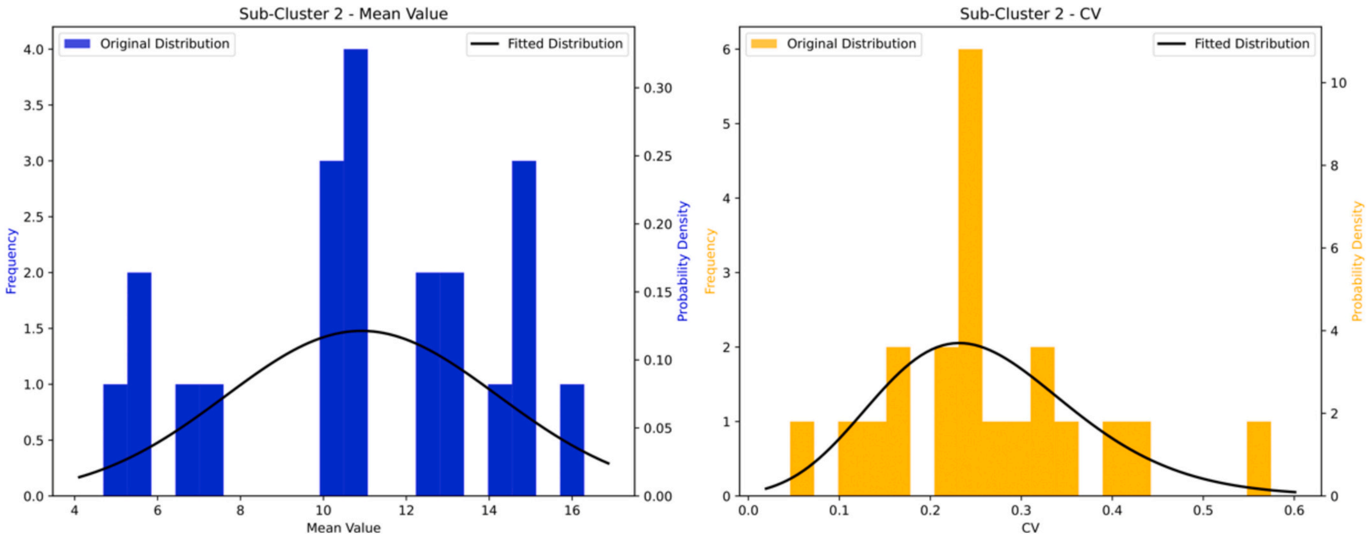

Fig. 3. Log-normal Distribution Fitted Results of CV and μ in cluster 0–2.

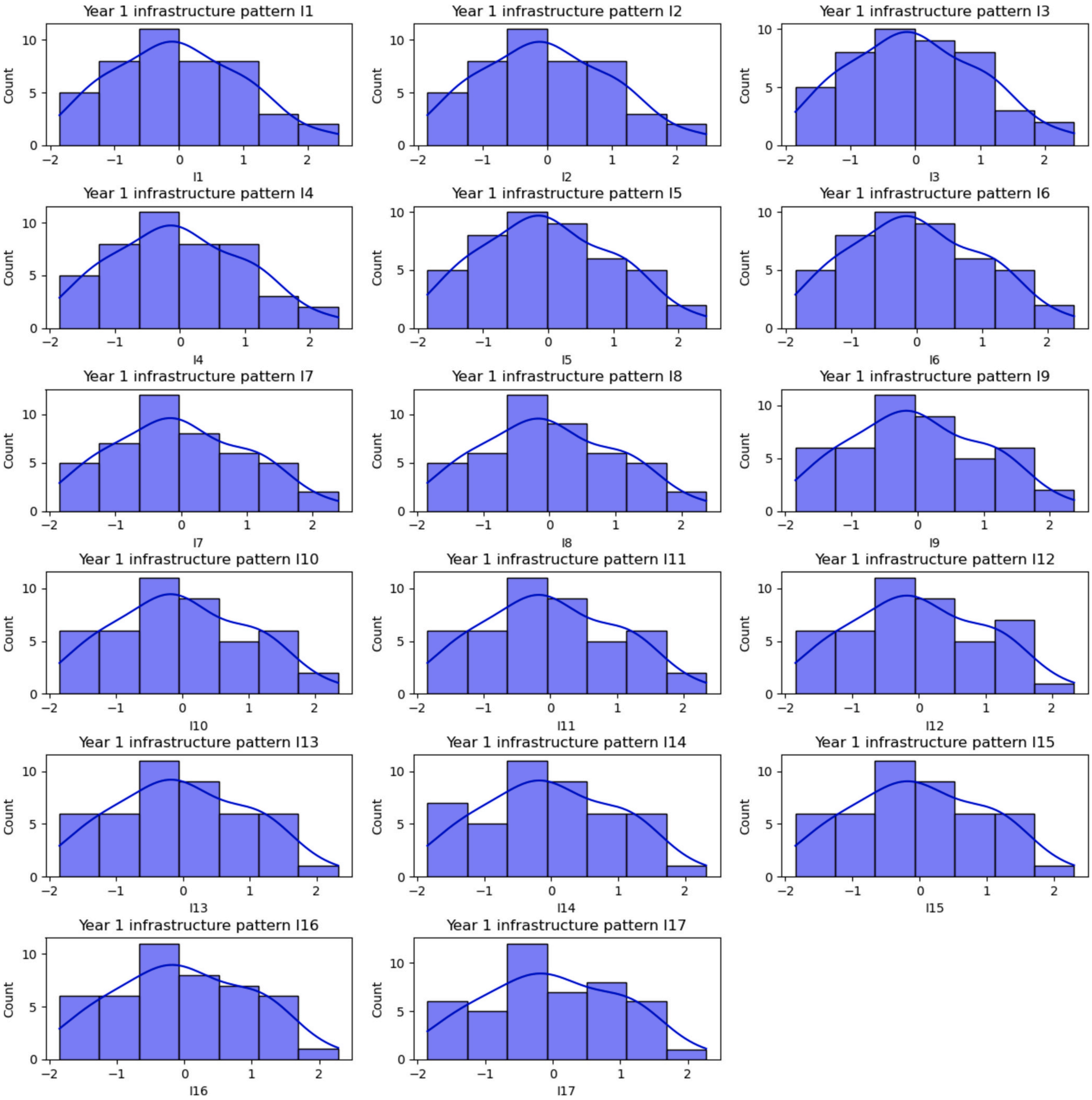

Fig. 4. Normal distribution fitted results for pattern parameters $(I_1, I_2, \ldots, I_{17})$.
The curves represent the fitted PDF, while the bars show the dataset's original distribution.

patterns based on the sub-cluster distributions and ratios. The development curves for synthetic infrastructure can be produced using the previously discussed method.

3.2.3.3. Model validation. The modeling procedure, which involves transitioning from infrastructure development data to the probability distribution, as well as the regeneration procedure, which is the reverse operation, requires a means of validation. To validate these two procedures, we will use two statistical parametric tests.

The Anderson-Darling test and Kolmogorov-Smirnov test are statistical tests used to assess the goodness of fit of simulated data to specific probability distributions. Both tests are widely used in various fields to validate simulation models and assess the accuracy of the simulated data. The Anderson-Darling test measures the discrepancy between the empirical distribution function of the simulated data and the theoretical distribution function, while the Kolmogorov-Smirnov test evaluates the maximum difference between the empirical distribution function and the reference distribution function. The expected results of the Anderson-Darling and Kolmogorov-Smirnov tests for fitting validation would be based on the test statistic and the corresponding *p*-value. The Kolmogorov-Smirnov and Anderson-Darling tests are performed on every distribution fitting procedure across every sub-cluster. Here, we illustrate the outcome using the sub-cluster 0–2 as an example. If the p-value is >0.05, the Anderson-Darling and Kolmogorov-Smirnov tests indicate an acceptable fit at the 95 % confidence level. All feature parameters passed the Anderson-Darling and Kolmogorov-Smirnov tests, as indicated in Figs. 5 and 6.

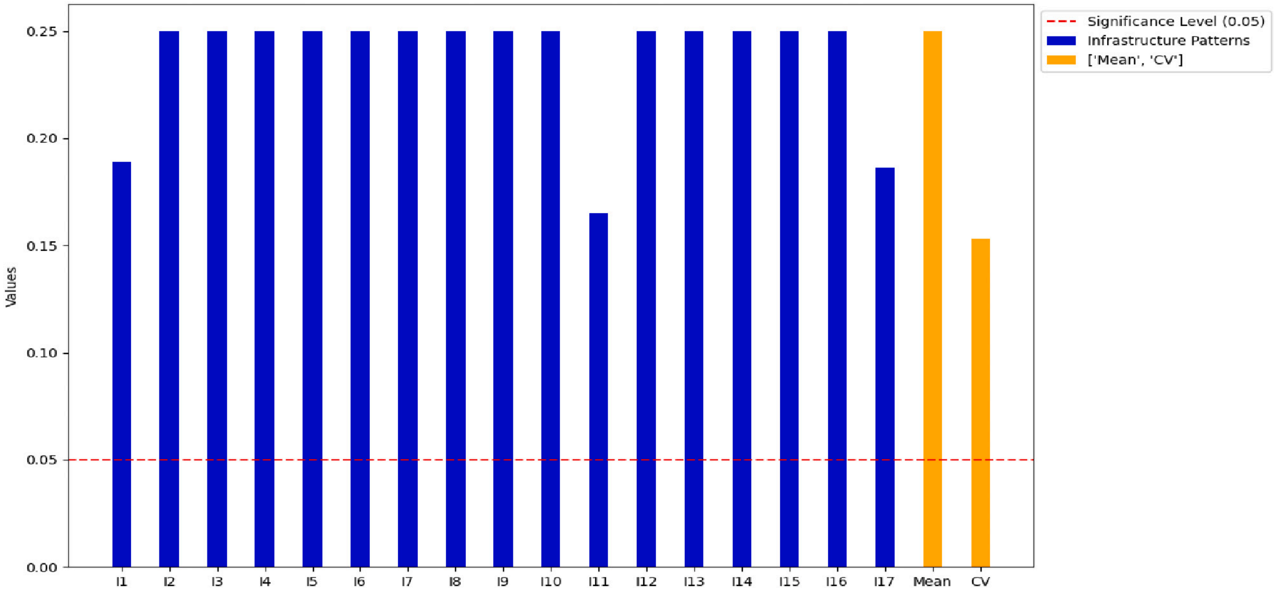

Fig. 5. Anderson-Darling Test Results for Sub-Cluster 0–2.

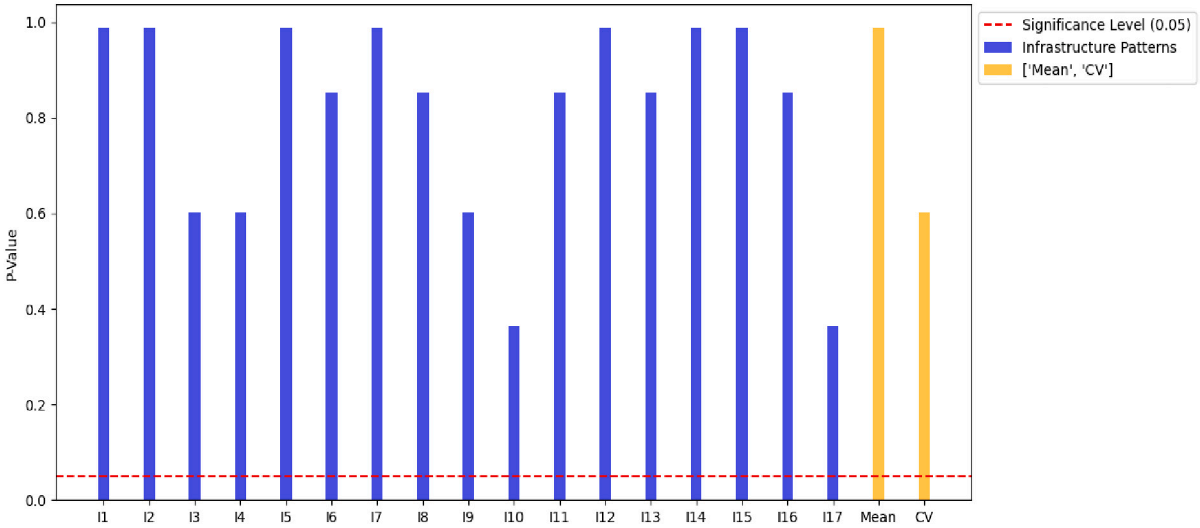

Fig. 6. Kolmogorov-Smirnov Test Results for Sub-Cluster 0–2

4. Concluding insights and future directions

This study underlined that urbanization and economic growth are crucial determinants of infrastructure development in the context of Africa, with varying impacts across different types of infrastructure. It examined the effects of urbanization on different infrastructure measures using panel dataset comprising 42 African during the period 2005–2021, using 2-step GMM technique. Four measures of infrastructure development were considered in the analysis: composite infrastructure index, electricity, water, and ICT. The empirical results reveal the following key findings: (i) Urbanization's impact on infrastructure development in Africa is multifaceted. Generally, urbanization positively influences the overall infrastructure index, electricity infrastructure, internet penetration, and fixed broadband subscriptions. However, it exerts a negative impact on water infrastructure and mobile subscriptions, indicating that urbanization affects different types of infrastructure in varied ways. (ii) Economic growth emerges as a consistent driver of infrastructure development across all categories, playing a crucial role in expanding ICT and electricity infrastructure. (iii) In contrast, the effects of industrial production are mixed; it generally promotes ICT infrastructure but shows either negative or insignificant impacts on other types of infrastructure. (iv) Trade openness has a dual influence, positively affecting overall infrastructure development and most ICT components while negatively impacting fixed broadband subscriptions and electricity infrastructure. These findings highlight the complex and varied role of urbanization and other factors in shaping infrastructure development in the African context.

The study reveals that urbanization and economic growth significantly influence infrastructure development in Africa, but with varying effects. To address the negative impact of urbanization on water infrastructure and mobile subscriptions, policymakers should prioritize expanding and modernizing these systems to keep pace with urban growth. Economic growth consistently boosts ICT and electricity infrastructure, so targeted investments in broadband expansion, mobile networks, and electrical grid modernization are essential. The mixed effects of industrial production on infrastructure suggest a need for policies that support ICT-focused industrial growth while addressing negative impacts on other infrastructure types. Additionally, while trade openness generally promotes infrastructure development, its adverse effects on fixed broadband and electricity infrastructure require balanced trade policies to mitigate these challenges. Finally, the study underscores the importance of developing resilient infrastructure capable of adapting to urban growth and environmental changes through regular assessments and sustainable practices. From the perspective of infrastructure development and investment, Africa has not reached the turning point of the relationship between urbanization and infrastructure development. Consequently, as infrastructure investment gradually increases, along with ongoing urbanization trends, the influence of urbanization on infrastructure development expected to grow, and the regional disparities within Africa may also widen.

As half of urban land in existence in 2030 is yet to be developed (Elmqvist et al., 2013), policymakers must integrate infrastructure planning with urban growth strategies to ensure that infrastructure development keeps pace with urbanization. This includes planning for future demand, upgrading existing infrastructure, and ensuring that new developments are supported by adequate services. Given the study's finding that urbanization can exacerbate infrastructure gaps, it is crucial to develop resilient infrastructure to manage urban growth effectively. This entails conducting regular assessments to identify and address emerging needs and investing in infrastructure that can withstand increased demand and adapt to environmental changes by reinforcing existing structures and incorporating flexibility in new projects.

The subsequent phase of this study involves using a clustering-based probability distribution model, which offers the benefit of capturing heterogeneity, identifying non-linear relationships, and generating synthetic data—features that complement but are not fully realizable through GMM. The primary goal of this phase is to categorize countries with similar infrastructure development trends and fit probability distributions to capture the statistical properties of these patterns within each sub-cluster. This approach allows for the creation of random feature parameters and regeneration of synthetic data that retain comparable statistical characteristics to the observed data. In our case study, two main clusters of infrastructure development levels were identified, with four sub-clusters of development patterns. Corresponding probability distributions were fitted for each feature parameter, reflecting the model's results. This model aids in regenerating infrastructure data from acquired knowledge, extending the research to both spatial and temporal dimensions.

5. Limitations and future recommendations

Transport infrastructure indicators were not included in the analysis due to unavailable data for the studied countries, although they are indirectly captured in the composite infrastructure index; thus, the impact of urbanization on transport infrastructure as a standalone component is not addressed in this study.

Another key limitation is the variation in urbanization levels across Africa. This study treats the continent as a whole, despite significant differences in urbanization rates between regions, such as the more urbanized Northern and Southern Africa compared to the rapidly growing West and East Africa. This broad approach, while providing general insights, risks over-generalization and may not fully account for regional variations in urbanization dynamics. Additionally, for GMM to be appropriately applied, the cross-sectional dimension should be greater than the time dimension, which is true for our dataset. However, dividing the dataset into two sub-periods (2005–2012 and 2012–2021), for instance, would provide a better basis for analysis, as it would ensure that the cross-sectional dimension is significantly greater than the time dimension.

CRediT authorship contribution statement

Rachida El-Bouayady: Writing – original draft, Methodology, Data curation, Conceptualization. **Hassan Radoine:** Writing – review & editing, Validation, Conceptualization. **Nour-Eddin El Faouzi:** Writing – review & editing, Validation, Methodology, Formal analysis. **Soukaina Tayi:** Writing – original draft, Methodology, Formal analysis. **Hakki Can Ozkan:** Writing – review & editing, Validation, Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

We would like to thank the anonymous two reviewers and the editor of this journal for their valuables inputs and insights.

References

AfDB. (2018). African economic outlook. In *Annual yearly review study*. https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/African_Economic_Outlook_2018_-_EN.pdf.

Andrianaivo, M., & Kpodar, K. (2012). Mobile phones, financial inclusion, and growth. *Review of Economics and Institutions, 3*(2). https://doi.org/10.5202/rei.v3i2.75

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29–51. https://doi.org/10.1016/0304-4076(94)01642-D

Aschauer, D. A. (1989). Is public expenditure productive. *23*(September 1988), 177–200.

Asongu, S. A., & Odhiambo, N. M. (2021). Enhancing governance for environmental sustainability in sub-Saharan Africa. *Energy Exploration and Exploitation, 39*(1), 444–463. https://doi.org/10.1177/0144598719900657

Baker, & Judy, L. (2012). Climate change, disaster risk, and the urban poor. In *Climate change, disaster risk, and the urban poor*. World Bank Publications. https://doi.org/10.1596/978-0-8213-8845-7.

Bank, W., & Flagship, G. (2016). World Development Report, 2016: digital dividends. *53* (11). https://doi.org/10.5860/choice.196952, 53-4889-53–4889.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115–143. https://doi.org/10.1016/j.jeconom.2023.03.001

Boateng, A., Asongu, S., Akamavi, R., & Tchamyou, V. (2018). Information asymmetry and market power in the African banking industry. *Journal of Multinational Financial Management, 44*, 69–83. https://doi.org/10.1016/j.mulfin.2017.11.002

Boserup, E. (1981). *Population and technology*. Oxford: Blackwell.

Calderón, C., & Servén, L. (2004). *The effects of infrastructure development on growth and income distribution*. Available at SSRN 625277.

Calderón, C., & Servén, L. (2010). Infrastructure and economic development in Sub-Saharan Africa. *Journal of African Economies, 19*(Supplement 1). https://doi.org/10.1093/jae/ejp022

Calinski, T., & Harabasz, J. (1974). Communications in Statistics A dendrite method for cluster analysis. *Communications in Statistics, 3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Chinoda, T., & Kwenda, F. (2019). Do mobile phones, economic growth, bank competition and stability matter for financial inclusion in Africa? *Cogent Economics and Finance, 7*(1), 1–20. https://doi.org/10.1080/23322039.2019.1622180

Collier, P., & Venables, A. J. (2016). Urban infrastructure for development. *Oxford Review of Economic Policy, 32*(3), 391–409. https://doi.org/10.1093/oxrep/grw016

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 2*, 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

DeGroot, M. H., & Schervish, M. J. (2011). *Probability and statistics*. Addison-Wesley.

Devoire, J. (2011). *Probability and statistics for engineering and the sciences* (8th ed.). Brooks/Cole.

Doc, C. F., Simply, D., The, W., & On, F. (2007). Does urban sprawl drive changes in the water balance and policy_the case of Leipzig (Germany) 1870–2003. *Landscape and Urban Planning*, 1–13.

El-Bouayady, R., & Radoine, H. (2023). Urbanization and sustainable urban infrastructure development in Africa. *Environment and Ecology Research, 11*(2), 385–391. https://doi.org/10.13189/eer.2023.110212

Elmqvist, T., Fragkias, M., Goodness, J., & Güneralp, B. (2013). *Urbanization, biodiversity and ecosystem services: Challenges and opportunities: A global assessment*. Springer Nature.

Foster, V., & Briceño-Garmendia, C. (2010). *Africa's infrastructure: A time for transformation*.

Gross, J., & Ouyang, Y. (2021). Types of urbanization and economic growth. *International Journal of Urban Sciences, 25*(1), 71–85. https://doi.org/10.1080/12265934.2020.1759447

Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., & Seto, K. C. (2018). Urbanization in Africa: Challenges and opportunities for conservation. *Environmental Research Letters, 13*(1). https://doi.org/10.1088/1748-9326/aa94fe

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*(1), 100–108. https://doi.org/10.2307/2346830

Henderson, V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic Growth, 8*(1), 47–71. https://doi.org/10.1023/A:1022860800744

Heshmati, A., & Rashidghalam, M. (2020). Measurement and analysis of urban infrastructure and its effects on urbanization in China. *Journal of Infrastructure Systems, 26*(1). https://doi.org/10.1061/(asce)is.1943-555x.0000513

Hirschman, A. O. (1958). *The strategy of economic development* (p. 83). New Havens: Yale University Press.

Holtz-Eakin, D., & Schwartz, A. E. (1995). Infrastructure in a structural model of economic growth. *Regional Science and Urban Economics, 25*(2), 131–151. https://doi.org/10.1016/0166-0462(94)02080-Z

Howard, B., Parshall, L., Thompson, J., Hammer, S., Dickinson, J., & Modi, V. (2012). Spatial distribution of urban building energy consumption by end use. *Energy and Buildings, 45*, 141–151. https://doi.org/10.1016/j.enbuild.2011.10.061

IEA. (2019). Africa energy outlook 2019. In *World Energy Outlook special report*. IEA. https://www.iea.org/reports/africa-energy-outlook-2019#energy-access%0Ahttps://www.iea.org/reports/africa-energy-outlook-2019#africa-case.

Journal of Infrastructure Systems. (2024). Aims & scope of journal of infrastructure systems. http://ascelibrary.org/page/jitse4/editorialboard.

Kasraian, D., Maat, K., & van Wee, B. (2016). Development of rail infrastructure and its impact on urbanization in the Randstad, the Netherlands. *Journal of Transport and Land Use, 9*(1), 151–170. https://doi.org/10.5198/jtlu.2015.665

Kengdo, A. A. N., Nchofoung, T. N., & Ntang, P. B. (2020). Effect of external debt on the level of infrastructure in Africa. *Economic Bulletin, 40*(4).

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy, 99*(3), 483–499. https://doi.org/10.1086/261763

Kumar, A., & Kober, B. (2012). Urbanization, human capital, and cross-country productivity differences. *Economics Letters, 117*(1), 14–17. https://doi.org/10.1016/j.econlet.2012.04.072

Li, C., Song, Y., & Chen, Y. (2017). Infrastructure development and urbanization in China. *China's Urbanization and Socioeconomic Impact*, 1–227. https://doi.org/10.1007/978-981-10-4831-9

Li, Y., Zheng, J., Li, F., Jin, X., & Xu, C. (2017). Assessment of municipal infrastructure development and its critical influencing factors in urban China: A FA and STIRPAT approach (PLoS ONE (2017) 12:8 (e0181917) DOI: 10.1371/journal.pone.0181917). *PLoS One, 12*(12), 1–14. https://doi.org/10.1371/journal.pone.0189530

Liddle, B., & Messinis, G. (2015). Which comes first—Urbanization or economic growth? Evidence from heterogeneous panel causality tests. *Applied Economics Letters, 22*(5), 349–355.

Maparu, T. S., & Mazumder, T. N. (2021). Investigating causality between transport infrastructure and urbanization: A state-level study of India (1991–2011). *Transport Policy, 113*(March 2020), 46–55. https://doi.org/10.1016/j.tranpol.2020.03.008

McFarlane, C. (2011). The city as assemblage: Dwelling and urban space. *Environment and Planning D: Society and Space, 29*(4), 649–671.

Miatto, A., Dawson, D., Nguyen, P. D., Kanaoka, K. S., & Tanikawa, H. (2021). The urbanisation-environment conflict: Insights from material stock and productivity of transport infrastructure in Hanoi, Vietnam. *Journal of Environmental Management, 294*(November 2020). https://doi.org/10.1016/j.jenvman.2021.113007

Momoh, J. (2016). Sustainable urbanism and its assessment in developing countries: The Nigerian case. https://core.ac.uk/reader/74252580.

Mustoe, R., & Walker, D. (1977). *Advanced Engineering Mathematics*. Chichester, UK: John Wiley & Sons.

Nchofoung, T. N., & Asongu, S. A. (2022). Effects of infrastructures on environmental quality contingent on trade openness and governance dynamics in Africa. *Renewable Energy, 189*, 152–163. https://doi.org/10.1016/j.renene.2022.02.114

Nchofoung, T. N., Asongu, S. A., Njamen Kengdo, A. A., & Achuo, E. D. (2022). Linear and non-linear effects of infrastructures on inclusive human development in Africa. *African Development Review, 34*(1), 81–96. https://doi.org/10.1111/1467-8268.12619

Ng, C. P., Law, T. H., Jakarni, F. M., & Kulanthayan, S. (2019). Road infrastructure development and economic growth. *IOP Conference Series: Materials Science and Engineering, 512*(1). https://doi.org/10.1088/1757-899X/512/1/012045

Nguyen, H. M., & Nguyen, L. D. (2016). The relationship between urbanisation and economic growth. *International Journal of Social Economics, 45*(2), 316–339.

Njoh, A. J. (2003). Urbanization and development in sub-Saharan Africa. *Cities, 20*(3), 167–174. https://doi.org/10.1016/S0264-2751(03)00010-6

Owusu-Manu, D. G., Jehuri, A. B., Edwards, D. J., Boateng, F., & Asumadu, G. (2019). The impact of infrastructure development on economic growth in sub-Saharan Africa with special focus on Ghana. *Journal of Financial Management of Property and Construction, 24*(3), 253–273. https://doi.org/10.1108/JFMPC-09-2018-0050

Palei, T. (2015). Assessing the impact of infrastructure on economic growth and global competitiveness. *Procedia Economics and Finance, 23*(October 2014), 168–175. https://doi.org/10.1016/s2212-5671(15)00322-6

Pradhan, R. P., Arvin, M. B., & Nair, M. (2021). Urbanization, transportation infrastructure, ICT, and economic growth: A temporal causal analysis. *Cities, 115* (April). https://doi.org/10.1016/j.cities.2021.103213

Rasiah, R., Salih, K., & Cheok, C. K. (2022). In S. Nature (Ed.), *Malaysia's leap into the future: The building blocks towards balanced development*.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal, 9*(1), 86–136. https://doi.org/10.1177/1536867x0900900106

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20*(C), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Saadaoui Mallek, R., Awad, A., Ozturk, I., & Ben Douissa, I. (2024). Infrastructure development and poverty eradication in sub-Saharan Africa; its effect and transmission channels. *Cities, 144*(2024). https://doi.org/10.1016/j.cities.2023.104658

Spence, M., Annez, P. C., & Buckley, R. M. (2009). Urbanization and growth: Setting the context. *Urbanization and Growth, February*, 1–46.

Srinivasu, B., & Srinivasa Rao, P. (2013). Infrastructure development and economic growth: Prospects and perspective. *Journal of Business Management & Social Sciences Research, 2*(1), 2319–5614.

Svirejeva-Hopkins, A. (2012). Urbanization as a global historical process: Theory and evidence from sub-Saharan Africa. *Population and Development Review, 38*(June), 3672–3678. https://doi.org/10.1016/B978-008045405-4.00764-3

Tayi, S., & Radoine, H. (2023). Mapping built-up area: combining Radar and Optical Imagery using Google Earth Engine. In *2023 Joint Urban Remote Sensing Event, JURSE 2023*. https://doi.org/10.1109/JURSE57346.2023.10144131

Turok, I., & McGranahan, G. (2013). Urbanization and economic growth: The arguments and evidence for Africa and Asia. *Environment and Urbanization, 25*(2), 465–482. https://doi.org/10.1177/0956247813490908

UN DESA. (2019). World urbanization prospects: The 2018 revision. In *, Vol. 12. Demographic research*. United Nations. https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.

UN-Habitat. (2022). Envisaging the future of cities. In *World city report 2022*.

Union Africaine. (2014). *Programme for infrastructure development in Africa (PIDA): Addressing the infrastructure gap in Africa, to speed up regional integration* (Issue July).

United Nations. (2019). World urbanization prospects: The 2018 revision. In *, Vol. 12. Demographic research*. https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf.

Wang, D., Zhou, T., & Wang, M. (2021a). Information and communication technology (ICT), digital divide and urbanization: Evidence from Chinese cities. *Technology in Society, 64*(January). https://doi.org/10.1016/j.techsoc.2020.101516

Wang, D., Zhou, T., & Wang, M. (2021b). Information and communication technology (ICT), digital divide and urbanization: Evidence from Chinese cities. *Technology in Society, 64*(December 2020). https://doi.org/10.1016/j.techsoc.2020.101516

Wang, Z., Ahmed, Z., Zhang, B., & Wang, B. (2019). The nexus between urbanization, road infrastructure, and transport energy demand: Empirical evidence from Pakistan. *Environmental Science and Pollution Research, 26*(34), 34884–34895. https://doi.org/10.1007/s11356-019-06542-8

Weinhold, D., & Reis, E. J. (2001). Model evaluation and causality testing in short panels: The case of infrastructure provision and population growth in the Brazilian Amazon. *Journal of Regional Science, 41*(4), 639–657. https://doi.org/10.1111/0022-4146.00236

World Bank. (1994). Infrastructure for development. In *World development*. https://doi.org/10.2307/j.ctt18fsc3v

Xu, R., & Wunsch, D. (2005). Survey of clustering algorithms. *IEEE Transactions on Neural Networks, 16*(3), 645–678. https://doi.org/10.1109/TNN.2005.845141

Rachida El-Bouayady is a PhD student at the School of Architecture Planning and Design (SAP+D) at University Mohammed VI Polytechnic (Um6p). She graduated from Rabat School of Mines with an energy engineering major. Her research is focused on developing frameworks that facilitate a deeper understanding of future infrastructure development in Africa, utilizing concepts such as integration, holism, and data-driven methods.

Hassan Radoine is Full Professor and Director of the School of Architecture, Planning and Design (SAP+D) at Mohammed VI Polytechnic University. He received his Architect Diploma from Ecole Nationale d'Architecture in Morocco. He completed his M.Sc. and Ph. D. in Architecture at the University of Pennsylvania. He has taught and conducted research at UPenn, Dartmouth College, EPFL, Ecole Nationale d'Architecture, University of Sharjah, American University of Sharjah, Al Akhawayn University and UM6P. He worked as an international expert for UN-Habitat, UNESCO, ICCROM, MCC, World Bank, and the Aga Khan Award. He contributed to several journals and books.

Nour-Eddin El Faouzi earned his Ph.D. in mathematics and statistics from the University of Montpellier, France, in 1992. He is a Professor of traffic modeling and engineering at the University of Lyon and has been the Director of Research at IFSTTAR in Bron, France, since 1992. Additionally, he directs the Transport and Traffic Engineering Laboratory (LICIT), a joint research unit of IFSTTAR and ENTPE, and serves as an Adjunct Professor at Queensland University of Technology, Australia. His research focuses on traffic modeling, flow prediction, and intelligent transport system applications. Prof. El Faouzi is also a member of several scientific and technical committees.

Soukaina Tayi is in her third year of the Ph.D. program in spatial data and urban planning at the School of Architecture Planning & Design (SAP+D), Mohammed VI Polytechnic University (UM6P) in Morocco. Her research interests are in urban planning and geospatial technologies. She holds a GIS engineer Diploma of geographic information sciences from the Hassania school of public works, in Casablanca, Morocco.

Hakkı Can Ozkan He completed his undergraduate studies in Architecture at Yıldız Technical University in 2008. He went on to earn his master's degree from the same university, specializing in Architecture, with a thesis titled “A Sustainable City Model: Slow City Movement”. In 2018, he obtained his Ph.D. with a thesis entitled “A United Nations-Focused Development Model Proposal for Urban Sustainability” from the Building Research and Planning program at Yıldız Technical University. Between 2008 and 2018, he worked in the private sector, achieving recognition through various national and international competitions. Since 2018, he has been serving as a lecturer at Sultan Mehmet Foundation University.