

Original Article

Iranian Remote Sensing GIS
Journal of Sensing

Evaluation and Analysis of the Evolution and Completeness of Building Data in Volunteered Geographic Information (Case Study: OpenStreetMap Building Information in Mashhad City)

Hadiyeh, Hassannia Badrabad¹, Babak Mirbagheri¹, Masoud Minaei^{2*} **Affiliation**

1. Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
2. Dep. of Geography, Geographic Information Science/System and Remote Sensing Laboratory (GISSRS: Lab), Ferdowsi University of Mashhad, Mashhad, Iran

ABSTRACT

Introduction and Objectives: The advent of digital sensors, smartphones, crowdsourced applications, and vibrant social media platforms has dramatically transformed the landscape of geographic data accessibility. This transformation is even more pronounced with the emergence of user-generated content, a hallmark of Web 2.0 technologies. This participatory phenomenon is termed Volunteered Geographic Information (VGI), of which OpenStreetMap (OSM) serves as a prominent illustration. OSM offers an extensive repository of detailed and frequently updated geospatial data concerning physical infrastructures, including buildings, roads, and other vital landmarks. Though OSM provides this wealth of data free of charge and continues to expand its database, the quality of the information necessitates meticulous evaluation to determine its suitability for various practical applications. Research across the globe has underscored the efficacy of VGI; however, there remains a significant gap in studies investigating the quality of spatial information pertaining specifically to buildings within OSM, particularly in regions like Iran. This study aims to address this gap by focusing on three critical aspects: first, it will evaluate the quality of building data captured in OSM; second, it will analyze the enhancements and overall completeness of this data from the years 2018 to 2023; and third, it will explore the intricate relationship between land use patterns and population density, utilizing the building data derived from OSM. Through this comprehensive examination, the study hopes to shed light on the integrity of VGI in representing urban landscapes.

Materials and Methods: The study area, encompassing districts 9 and 11 of Mashhad, was divided into a grid of 1 x 1 km cells, serving as the foundational framework for our analysis. Within each cell, we assessed three key metrics from 2018 to 2023: building density, completeness, and the quantity of buildings documented in OpenStreetMap (OSM) data. Notably, the ratio of officially registered charges to the overall count of charges within the voluntary dataset indicates the extent to which official data is represented in the voluntary dataset. This metric reflects the completeness of the voluntary data relative to official sources. Subsequently, we analyzed the correlation between building density and data completeness for each grid cell. Additionally, demographic parameters, including population and land use, were incorporated into the evaluation process.

Results and Discussion: The construction data sourced from OpenStreetMap (OSM) across the 9th and 11th districts of Mashhad has witnessed an impressive twentyfold increase from 2018 to 2023. This extraordinary expansion is particularly evident in the eastern and northeastern sectors of the city, where development has surged. Among the various land use categories identified, residential areas, followed by residential-commercial zones, commercial spaces, and educational institutions, display the highest frequency of construction activities in that specific order. Moreover, there exists a notable and robust correlation between the OSM building data and the population residing within each building block. However, this connection has exhibited significant fluctuations throughout the analyzed period, indicating varying patterns of growth and development. Furthermore, an intriguing relationship emerges between the density index and the completeness of the data; in regions where the density of OSM building data is greater, users tend to exhibit a heightened willingness to create and delineate new buildings, reflecting an active engagement in urban development.

Conclusion: An in-depth examination of completeness index values reveals a noteworthy discrepancy in the OpenStreetMap (OSM) dataset, which significantly underrepresents the actual number of buildings when compared to official records. This research compellingly illustrates that, although OSM data can serve as a valuable resource for geographical analyses in urban environments—especially within bustling metropolises such as Mashhad—there exists a pressing necessity for enhancements and validation of the data to attain a higher level of accuracy. Moreover, a meticulous evaluation of key performance indicators can provide profound insights that not only enhance the reliability of the OSM dataset but also bolster its applications in urban planning and other related fields, thereby paving the way for more informed decision-making processes.

Keywords: VGI, OpenStreetMap, Building, Completeness, Mashhad.

Citation:

Hassannia Badrabad, H., Mirbagheri, B., Minaei, M., Evaluation and Analysis of the Evolution and Completeness of Building Data in Volunteered Geographic Information (Case Study: OpenStreetMap Building Information in Mashhad City), *Iran J Remote Sens GIS*. 18(1): 23-36.

* Corresponding Author: m.minaei@um.ac.ir
DOI: <https://doi.org/10.48308/gisj.2025.234725.1203>

Received: 2024.02.05
Accepted: 2025.02.20





ارزیابی و تحلیلی بر ارتقا و کامل بودن داده‌های ساختمانی در اطلاعات جغرافیایی داوطلبانه (مطالعه موردی: اطلاعات ساختمانی OpenStreetMap در شهر مشهد)

هدیه حسن‌نیا بدرآباد^۱، بابک میرباقری^۱، مسعود مینائی^{۲*}

چکیده

مقدمه و هدف: افزایش فزاینده حسگرهای دیجیتال، تلفن‌های هوشمند، برنامه‌های کاربردی جمع‌سپاری و رسانه‌های اجتماعی موجب افزایش سریع میزان داده‌های جغرافیایی شده است. در این راستا، منبعی نوین و شایان توجه از اطلاعات جغرافیایی، به‌صورت محتوای وب، ازسوی کاربران در دسترس قرار گرفته است. این منبع اغلب با فناوری‌هایی با عنوان Web 2.0 پشتیبانی می‌شود. در چارچوب جغرافیای نوین، افراد داده‌های دریافتی از محیط خود را گردآوری و بارگذاری می‌کنند و از این طریق، به بهبود داده‌هایی که دیگران مطرح کرده‌اند، کمک می‌کنند. این روش نوین با عناوینی همچون اطلاعات مکانی مردم‌گستر یا Volunteered Geographic Information (VGI) شناخته می‌شود. یکی از منابع VGI داده‌های وب‌سایت OpenStreetMap (OSM) است که انواع گوناگونی از داده‌های مکان‌مبنا، ازجمله ساختمان‌ها را شامل می‌شود. استفاده از داده‌های OSM مزیت‌های بسیاری دارد زیرا این داده‌ها به‌رايگان در دسترس‌اند و به‌سرعت به‌روز می‌شوند. پایگاه داده OSM، هم از نظر تعداد کاربران و هم از نظر حجم داده‌های مکانی، به‌سرعت در حال توسعه است. از این رو به مجموعه‌داده‌های جهانی، پرکاربرد و باکیفیت تبدیل شده است. با وجود اینکه دسترسی به چنین داده‌هایی جالب‌توجه به نظر می‌رسد، کیفیت این داده‌ها به ارزیابی نیاز دارد؛ بنابراین تناسب آن‌ها برای اهدافی خاص باید ارزیابی شود. به‌رغم مطالعات انجام‌شده در سطح جهان و مزایای بهره‌برداری از اطلاعات مکانی مردم‌گستر در ایران، تحقیقات بسیار اندکی در زمینه داده‌های فضایی مرتبط با ساختمان‌ها در OSM انجام شده است. از این رو در پژوهش حاضر: (۱) کیفیت داده‌های ساختمانی OSM ارزیابی می‌شود؛ (۲) روند ارتقا و کامل بودن ساختمان‌ها در داده‌های OSM، طی دوره زمانی شش‌ساله‌ای (۲۰۱۸-۲۰۲۳)، بررسی می‌شود و (۳) ارتباط بین کاربری اراضی و جمعیت با داده‌های ساختمانی OSM ارزیابی می‌شود.

مواد و روش‌ها: منطقه مورد مطالعه (مناطق ۹ و ۱۱ مشهد) به شبکه‌ای سلولی با اندازه ۱×۱ کیلومتر تقسیم شد و این مساحت، به‌منزله واحد جغرافیایی، مبنای محاسبات قرار گرفت. برای هر واحد جغرافیایی، سه معیار (تراکم ساختمانی، کامل بودن و تعداد ساختمان در داده‌های OSM) در منطقه مورد مطالعه، متعلق به سال‌های ۲۰۱۸ تا ۲۰۲۳، بررسی شد. شایان ذکر است که تعداد عوارض رسمی متناظر در مجموعه‌داده جغرافیایی داوطلبانه، به تعداد کل عوارض مجموعه‌داده رسمی، میزان حضور داده‌های رسمی را در مجموعه‌داده جغرافیایی داوطلبانه نشان می‌دهد. این مقدار همان میزان کامل بودن داده‌های داوطلبانه در قیاس با داده‌های رسمی است. در گام بعدی، رابطه بین تراکم و کامل بودن داده‌های ساختمانی، درمورد هر سلول شبکه، محاسبه شد و دو پارامتر جمعیت و کاربری اراضی نیز در بررسی‌ها به کار رفت.

نتایج و بحث: داده ساختمانی OSM در مناطق ۹ و ۱۱ مشهد، از ۲۰۱۸ تا ۲۰۲۳، بیست‌برابر افزایش یافته و این افزایش، به‌ویژه درمورد مناطق شرق و شمال‌شرق، مشهودتر است. همچنین کاربری اراضی مسکونی، مسکونی-تجاری، تجاری و آموزشی، به‌ترتیب، بیشترین فراوانی را به خود اختصاص داده‌اند. به‌رغم همبستگی مثبت و معنادار بین داده‌های ساختمانی OSM و جمعیت هر بلوک ساختمانی، این همبستگی در بازه زمانی مورد بررسی به‌صورت نوسانی متغیر بوده است. مقدار همبستگی‌ها میان شاخص تراکم و کامل بودن نشان می‌دهد، در مکان‌هایی که تراکم داده ساختمانی در داده OSM بیشتر بوده است، کاربران تمایل بیشتری به ایجاد و ترسیم ساختمان جدید داشته‌اند.

نتیجه‌گیری: با توجه به مقادیر شاخص کامل بودن، می‌توان نتیجه گرفت که ساختمان‌های به‌ثبت‌رسیده در OSM بسیار کمتر از تعداد واقعی ساختمان‌ها در داده رسمی است. با توجه به این تحقیق، می‌توان گفت که داده‌های OSM برای تحلیل جغرافیایی ساختمان‌ها در شهرهای بزرگ، مانند مشهد، مناسب‌اند اما، برای استفاده‌های بیشتر و دقیق‌تر، به تکمیل و بهبود این داده‌ها نیاز است. همچنین بررسی شاخص‌های تأثیرگذار می‌تواند، به‌منظور استفاده مطمئن‌تر از این داده‌ها، اطلاعات سودبخشی در اختیار کاربران قرار دهد.

واژه‌های کلیدی: OpenStreetMap، VGI، ساختمان، کامل بودن، مشهد.

استناد:

حسن‌نیا بدرآباد، ه. میرباقری، ب. مینائی، م. ارزیابی و تحلیلی بر ارتقا و کامل بودن داده‌های ساختمانی در اطلاعات جغرافیایی داوطلبانه (مطالعه موردی: اطلاعات ساختمانی OpenStreetMap در شهر مشهد)، نشریه سنجش از دور و GIS ایران، سال ۱۸، شماره ۱، بهار ۱۴۰۵: ۲۳-۳۶.



۱- مقدمه

افزایش فزاینده حسگرهای دیجیتالی، تلفن‌های هوشمند، برنامه‌های کاربردی جمع‌سپاری^۱ و رسانه‌های اجتماعی موجب افزایش سریع حجم داده‌های جغرافیایی شده است (Mobasher, 2017). در این راستا، منبعی نوین و درخور توجه شامل اطلاعات جغرافیایی، به صورت محتوای وب، از سوی کاربران در دسترس قرار گرفته است. این منبع اغلب با فناوری‌هایی با عنوان Web 2.0 تجهیز می‌شود (Elwood et al., 2012). در چارچوب جغرافیایی نوین^۲، افراد داده‌های محیط خود را گردآوری و آپلود می‌کنند و از این راه، به بهبود داده‌های موجود ارائه‌شده از سوی دیگران کمک می‌کنند. این روش نوین با عناوینی همچون اطلاعات جغرافیایی داوطلبانه (VGI)^۳ و داده‌های جغرافیایی جمع‌سپاری (CSGD)^۴ شناخته می‌شود (Minaci, 2020). بنابراین داده‌های جغرافیایی داوطلبانه آنلاین و رایگان، با رویکردی مبتنی بر جمع‌سپاری، پدید آمده‌اند. پدیده جمع‌سپاری، به طور کلی و اطلاعات جغرافیایی داوطلبانه، به طور خاص، الگوواره‌ای جدیدند که می‌توانند به غنی‌سازی چارچوب‌های کنونی در علم اطلاعات جغرافیایی^۵ کمک کنند (Bakillah & Liang, 2016; Sun & Du, 2017). یکی از منابع VGI داده‌های وبسایت OpenStreetMap (OSM) است که انواع گوناگونی از داده‌های مکان‌مبنا، از جمله جاده‌ها، ساختمان‌ها، کاربری اراضی و عوارض طبیعی همچون رودخانه‌ها را شامل می‌شود (Minaci, 2020). پایگاه داده OSM، هم از نظر تعداد کاربران و هم از نظر حجم داده‌های جغرافیایی، به سرعت در حال توسعه است و کاربران و سازمان‌های بسیاری در به‌روزرسانی داده‌ها مشارکت دارند؛ از این رو به مجموعه داده‌ای جهانی، پرکاربرد و باکیفیت تبدیل شده است. در ضمن در حال حاضر، علاوه بر معابر، سایر داده‌ها همچون پارسل‌های شهری و بسیاری دیگر از نقاط مهم شهری را نیز دربرمی‌گیرد و بنابراین مرجعی معتبر در پژوهش‌های جغرافیایی و برنامه‌ریزی شهری به شمار می‌رود. به همین دلیل، در این مقاله این بستر برای بررسی انتخاب شده

است (Minaci, 2020; Zhou et al., 2022). با اینکه دسترسی به چنین داده‌هایی جالب توجه به نظر می‌رسد، کیفیت آن‌ها به ارزیابی نیاز دارد؛ بنابراین تناسبشان برای اهداف خاص باید ارزیابی شود (Goodchild, 2007). طبق تعریف سازمان جهانی استاندارد (ISO)، کیفیت داده‌های جغرافیایی شامل شش عنصر اصلی کامل بودن^۶، سازگاری منطقی^۷، دقت موقعیتی، دقت معنایی، کیفیت زمانی و امکان استفاده می‌شود (Kresse & Fadaie, 2004). یکی از داده‌های OSM ساختمان‌هاست که با توجه به اهمیتشان، مطالعات متعددی برای ارزیابی کیفیت آن‌ها براساس معیارهای کیفی و نیز اهداف کاربردی گوناگون انجام شده است. هخت^۸ و همکاران (۲۰۱۳) کیفیت داده‌های ساختمانی OSM را برای ایالت‌هایی در آلمان، در مقایسه با داده رسمی، ارزیابی کردند. برطبق این پژوهش، کامل بودن ساختمان‌های OSM در نواحی شهری بیشتر از مناطق روستایی است. فن^۹ و همکاران (۲۰۱۴) کیفیت داده‌های ساختمانی OSM را در شهر مونیخ کشور آلمان ارزیابی کردند. آن‌ها، با استفاده از داده‌های رسمی، داده‌ها را بابت چهار شاخص کیفیت (کامل بودن، دقت معنایی، دقت موقعیتی و دقت شکلی) بررسی کردند و به این نتیجه رسیدند که دقت معنایی و کامل بودن در داده‌های مونیخ چشمگیر است؛ در صورتی که از نظر دقت موقعیتی، به طور متوسط، چهار متر انحراف وجود دارد. تورنروز^{۱۰} و همکاران (۲۰۱۵)، به دنبال ارزیابی کامل بودن داده‌های ساختمانی در جنوب غرب آلمان، دو روش سنجش کامل بودن براساس تعداد ساختمان‌ها و نیز براساس مساحت را مقایسه کردند تا داده‌های ساختمانی OSM و داده‌های

1. Crowdsourcing
2. Neogeography
3. Volunteered Geographic Information (VGI)
4. Crowdsourcing Geographic Data (CSGD)
5. Geographic Information Science (GIScience)
6. Completeness
7. Logical Consistency
8. Hecht
9. Fan
10. Törnros

ارزیابی روند تکمیل بلوک‌های ساختمانی کلان‌شهر تهران در داده‌های OSM طی سال‌های ۲۰۱۱ تا ۲۰۲۰، پژوهشی انجام داده‌اند. بر مبنای نتایج این تحقیق، در دو سال آخر بازه زمانی منتخب، داده‌های بلوک ساختمانی OSM از نظر تعداد عوارض و کامل بودن اطلاعات هندسی افزایش چشمگیری داشته و کامل بودن این داده‌ها، از ۰/۱۸ در سال ۲۰۱۱، به ۲/۷ در سال ۲۰۲۰ رسیده است. سدید^۶ و همکاران (۲۰۲۴) در مقاله‌ای، با هدف ارزیابی استفاده از هوش مصنوعی در تکمیل داده‌های داوطلبانه کاربری اراضی OSM در شهر کرج، از روش شبکه عصبی کانولوشنی و سپس روش واحد مبنا استفاده کردند تا میزان کامل بودن داده‌های OSM را ارزیابی کنند. نتایج این تحقیق نشان می‌دهد میزان کامل بودن بلوک‌های ساختمانی OSM در کل منطقه مطالعاتی برابر با ۳/۶٪ است. همچنین این پژوهش نرخ پایین کامل بودن بلوک‌های ساختمانی و اراضی جنگلی و نرخ بالای کاربری کشاورزی و زمین‌های میوه را مشخص کرد. به‌رغم مطالعات انجام شده در سطح جهان و مزایای بهره‌برداری از اطلاعات جغرافیایی داوطلبانه در ایران، تحقیقات بسیار اندکی در زمینه داده‌های ساختمان‌ها در OSM انجام شده است. از این رو، در پژوهش حاضر، روند ارتقا و کامل بودن ساختمان‌ها در داده‌های OSM طی دوره زمانی شش‌ساله‌ای (۲۰۱۸-۲۰۲۳) بررسی می‌شود. همچنین ارتباط بین کاربری اراضی و جمعیت با داده‌های ساختمانی OSM ارزیابی می‌شود. در ادامه مقاله، بخش‌های منطقه مورد مطالعه و داده‌های مورد استفاده بیان خواهد شد. پس از آن، بخش روش‌شناسی و در نهایت، بخش‌های یافته‌ها و نیز بحث و نتیجه‌گیری در مقاله مطرح خواهد شد.

رسمی را با هم مقایسه کنند. نتایج این تحقیق نشان داد، در روش سنجش براساس تعداد، معمولاً نرخ کامل بودن ساختمان‌های OSM کمتر از مقدار واقعی برآورد می‌شود. در روش سنجش براساس مساحت نیز معمولاً نرخ کامل بودن داده‌های OSM بیشتر از مقدار واقعی به دست می‌آید. بنابراین، در این مقاله استفاده از روش‌های پیش‌پردازش داده‌ها، برای دستیابی به نرخ دقیق‌تر، توصیه شده است. ژو^۱ (۲۰۱۸)، در مقاله‌ای، به رابطه بین تراکم و کامل بودن داده‌های ساختمان OSM در نواحی شهری برای تخمین کیفیت این داده‌ها پرداخت. نتایج این مقاله رابطه‌ای خطی را بین کامل بودن و تراکم ساختمانی^۲ داده‌های ساختمانی OSM نشان داد؛ به‌علاوه اینکه مشخص کرد تراکم ساختمانی معیاری تأثیرگذار برای تخمین کامل بودن داده‌های ساختمانی OSM است. تیان^۳ و ژو (۲۰۱۹)، در پژوهشی، ارتقا و کامل بودن و الگوی جغرافیایی ساختمان‌های OSM را در چین تحلیل کردند. طبق نتایج این تحقیق، تعداد ساختمان‌های OSM از سال ۲۰۱۲ تا ۲۰۱۷ تقریباً بیست‌برابر افزایش داشته است. طول جاده و شاخص اقتصادی دو پارامتر تأثیرگذار در افزایش ساختمان‌های OSM در این کشور شمرده می‌شود. ژانگ^۴ و همکاران (۲۰۲۲) کامل بودن ساختمان‌ها در OSM را با استفاده از اطلاعات جمعیتی، ارزیابی کردند. در این مطالعه، با به‌کاربردن داده‌های جمعیتی، کامل بودن ساختمان‌های OSM ارزیابی شد و نتایج نشان داد که استفاده از این داده‌ها، با هدف ایجاد داده‌های مورد اعتماد و معتبر ساختمانی، روشی مؤثر برای ارزیابی کامل بودن ساختمان‌ها در OSM است. ژو و همکاران (۲۰۲۲) وضعیت کامل بودن ساختمان‌ها را در مورد ۱۲۰۹۷۵ شهر، در سطح جهانی، ارزیابی کردند و در پژوهش خود، از داده‌های جمعیت بهره بردند. نتایج این تحقیق نشان داد، در مورد ۷۵٪ از این شهرها، میزان کامل بودن کمتر از ۲۰٪ است و شاخص کامل بودن، فقط در مورد ۹٪ از شهرهای مورد مطالعه، بیشتر از ۸۰٪ تخمین زده شد. ادبی^۵ و همکاران (۲۰۲۲)، با هدف

1. Zhou

2. OSM Building Density

3. Tian

4. Zhang

5. Adabi

6. Sadidi

۲- مواد و روش‌ها

۲-۱- منطقه مورد مطالعه

محدوده مورد مطالعه، در این پژوهش، مناطق ۹ و ۱۱ شهر مشهد در استان خراسان رضوی است. این مناطق دارای داده‌های OSM ثبت‌شده بیشتری‌اند و بنابراین، در فرایند تحقیق، امکان ارزیابی و مقایسه دقیق‌تری را فراهم می‌آورند. منطقه ۹ با ۴۱/۳۵۷ و منطقه ۱۱ با ۱۵/۸۰۸ کیلومترمربع مساحت، به ترتیب، ۱۳/۲ و ۵/۰۵٪ از مساحت کل شهر را تشکیل می‌دهند (شکل ۱). طبق گزارش رسمی، مجموع جمعیت این دو منطقه ۵۲۷.۲۰۳ نفر است. کاربری اراضی غالب در این مناطق نیز، به ترتیب، مسکونی، مسکونی-تجاری، تجاری و آموزشی است (به‌استثنای مناطق ثبت‌نشده). در این مطالعه، داده‌های مورد نیاز به‌صورت شیب فایل‌های جامع از وبسایت OSM، به‌نشانی <https://download.geofabrik.de>، متعلق به سال‌های ۲۰۱۸ تا ۲۰۲۳ بارگیری شد. پس از بارگیری داده‌ها، ابتدا لایه‌های اطلاعاتی مازاد حذف و فقط اطلاعات درباره ساختمان‌ها استخراج شد. پس از آن پیش‌پردازش‌های لازم، از جمله استخراج ساختمان‌های

منطقه مورد مطالعه، یکسان‌سازی فرمت داده‌های رسمی و داده‌های OSM، یکسان‌سازی سیستم مختصات و حذف خطاهای توپولوژیکی مطالعه، انجام شد.

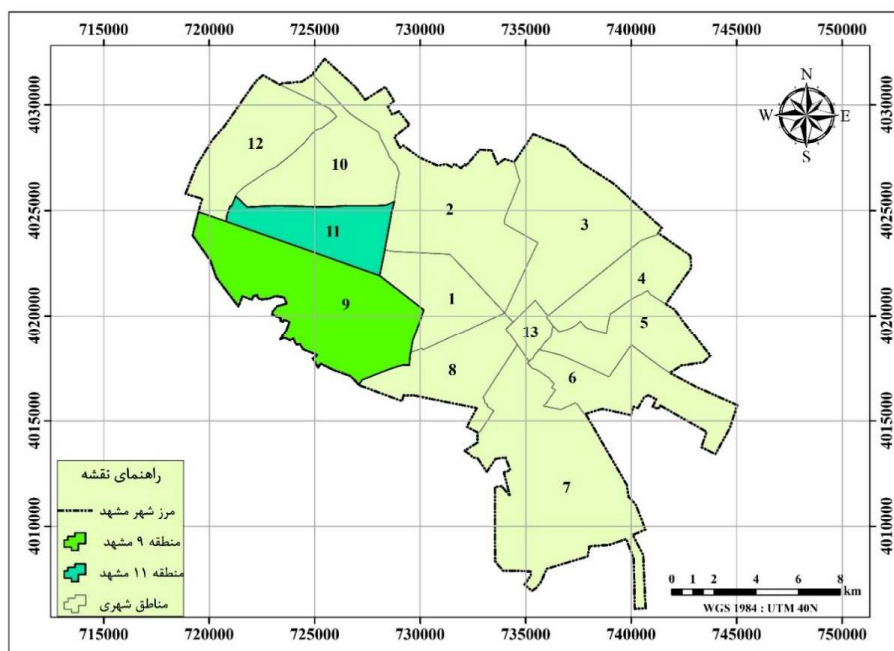
۲-۲- روش‌شناسی

۲-۲-۱- بررسی تعداد ساختمان‌ها در داده OSM

با توجه به افزایش تعداد ساختمان در طول زمان، استفاده از معیار تعداد، برای ارزیابی کامل بودن ساختمان‌ها در OSM، پیشنهاد شده است. به‌لحاظ نظری، اگرچه تعداد نمی‌تواند میزان کامل بودن را به‌طور مشخص نشان دهد، تعداد ساختمان‌های OSM ارتباط مستقیمی با کامل بودن داده ساختمانی OSM در یک منطقه دارد (Tian & Zhou, 2019).

۲-۲-۲- بررسی تراکم ساختمان‌ها در داده OSM

تراکم ساختمان (نسبت پوشش ساختمان) به درصد کل مساحت ساختمانی در هر واحد جغرافیایی اشاره دارد و اغلب به‌منزله معیاری در زمینه طراحی و برنامه‌ریزی شهری استفاده می‌شود (Tian & Zhou, 2019).



شکل ۱. موقعیت منطقه ۹ و ۱۱ در شهر مشهد

کامل بودن داده‌های داوطلبانه در قیاس با داده‌های رسمی است. شاخص کامل بودن ساختمان‌ها در داده OSM با استفاده از رابطه (۲) بررسی می‌شود.

$$\text{رابطه (۲)} \quad \text{کامل بودن (ساختمان)} = \frac{\text{مجموع مساحت ساختمان در داده OSM}}{\text{مجموع مساحت ساختمان در داده رسمی}} \times 100$$

کامل بودن ساختمان‌ها در OSM ممکن است بین ۰٪ (یعنی، در این واحد جغرافیایی، هیچ ساختمانی وجود ندارد) تا ۱۰۰٪ (به این مفهوم که کل مساحت لایه OSM و ساختمان‌های لایه رسمی، در یک واحد جغرافیایی، برابر است) متفاوت باشد. به طور مثال، کامل بودن ساختمان‌ها در OSM در شکل ۲-b به صورت $(0.2 + 0.1) \div 0.1 = 33\%$ و نیز در شکل ۲-c به صورت $(0.2 \div (0.1 + 0.2)) = 66.7\%$ محاسبه می‌شود. هرچند، در صورتی که در واحدی جغرافیایی، چه در لایه رسمی و چه در لایه OSM، هیچ ساختمانی وجود نداشته باشد، کامل بودن برابر ۱۰۰٪ است.

۲-۲-۴- رابطه بین شاخص تراکم و کامل بودن ساختمان‌ها در داده OSM

به منظور بررسی رابطه بین کامل بودن و تراکم ساختمان‌های OSM، یک لایه شبکه 1×1 (کیلومتر) برای هم‌پوشانی به کار می‌رود. پژوهش ژو (۲۰۱۸) نشان

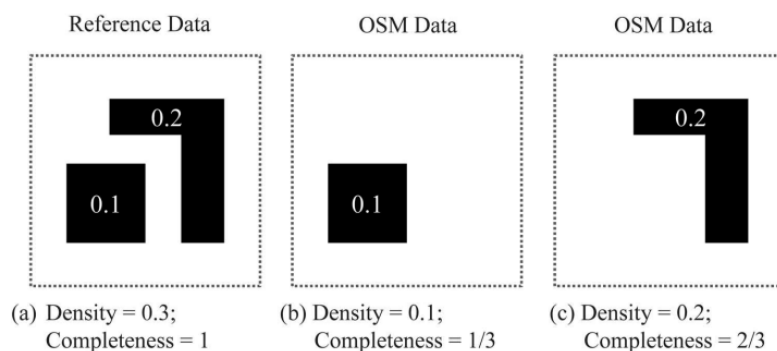
این شاخص مطابق رابطه (۱) محاسبه می‌شود.

$$\text{رابطه (۱)} \quad \text{تراکم (ساختمان)} = \frac{\text{مجموع مساحت ساختمان}}{\text{مساحت واحد جغرافیایی}} \times 100$$

تراکم ساختمان ممکن است بین ۰٪ (به این معنی که هیچ ساختمانی در آن واحد جغرافیایی وجود ندارد) تا ۱۰۰٪ (یعنی این واحد جغرافیایی کاملاً پوشیده از ساختمان است) متفاوت باشد. برای مثال، شکل ۲-a دو ساختمان را در یک واحد جغرافیایی نشان می‌دهد. با فرض اینکه مساحت واحد جغرافیایی ۱ باشد، تراکم ساختمان به این صورت محاسبه می‌شود:

$$((0.1 + 0.2) \div 1) \times 100\% = 30\%$$

۲-۲-۳- بررسی کامل بودن ساختمان‌ها در داده OSM یکی از روش‌های پرکاربرد ارزیابی معیار کامل بودن، مقایسه ساختمان‌ها در داده‌های OSM با داده‌های تولیدشده در منابع رسمی است (Adabi et al., 2022). به طور کلی، عوارض در دو مجموعه داده (رسمی و داوطلبانه) به دو دسته متناظر و نامتناظر تقسیم می‌شوند و از این داده‌ها برای تعیین پارامتر کامل بودن استفاده می‌شود. در واقع، تعداد عوارض رسمی متناظر در مجموعه داده داوطلبانه به تعداد کل عوارض مجموعه داده رسمی، میزان حضور داده‌های رسمی را در مجموعه داده داوطلبانه نشان می‌دهد که این مقدار همان میزان



شکل ۲. محاسبه کامل بودن و تراکم ساختمان

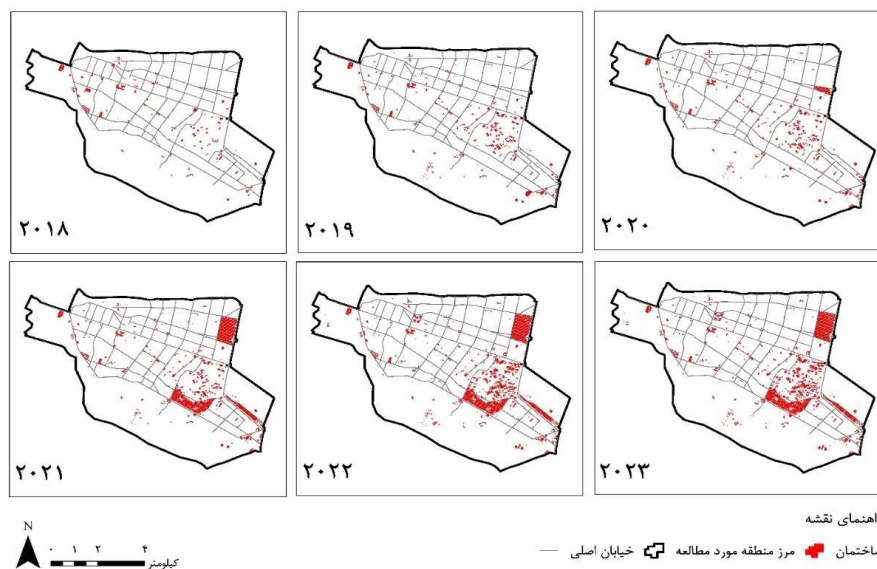
1. Grid

می‌دهد که این اندازه سلول در محدوده شهری را می‌توان آستانه مناسبی برای اندازه واحد جغرافیایی در نظر گرفت. از این رو ساختمان‌های داده OSM در سطح شبکه‌ای به این اندازه در نظر گرفته شده است. سپس هر دو معیار بررسی (کامل بودن و تراکم ساختمانی) برای هر سلول شبکه، مطابق با روابط (۱) و (۲) محاسبه و رابطه بین این دو شاخص تحلیل می‌شود.

۳- نتایج و بحث

۳-۱- تحلیل بر پایه تعداد ساختمان‌های داده OSM

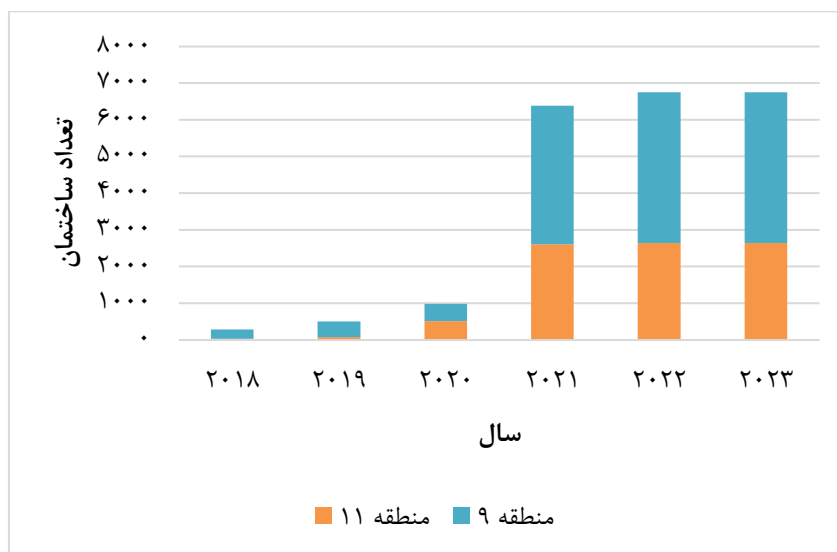
شکل ۳ روند افزایش تعداد ساختمان‌های OSM را، از سال ۲۰۱۸ تا ۲۰۲۳، نشان می‌دهد. در سال ۲۰۱۸، تعداد ۲۸۷ ساختمان وجود دارد. با اندکی افزایش در سال‌های



شکل ۳. ساختمان‌های OSM از سال ۲۰۱۸ تا ۲۰۲۳ در منطقه مورد مطالعه

جدول ۱. تعداد و مساحت ساختمان‌های OSM در منطقه‌های (۹ و ۱۱) شهر مشهد، از سال ۲۰۱۸ تا ۲۰۲۳

سال	۲۰۱۸	۲۰۱۹	۲۰۲۰	۲۰۲۱	۲۰۲۲	۲۰۲۳
تعداد	۲۸۷	۵۰۲	۹۷۹	۶۳۸۱	۶۷۵۲	۶۷۵۲
مساحت (m ²)	۳۷۷۴۳۳.۴۲	۵۷۳۸۲۸.۶۷	۶۰۹۷۳۲.۶۹	۱۳۲۷۱۸۹.۴۵	۱۵۵۸۲۱۷.۸۹	۱۵۶۲۷۰۴.۱۸



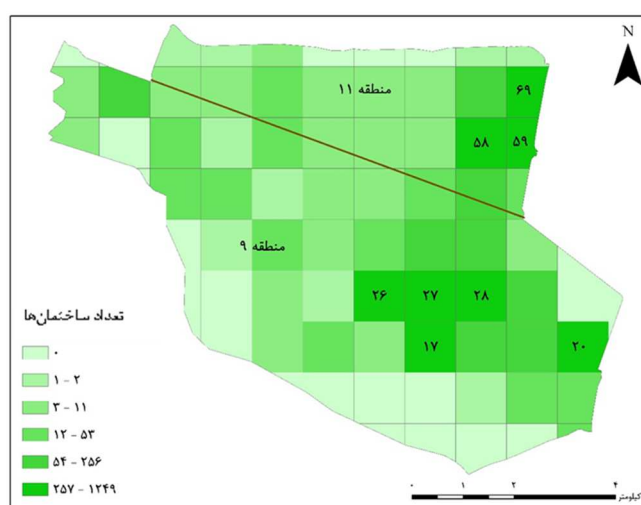
شکل ۴. تعداد ساختمان‌های OSM در منطقه ۹ و ۱۱) شهر مشهد، از سال ۲۰۱۸ تا ۲۰۲۳

۳-۲- پراکندگی، جمعیت و کاربری اراضی ساختمان‌ها

در داده OSM

با توجه به شکل که در آن شماره سلول‌های منتخب مشخص است، در بیشتر سلول‌ها، کمتر از یکصد ساختمان در داده OSM وجود دارد. در صورتی که، هشت سلول از ۵۱ سلولی که داده در آن‌ها ثبت شده است، بیش از ۱۰۰۰ داده ساختمانی را شامل می‌شوند. این سلول‌ها اغلب در شرق و شمال‌شرق منطقه مورد مطالعه واقع شده‌اند. تعداد ساختمان‌های به‌ثبت‌رسیده و

جمعیت این هشت سلول منتخب در جدول ۲ نشان داده شده و کاربری اراضی آن‌ها در جدول ۳ مشخص شده است. به‌طور کلی، توزیع کاربری‌های غالب در این هشت سلول به‌ترتیب، مسکونی، مسکونی-تجاری و تجاری و آموزشی است. به‌بیان‌دیگر، بیشترین ترسیم معمولاً در پیرامون و داخل دانشگاه فردوسی مشهد و اطراف پارک ملت و چهارراه میلاد است که از نظر آموزشی و تجاری، دارای اهمیتی چشمگیرند.



شکل ۵. توزیع تعداد داده‌های ساختمانی OSM در منطقه‌های ۹ و ۱۱) مشهد طی سال ۲۰۲۳، در هر سلول شبکه

جدول ۲. جمعیت و تعداد ساختمان ثبت‌شده داده OSM طی سال ۲۰۲۳، در سلول‌های منتخب

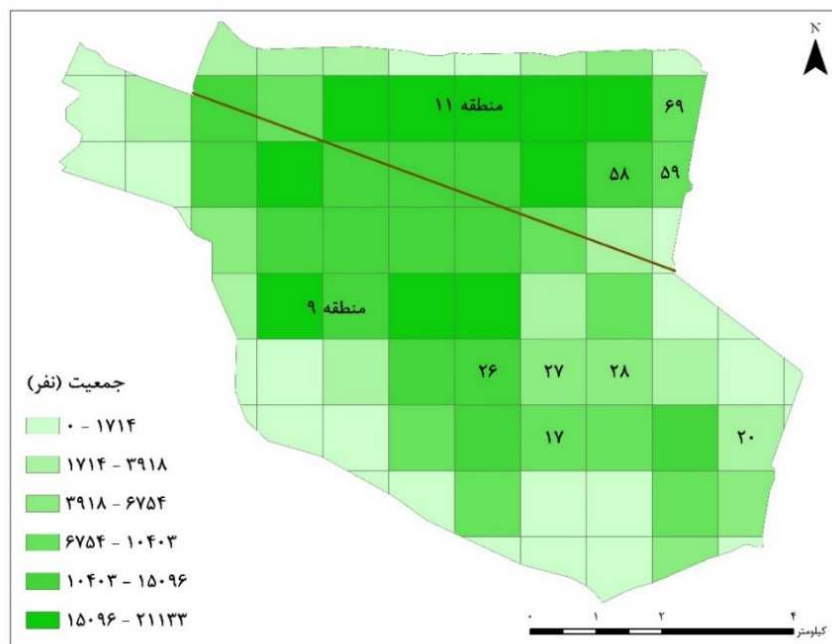
شماره سلول	۶۹	۲۷	۵۹	۲۶	۲۸	۵۸	۱۷	۲۰
تعداد ساختمان	۱۵۵۳	۱۲۴۹	۱۱۱۴	۶۲۴	۵۳۶	۵۲۰	۳۱۱	۲۹۲
رتبه	۱	۲	۳	۴	۵	۶	۷	۸
جمعیت (نفر)	۱۰۲۸۵	۵۶۰۸	۸۱۴۹	۱۳۸۸۲	۵۲۱۰	۱۴۲۵۱	۹۹۸۵	۳۸۴۸

جدول ۳. کاربری اراضی ثبت‌شده در سلول‌های منتخب

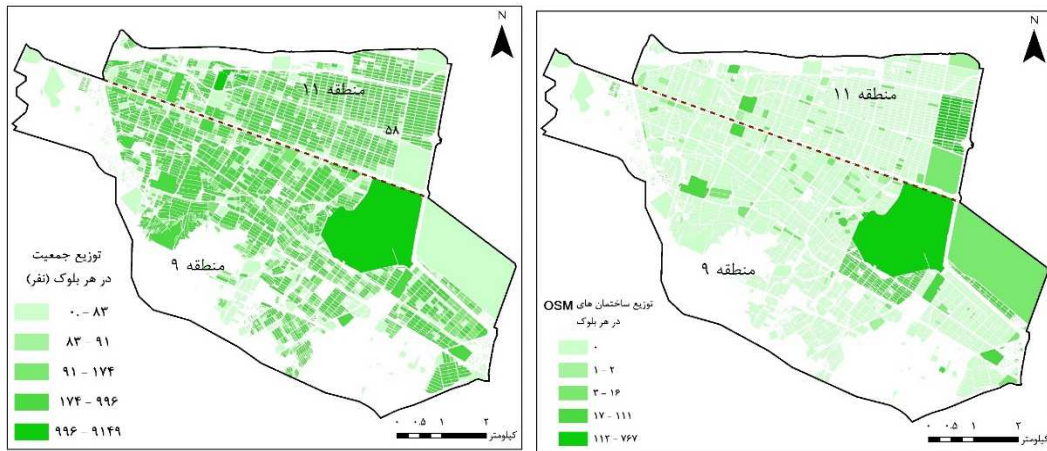
شماره سلول	۱۷	۲۰	۲۶	۲۷	۲۸	۵۸	۵۹	۶۹
آموزشی	۲	۱	۵	۱	۰	۰	۰	۰
تجاری	۱۴	۱۳	۹	۲۱	۱۳	۲۸	۴	۶
مسکونی	۱۰۸۶	۵۳۹	۱۴۳۳	۸۲۵	۵۶۸	۱۴۷۷	۸۱۱	۸۹۸
مسکونی-تجاری	۱۳۵	۵۱	۲۴۷	۲۲۸	۸۸	۴۰۸	۹۱	۷۹

سلول منتخب سکونت دارند. این نکته می‌تواند نشان‌دهنده تأثیر جمعیت در ترسیم ساختمان در داده‌های OSM باشد.

در بررسی رابطه بین داده‌های ساختمانی OSM و جمعیت، با توجه به شکل‌های ۶ و ۷، جمعیت درخور توجهی از منطقه مورد مطالعه (بیش از ۴۰٪) در هشت



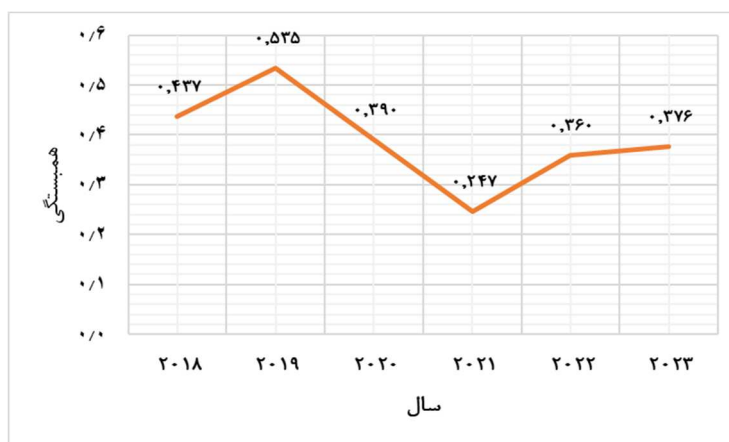
شکل ۶. توزیع جمعیت در منطقه‌های (۹ و ۱۱) مشهد در سال ۲۰۲۳، در هر سلول شبکه



شکل ۷. توزیع تعداد داده‌های ساختمانی OSM در سال ۲۰۲۳، در منطقه‌های (۹ و ۱۱) مشهد در هر بلوک (سمت راست)؛ توزیع جمعیت در منطقه‌های (۹ و ۱۱) شهر مشهد در هر بلوک (سمت چپ)

جمعیت در هر بلوک شهری است. باین‌حال ضریب همبستگی سال ۲۰۲۳ قدری بیشتر از دو سال پیش است. می‌توان نتیجه گرفت که بین تعداد ساختمان‌های ذخیره‌شده در OSM و جمعیت در هر بلوک شهری، طی سال‌های ۲۰۱۸-۲۰۱۹، همبستگی مثبتی وجود دارد. باین‌حال این همبستگی در سال‌های بعدی (۲۰۲۰-۲۰۲۳) ضعیف شده و رابطه میان آن‌ها کمتر مشخص است. بنابراین میان جمعیت هر بلوک و افزایش میزان ترسیم ساختمان‌های OSM رابطه معناداری وجود دارد اما این همبستگی، در سال‌های گوناگون، دارای تغییرات نوسانی است.

با توجه به شکل ۸، در سال ۲۰۱۸ میان تعداد ساختمان‌های ذخیره‌شده در OSM و جمعیت هر بلوک، همبستگی مثبت وجود دارد. بنابراین به‌طور کلی، با افزایش تراکم جمعیتی در هر بلوک شهری، تعداد ترسیم‌ها افزایش یافته است. در سال ۲۰۱۹، ضریب همبستگی نشان‌دهنده رابطه مثبت قوی‌تری در مقایسه با سال قبل است؛ به این معنی که تعداد ساختمان‌های ذخیره‌شده در OSM بیشتر با جمعیت هر بلوک مرتبط است. ضریب همبستگی در مورد سال‌های ۲۰۲۰ تا ۲۰۲۳ کاهش می‌یابد اما همچنان گویای همبستگی مثبتی بین تعداد ساختمان‌های ذخیره‌شده در OSM و

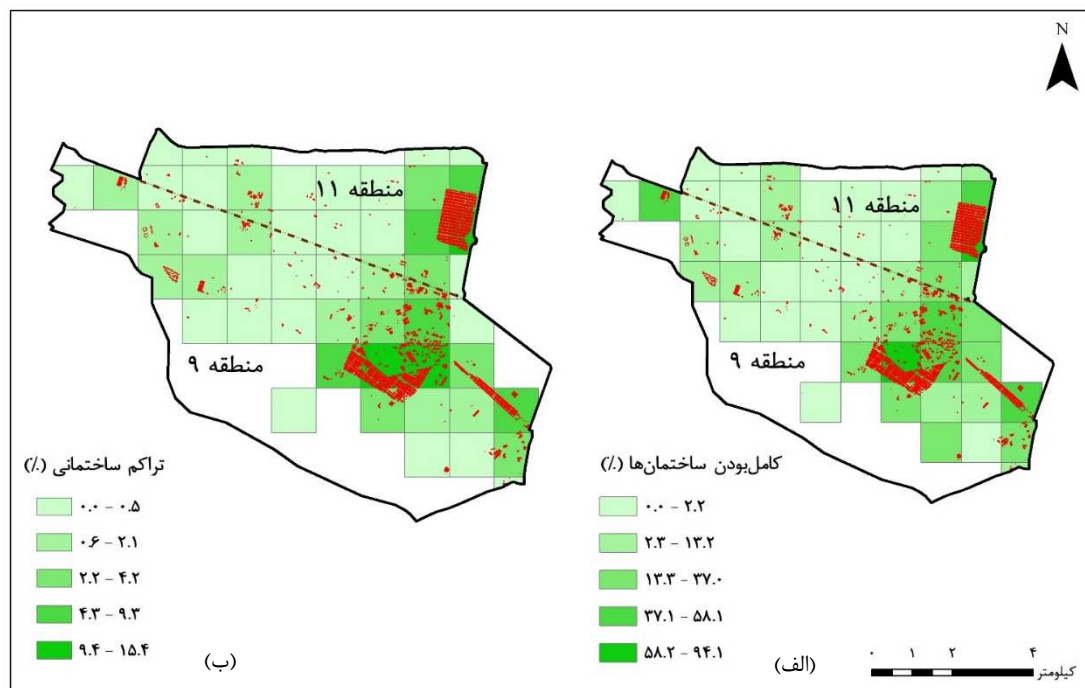


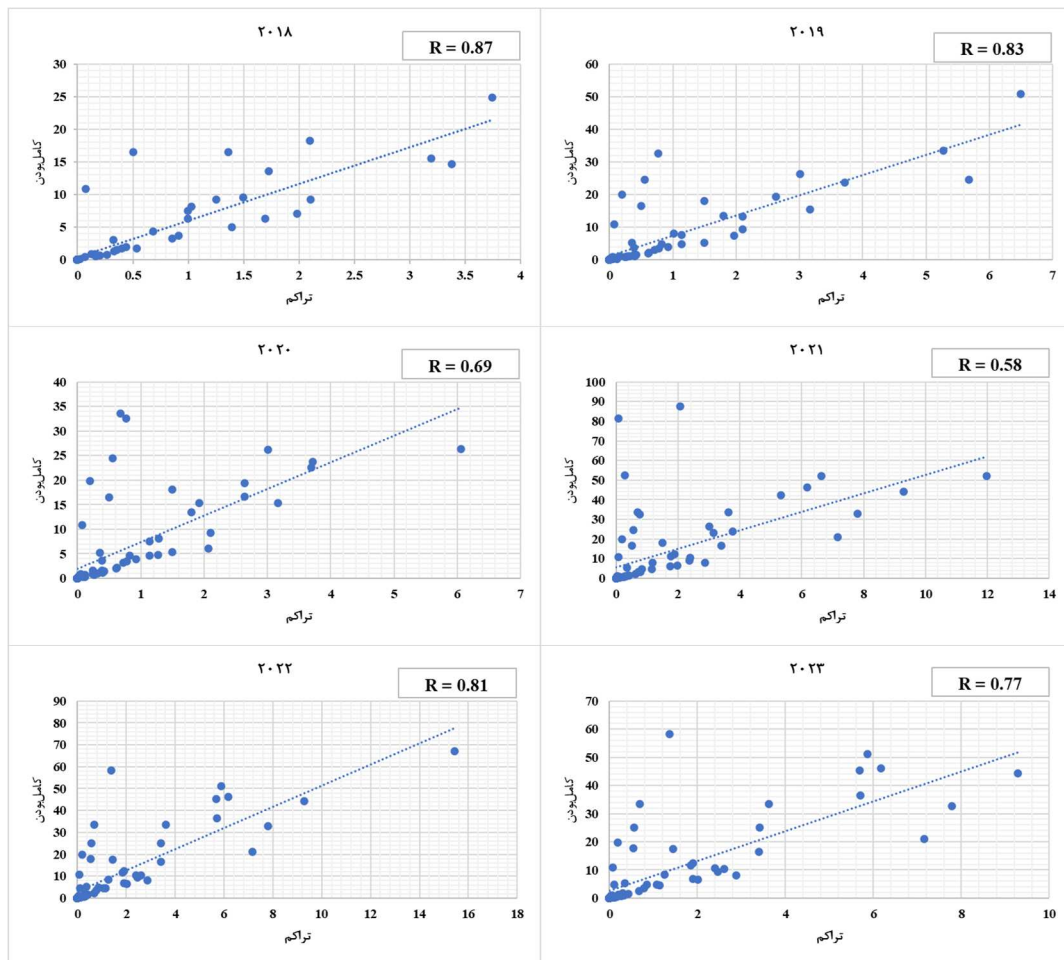
شکل ۸. همبستگی میان تعداد ساختمان‌ها در داده OSM و جمعیت هر بلوک از سال ۲۰۱۸ تا ۲۰۲۳

۳-۳- تحلیل برپایه شاخص تراکم و کامل بودن

با توجه به شکل ۱۱ که تعداد ساختمان‌های OSM را در مقایسه با داده‌های رسمی موجود نشان می‌دهد و نیز شکل ۹-الف که براساس شاخص، کامل بودن منطقه مورد مطالعه را نمایش می‌دهد، به‌خوبی مشخص است که ساختمان‌های داده OSM، در قیاس با داده رسمی، بسیار ناقص‌اند. تحقیقات هخت و همکاران (۲۰۱۳) و هرفورت و همکاران (۲۰۲۳) نیز همین نتیجه را به دست داده‌اند. مطابق شکل ۹، در بخش‌های شرق و شمال‌شرق مناطق ۹ و ۱۱ مشهد، شاخص تراکم و شاخص کامل بودن داده‌های ساختمانی در OSM شایان توجه است. برای بررسی رابطه بین تراکم ساختمانی و کامل بودن داده‌های ساختمانی OSM طی سال ۲۰۱۸ تا ۲۰۲۳، ضریب همبستگی بین این دو شاخص در هر سلول شبکه محاسبه شد (شکل ۱۰). مقدار همبستگی‌ها که حداکثر ۰/۸۷ در سال ۲۰۱۸ و حداقل ۰/۵۸ در سال ۲۰۲۱ است، نشان می‌دهد در مکان‌هایی که تراکم داده ساختمانی در داده OSM بیشتر بوده است، کاربران

تمایل بیشتری برای ایجاد و ترسیم ساختمان جدید داشته‌اند. این پدیده نشان می‌دهد زمانی که در مکانی مشخص از نقشه، عملیات ترسیم و ایجاد ساختمان آغاز می‌شود، این فرایند در سال‌های بعدی نیز ادامه می‌یابد. این مسئله بیانگر آن است که کاربران OSM انگیزه بیشتری برای مشارکت در ایجاد و به‌روزرسانی مناطق شناخته‌شده دارند و به‌مرور زمان، جزئیات بیشتری را به این مناطق می‌افزایند. این یافته‌ها با تحقیقات ژو و همکاران (۲۰۱۸) مطابقت دارد. در پژوهش آن‌ها نیز، رابطه بین تراکم و کامل بودن داده‌های ساختمان شهری در OpenStreetMap بررسی شده است. آن‌ها دریافتند، در مناطق دارای تراکم ساختمانی بیشتر، داده‌های کامل‌تری نیز وجود دارد. این بیان می‌کند کاربران به‌احتمال بسیار، در مناطقی که تراکم داده‌های ساختمانی بالاست، مشارکت بیشتری دارند. همچنین بر اهمیت تراکم، به‌منزله عاملی مؤثر در کامل بودن داده‌های ساختمانی OSM، تأکید می‌کند.





شکل ۱۰. همبستگی میان شاخص کامل بودن و تراکم در منطقه‌های (۹ و ۱۱) شهر مشهد، از سال ۲۰۱۸ تا ۲۰۲۳



شکل ۱۱. ارتقای داده‌های ساختمانی OSM در منطقه‌های (۹ و ۱۱) شهر مشهد، از ۲۰۱۸ تا ۲۰۲۳

نشریه سنجش از دور و GIS ایران

سال ۱۸، شماره ۱، بهار ۱۴۰۵

۴- نتیجه‌گیری

از آنجا که مجموعه داده‌های رسمی از ساختمان‌ها در مورد بسیاری از شهرها وجود ندارد و یا به‌رایگان در دسترس نیست، مجموعه‌ای از داده‌های ساختمانی رایگان می‌تواند برای پژوهشگران بسیار سودمند باشد. در این پژوهش، ارتقا و کامل بودن داده‌های ساختمانی در OSM تحلیل شد. بدین منظور، ابتدا داده‌های ساختمان‌ها در OSM و روند تهیه و پیش‌پردازش آن‌ها ارزیابی شد. در ادامه، منطقه مورد مطالعه (مناطق ۹ و ۱۱ مشهد) به شبکه‌ای سلولی به اندازه ۱×۱ کیلومتر تقسیم شد و این مساحت، به‌منزله واحد جغرافیایی، مبنای محاسبات قرار گرفت. در مورد هر واحد جغرافیایی، سه معیار (تراکم ساختمانی، کامل بودن و تعداد ساختمان در داده‌های OSM) در منطقه مورد نظر، برای سال‌های ۲۰۱۸ تا ۲۰۲۳، بررسی شد. سپس رابطه بین تراکم و کامل بودن داده‌های ساختمانی، در مورد هر سلول شبکه، محاسبه شد. همچنین دو پارامتر جمعیت و کاربری اراضی در بررسی‌ها به کار رفت. طبق نتایج، داده ساختمانی OSM در مناطق ۹ و ۱۱ مشهد، از ۲۰۱۸ تا ۲۰۲۳، بیست‌برابر افزایش داشته و این افزایش، به‌ویژه در مورد مناطق شرق و شمال‌شرق، مشهودتر است. کاربری اراضی مسکونی، مسکونی- تجاری، تجاری و آموزشی نیز، به‌ترتیب، بیشترین فراوانی را به خود اختصاص داده‌اند. با وجود همبستگی مثبت و معنادار بین داده‌های ساختمانی OSM و جمعیت هر بلوک ساختمانی، این همبستگی در بازه زمانی مورد بررسی دچار تغییرات نوسانی شده است. مقدار همبستگی‌ها میان شاخص تراکم و کامل بودن نشان می‌دهد، در مکان‌هایی که تراکم داده ساختمانی در داده OSM بیشتر بوده، کاربران تمایل بیشتری برای ایجاد و ترسیم ساختمان جدید داشته‌اند. همچنین با توجه به مقادیر شاخص کامل بودن می‌توان نتیجه گرفت که در مقایسه با داده‌های رسمی، ساختمان‌های به‌ثبت‌رسیده در OSM بسیار کمتر از تعداد واقعی ساختمان‌ها در داده رسمی‌اند.

با توجه به این تحقیق، می‌توان گفت که داده‌های OSM برای تحلیل جغرافیایی ساختمان‌ها در شهرهای

بزرگ، مانند مشهد، مناسب‌اند اما، برای استفاده‌های بیشتر و دقیق‌تر، این داده‌ها باید تکمیل و بهبود یابند. همچنین بررسی شاخص‌های تأثیرگذار می‌تواند اطلاعات سودبخشی برای استفاده مطمئن‌تر از این داده‌ها در اختیار کاربران قرار دهد.

در این پژوهش سعی شده است شاخص‌هایی به کار رود که اعتبار آن‌ها، در تحقیقات معتبر پیشین، تأیید شده باشد؛ بدیهی است که این شاخص‌ها کامل نیستند و نقاط ضعفی نیز خواهند داشت؛ از جمله ناتوانی در تشخیص ساختمان‌های ترسیم‌شده‌ای که در محیط خارجی وجود ندارند. بنابراین پیشنهاد می‌شود این مسئله در تحقیقات آتی مورد توجه قرار گیرد.

همچنین شاخص جمعیت نشان داد، در بلوک‌هایی که جمعیت بیشتری در آن‌ها ساکن بوده‌اند، میزان ترسیم ساختمان‌های جدید OSM نیز بیشتر بوده است. به‌رغم وجود همبستگی مثبت بین تعداد ترسیم‌ها و جمعیت در سال‌های متفاوت، این همبستگی نوسانی بوده است که نشان می‌دهد، علاوه بر جمعیت، شاخص‌های دیگری نیز باید بررسی شود. از این رو پیشنهاد می‌شود، در پژوهش‌های آتی، رابطه ترسیم ساختمان‌ها با سایر شاخص‌ها مد نظر قرار گیرد.

۵- منابع

- Adabi, R., Ali Abbaspour, R. & Chehreghan, A. 2022, **An Assessment on Building Blocks Completion Process in the Volunteered Geographic Information (VGI) for Using in Smart Cities**, Scientific- Research Quarterly of Geographical Data (SEPEHR), 30(120), PP. 27-42, <https://doi.org/10.22131/sepehr.2022.251052>.
- Bakillah, M. & Liang, S., 2016, **Open Geospatial Data**, Software and Standards, 1(1), PP. 1-2. <https://doi.org/10.1186/s40965-016-0004-1>.
- Elwood, S., Goodchild, M.F. & Sui, D.Z., 2012, **Researching Volunteered Geographic Information: Spatial Data, Geographic Research, and New Social Practice**, Annals of the Association of American Geographers, 102(3), PP. 571-590, <https://doi.org/10.1080/00045608.2011.595657>.

- Fan, H., Zipt, A., Fu, Q. & Neis, P., 2014, **Quality Assessment for Building Footprints Data on OpenStreetMap**, International Journal of Geographical Information Science, 28(4), PP. 700-719, <https://doi.org/10.1080/13658816.2013.867495>.
- Goodchild, F.M., 2007, **Citizens as Sensors: The World of Volunteered Geography**, GeoJournal, 69, PP. 211-221, <https://doi.org/10.1007/s10708-007-9111-y>.
- Hecht, R., Kunze, C. & Hahmann, S., 2013, **Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time**, International Journal of Geo-Information, 2(4), PP. 1066-1091, <https://doi.org/10.3390/ijgi2041066>.
- Herfort, B., Lautenbach, S., Porto de Albuquerque, J., Anderson, J. & Zipf, A., 2023, **A Spatio-Temporal Analysis Investigating Completeness and Inequalities of Global Urban Building Data in OpenStreetMap**, Nature Communications, 14(1), P. 3985, <https://doi.org/10.1038/s41467-023-39698-6>.
- Kresse, W. & Fadaie, K., 2004, **ISO Standards for Geographic Information**, Springer: Berlin, <https://doi.org/10.1007/978-3-662-08039-9>.
- Minaei, M., 2020, **Evolution, Density, and Completeness of OpenStreetMap Road Networks in Developing Countries: The Case of Iran**, Applied Geography, 119, P. 102246, <https://doi.org/10.1016/j.apgeog.2020.102246>.
- Mobasheri, A., 2017, **A Rule-Based Spatial Reasoning Approach for OpenStreetMap Data Quality Enrichment; Case Study of Routing and Navigation**, Sensors, 17(11), P. 2498, <https://doi.org/10.3390/s17112498>.
- Sadidi, J., tamnia, F. & Rezaian, H., 2024, **Assessment of Using Artificial Intelligence in Completeness of Volunteer Geographic Information. A Case Study for Open Street Map (OSM) Landuse Data**, Journal of Spatial Analysis Environmental Hazards, 11(1), P. 1, <https://doi.org/10.61186/jsaeh.11.1.3>.
- Sun, Y. & Du, Y., 2017, **Big Data and Sustainable Cities: Applications of New and Emerging Forms of Geospatial Data in Urban Studies**, Open Geospatial Data Software and Standards, 2(1), PP. 1-4, <https://doi.org/10.1186/s40965-017-0037-0>.
- Tian, Y. & Zhou, Q.F.X., 2019, **An Analysis of the Evolution, Completeness, and Spatial Patterns of OpenStreetMap Building Data in China**, ISPRS International Journal of Geo-Information, 8(1), P. 35, <https://doi.org/10.3390/ijgi8010035>.
- Törnros, T., Dorn, H., Hahmann, S. & Zipf, A., 2015, **Uncertainties of Completeness Measures in OpenStreetMap—A Case Study for Buildings in a Medium-Sized German City**, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2, PP. 353-357, <https://doi.org/10.5194/isprsannals-II-3-W5-353-2015>.
- Zhang, Y., Zhou, Q., Brovelli, M.A. & Li, W., 2022, **Assessing OSM Building Completeness Using Population Data**, International Journal of Geographical Information Science, 36(7), PP. 1443-1466, <https://doi.org/10.1080/13658816.2021.2023158>.
- Zhou, Q., 2018, **Exploring the Relationship between Density and Completeness of Urban Building Data in OpenStreetMap for Quality Estimation**, International Journal of Geographical Information Science, 32(2), PP. 257-281, <https://doi.org/10.1080/13658816.2017.1395883>.
- Zhou, Q., Zhang, Y., Chang, K. & Brovelli, M.A., 2022, **Assessing OSM Building Completeness for Almost 13,000 Cities Globally**, International Journal of Digital Earth, 15(1), PP. 2400-2421, <https://doi.org/10.1080/17538947.2022.2159550>.