

## Đánh giá tiềm năng của dữ liệu LiDAR không gian trong ước tính chiều cao công trình đô thị tại Hà Nội

Hoàng Tuấn Kiệt<sup>1</sup>, Vũ Phương Lan<sup>1</sup>, Nguyễn Phương Bắc<sup>1,\*</sup>

<sup>1</sup>Khoa Địa Lý, Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, Hà Nội, Việt Nam \*Email tác giả liên hệ: [nguyenphuongbac@hus.edu.vn](mailto:nguyenphuongbac@hus.edu.vn)

<https://doi.org/10.67484/mgc.14>

**Tóm tắt:** Nghiên cứu này khẳng định tiềm năng của dữ liệu LiDAR không gian từ hai hệ thống cảm biến ICESat-2 và GEDI trong việc tự động hóa trích xuất và ước tính chiều cao cấu trúc công trình tại khu vực nội thành Hà Nội. Quy trình công nghệ được thiết lập dựa trên sự tích hợp dữ liệu ranh giới nền công trình kết hợp quy trình lọc nhiễu đám mây điểm để phân loại chính xác các điểm phản hồi từ mái công trình và mặt đất tham chiếu. Dựa trên bộ dữ liệu nghiệm chứng gồm 200 điểm đo thực địa, kết quả phân tích thống kê cho thấy dữ liệu dạng sóng toàn vẹn của GEDI đạt độ tin cậy cao với hệ số xác định  $R^2 = 0.89$ , trong khi công nghệ đếm photon của ICESat-2 thể hiện ưu thế về độ phân giải không gian dài quét chi tiết với  $R^2 = 0.86$ . Kết quả nghiên cứu khẳng định ưu thế của nguồn dữ liệu ICESat-2 và GEDI có thể hỗ trợ cho các phép đo thực địa. Nghiên cứu mở ra triển vọng ứng dụng quy mô lớn cho việc tối ưu hóa công tác quy hoạch, giám sát động thái phát triển kiến trúc và đẩy nhanh tiến trình xây dựng cơ sở dữ liệu đô thị 3D tại các siêu đô thị ở Việt Nam hiện nay.

**Từ khóa:** LiDAR không gian, ICESat-2, GEDI, chiều cao công trình, cơ sở dữ liệu 3D, Hà Nội.

Ngày nhận bài: 26/05/2026 Ngày sửa lại: 17/06/2026 Ngày chấp nhận đăng: 18/06/2026 Ngày xuất bản: 30/08/2026

## Accuracy of the potential of spaceborne LiDAR data in estimating urban building heights in Hanoi

Tuan Kiet HOANG<sup>1</sup>, Phuong Lan Vu<sup>1</sup>, Phuong Bac NGUYEN<sup>1,\*</sup>

<sup>1</sup>Faculty of Geography, VNU University of Science, 334 Nguyen Trai, Thanh Xuan, Ha Noi, Vietnam

\*Corresponding Author Email: [nguyenphuongbac@hus.edu.vn](mailto:nguyenphuongbac@hus.edu.vn)

**Abstract:** This study evaluates the potential of spaceborne LiDAR data from the ICESat-2 and GEDI sensor systems for the automated extraction and estimation of urban building heights in the inner wards of Hanoi, Vietnam. The proposed workflow integrates building footprint boundary data with a point-cloud noise filtering procedure to accurately classify return signals from building rooftops and reference ground surfaces. Based on an independent validation dataset consisting of 200 field-survey points, the statistical analysis indicates that GEDI full-waveform LiDAR data achieved high reliability with a coefficient of determination ( $R^2$ ) of 0.89, whereas ICESat-2 photon-counting technology demonstrated advantages in detailed along-track spatial resolution with an  $R^2$  value of 0.86. The results highlight the capability of ICESat-2 and GEDI datasets to effectively supplement or partially replace conventional field-based surveying methods. This study also demonstrates strong potential for large-scale applications in optimizing urban planning, monitoring architectural development dynamics, and accelerating the development of 3D urban spatial databases in rapidly growing megacities across Vietnam.

**Key words:** Spaceborne LiDAR, ICESat-2, GEDI, building height, 3D databases, Hanoi.

Submission received: 26/05/2026

Revised: 17/06/2026

Accepted: 18/06/2026

Published: 30/08/2026

### 1. Giới thiệu

Sự gia tăng mật độ các công trình cao tầng đang trở thành đặc trưng hình thái nổi bật tại các siêu đô thị lớn ở các quốc gia đang phát triển, đặt ra những thách thức phức tạp cho công tác quy hoạch, quản lý không gian và mục tiêu phát triển bền vững [1]. Tại các thành phố như Hà Nội, nơi mật độ xây dựng cao và hình thái công trình phức tạp chiếm ưu thế, nhu cầu về dữ liệu không gian ba chiều chính xác và cập nhật ngày càng trở nên cấp thiết. Thông tin đáng tin cậy về chiều cao công trình không chỉ hỗ trợ công tác quản lý đô thị và quy hoạch hạ tầng [2], mà còn đóng vai trò quan trọng trong nhiều ứng dụng như ước tính dân số [3], mô hình hóa năng lượng và đánh giá rủi ro thiên tai [4].

Trong hệ thống các thông số cấu trúc không gian đô thị, chiều cao công trình là một biến số độc lập cốt lõi, chi phối trực tiếp đến các bài toán mô phỏng vi khí hậu nội đô, phân tích xu hướng đảo nhiệt đô thị và tối ưu hóa quy hoạch kiến trúc cảnh quan [5]. Theo cách tiếp cận truyền thống, việc thu thập dữ liệu độ cao tòa nhà thường dựa trên kỹ thuật đo đạc trắc địa mặt đất hoặc

viễn thám hàng không. Mặc dù các công nghệ này đảm bảo độ chính xác ở quy mô cục bộ, nhưng chúng lại bộc lộ những rào cản lớn về chi phí vận hành, giới hạn phạm vi bao phủ không gian và rất khó để cập nhật dữ liệu thường xuyên trên quy mô toàn thành phố hoặc quốc gia.

Các nghiên cứu xây dựng thành phố 3D ở khu vực lớn hiện dựa trên dữ liệu ảnh vệ tinh quang học cho kết quả khá tốt trên những khu vực rộng lớn, tuy nhiên phương pháp này cũng phụ thuộc mạnh vào điều kiện khí quyển như mây, sương mù, khói bụi và sự thay đổi điều kiện chiếu sáng, đặc biệt tại các khu vực nhiệt đới. Ngoài ra, quá trình ghép cặp điểm ảnh thường gặp khó khăn tại các bề mặt đồng nhất, khu vực đô thị mật độ cao và vùng có thảm thực vật dày, nơi hiện tượng che khuất, bóng đổ và tán cây có thể gây ra sai lệch, khoảng trống dữ liệu và nhiễu trong mô hình tái tạo. Chất lượng mô hình 3D còn phụ thuộc đáng kể vào hình học chụp ảnh, tỷ lệ baseline/height và sự hiện diện của các điểm khống chế mặt đất phục vụ hiệu chỉnh hình học. Bên cạnh đó, dữ liệu ảnh vệ tinh độ phân giải cao thường có chi phí lớn, trong khi các bước xử lý ảnh dày đặc và tái tạo 3D đòi hỏi năng lực tính toán và lưu trữ đáng kể.

Sự ra đời của các hệ thống viễn thám LiDAR không gian trong những năm gần đây đã thiết lập một phương thức hoàn toàn mới trong công tác quan trắc cấu trúc bề mặt thẳng đứng của Trái Đất một cách liên tục và trên diện rộng. Hai sứ mệnh tiên tiến và mang tính hỗ trợ cao nhất hiện nay là vệ tinh ICESat-2 [6] và thiết bị GEDI hoạt động trên Trạm Không gian Quốc tế ISS [7]. Mặc dù mục tiêu thiết kế ban đầu của hai cảm biến này hướng tới việc đo đặc độ dày dải băng và ước tính sinh khối tán rừng, các thử nghiệm phân tích gần đây đã chứng minh tiềm năng to lớn trong việc ứng dụng nguồn dữ liệu mở này để tự động hóa trích xuất chiều cao tòa nhà tại các khu vực đô thị [8-12]. Việc kết hợp các điểm LiDAR này với dữ liệu ranh giới tòa nhà 2D đang trở thành một giải pháp tối ưu để tái tạo các mô hình thành phố 3D LOD1 hoàn chỉnh và bao quát. Tuy nhiên, giới hạn khoa học hiện tại nằm ở chỗ bản chất vật lý và cơ chế phản xạ tín hiệu của hai cảm biến có sự khác biệt rất lớn. Sự tương tác của chùm tia laser vi xung cực nhạy của ICESat-2 và vết quét dạng sóng có đường kính phủ 25 m của GEDI đối với hình thái kiến trúc đô thị nên vẫn chưa được đánh giá đầy đủ. Đặc biệt với ICESat-2, lượng nhiễu nền sinh ra do bức xạ mặt trời vào ban ngày tại các khu vực đô thị là rất lớn [13], đòi hỏi phải có các thuật toán lọc nhiễu tự thích ứng phức tạp.

Tại khu vực lõi nội thành Hà Nội, đặc trưng bởi sự đan xen của nhà ống san sát, mái dốc, và các tổ hợp chung cư vách kính hiện đại, hiện tượng đa tán xạ và hiệu ứng rìa công trình (edge effects) xảy ra vô cùng phức tạp. Các thuật toán khử nhiễu và trích xuất độ cao hiện hành phân lớn được thiết kế dựa trên hình thái đô thị phương Tây, nên việc áp dụng chúng vào môi trường nhiệt đới gió mùa tại Việt Nam vẫn còn tiềm ẩn nhiều sai số.

Xuất phát từ những khoảng trống nghiên cứu nêu trên, bài báo này nhằm mục tiêu đánh giá tiềm năng ước tính chiều cao công trình của dữ liệu ICESat-2 và GEDI tại khu vực nội thành Hà Nội, thông qua so sánh với dữ liệu nghiệm chứng thực địa độc lập. Trên cơ sở đó, nghiên cứu phân tích ưu điểm và hạn chế của từng loại cảm biến, qua đó đề xuất cơ sở khoa học cho việc tự động hóa xây dựng mô hình đô thị 3D (LOD1), phục vụ hiệu quả công tác quản lý không gian và quy hoạch đô thị.

## **2. Khu vực nghiên cứu và dữ liệu sử dụng**

### **2.1. Khu vực nghiên cứu**

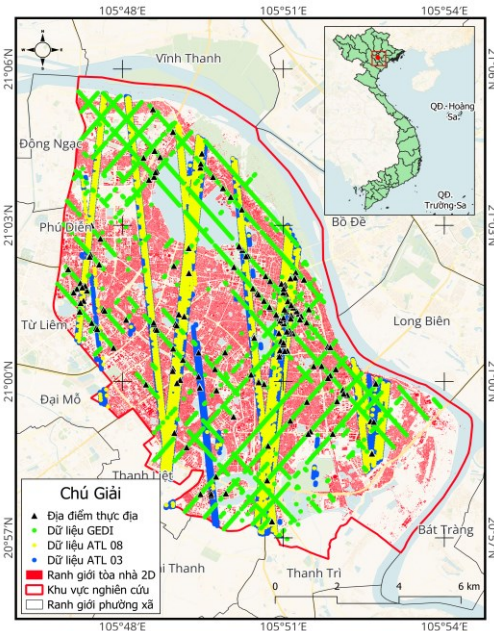
Khu vực nghiên cứu được thiết lập trong phạm vi 30 phường nội thành Thủ đô Hà Nội. Đây là khu vực có mật độ xây dựng cao nhất cả nước với hình thái kiến trúc đa dạng, đan xen giữa các khu phố cổ, phố cũ và các khu đô thị mới với nhiều nhà cao tầng. Đặc điểm bề mặt đô thị tại đây cực kỳ phức tạp với các khối nhà san sát, tạo ra những thách thức lớn cho việc phân tách tín hiệu LiDAR vệ tinh giữa nóc công trình và mặt đất tự nhiên.

### **2.2. Dữ liệu sử dụng**

#### **2.2.1 Dữ liệu ICESat-2**

Vệ tinh ICESat-2 được trang bị thiết bị đo cao laser địa hình (Advanced Topographic

Laser Altimeter System - ATLAS). ATLAS hoạt động dựa trên công nghệ đếm photon với chùm tia laser màu lục có bước sóng 532 nm. Hệ thống phát ra 6 vết sáng chia thành 3 cặp mạnh - yếu, cung cấp mật độ điểm cực kỳ dày đặc dọc theo quỹ đạo quét.



Hình 1: Khu vực nghiên cứu và phân bố dữ liệu vệ tinh ICESat-2 và GEDI

Trong nghiên cứu này, sản phẩm dữ liệu đám mây điểm ATL03 [14] được sử dụng để xác định độ cao công trình, trong khi sản phẩm ATL08 [15] được khai thác nhằm xác định độ cao nền đất. Dữ liệu ATL03 có ưu điểm về mức độ chi tiết cao, cho phép mô tả tốt đặc trưng hình thái đô thị; tuy nhiên, dữ liệu này cũng rất nhạy cảm với nhiễu gây ra bởi bức xạ mặt trời phản xạ từ các bề mặt vật liệu đô thị như kính và mái tôn, dẫn đến sự xuất hiện của số lượng lớn điểm nhiễu xen lẫn trong tín hiệu địa vật [13]. Trong khi đó, ATL08 là sản phẩm dẫn xuất đã được xử lý và phân loại tín hiệu photon từ ATL03, cho phép tách lọc tương đối ổn định các điểm địa hình và thảm phủ bề mặt, qua đó phù hợp cho việc xác định độ cao nền đất trong môi trường đô thị. Trong phạm vi nghiên cứu, tổng cộng 36 track ATL03 được thu thập trong giai đoạn từ 19/11/2018 đến 13/06/2025, cùng với 14 track ATL08 được thu thập từ 25/06/2019 đến 06/11/2024. Phân bố các vết quét của dữ liệu được minh họa trong Hình 1.

#### 2.2.2 Dữ liệu GEDI

Thiết bị GEDI (Global Ecosystem Dynamics Investigation) hoạt động trên Trạm Vũ trụ Quốc tế, sử dụng công nghệ LiDAR dạng sóng toàn phần với tia laser cận hồng ngoại có bước sóng 1064 nm. Hệ thống này quét bề mặt Trái Đất thông qua các vết quét dạng tròn có đường kính xấp xỉ 25 m. Trong nghiên cứu này, sản phẩm dữ liệu GEDI Level 2A được sử dụng nhằm khai thác thông tin phân bố năng lượng tín hiệu theo dạng sóng, đặc biệt là các đặc trưng liên quan đến đỉnh và đáy sóng phản xạ. So với dữ liệu đếm photon như ICESat-2 ATL03, vết quét có kích thước lớn của GEDI cho phép bao phủ phần lớn khối công trình trong môi trường đô thị, từ đó giảm ảnh hưởng của các nhiễu cục bộ gây ra bởi các chi tiết kiến trúc nhỏ và cung cấp tín hiệu phản xạ tổng thể ổn định hơn [16]. Trong phạm vi nghiên cứu, tổng cộng 36 track dữ liệu GEDI L2A được thu thập trong giai đoạn từ 12/08/2019 đến 07/01/2025 (Hình 1).

#### 2.2.3 Dữ liệu Google Open Buildings

Bên cạnh dữ liệu vệ tinh, nghiên cứu khai thác bộ cơ sở dữ liệu mở Google Open Buildings, đây một dự án lập bản đồ quy mô toàn cầu do Google Research phát triển, sử dụng các mô hình học sâu dựa trên mạng nơ-ron tích chập để tự động phát hiện và số hóa đa giác công trình từ ảnh vệ tinh quang học độ phân giải cao [17]. Bộ dữ liệu này được sử dụng như một lớp mặt nạ không gian nhằm khoanh vùng chính xác các tín hiệu LiDAR phản xạ từ bề mặt mái công trình, đồng thời giảm thiểu ảnh hưởng của các tín hiệu nhiễu phát sinh từ mặt đường, thảm thực vật hoặc mặt nước. Nhờ đó, các giá trị cao độ thu nhận từ dữ liệu ICESat-2 và GEDI có thể được gán chính xác cho từng công trình độc lập, góp phần nâng cao độ tin cậy của quá trình trích xuất độ cao đô thị.

#### 2.2.4 Dữ liệu thực địa

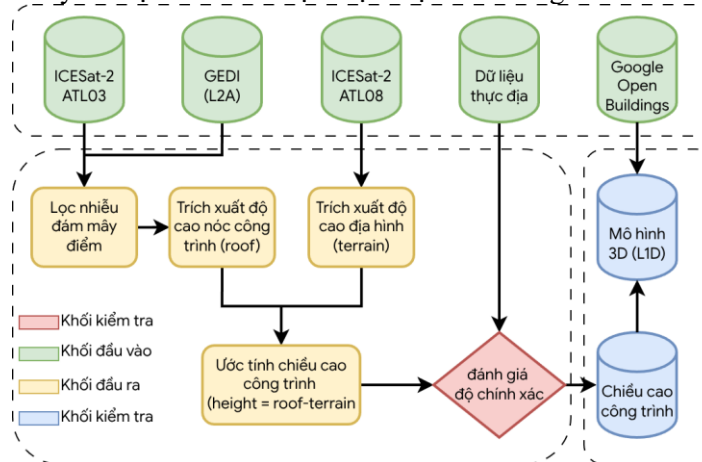
Để đánh giá độ chính xác cũng như kiểm định tính khách quan của kết quả nghiên cứu,

nhóm tác giả đã tiến hành khảo sát thực địa và thu thập dữ liệu tại 200 điểm công trình phân bố trong khu vực nghiên cứu. Các điểm khảo sát này đều nằm trong vùng chồng phủ của hai vệ tinh, qua đó đảm bảo tính đồng nhất và khả năng đối sánh trực tiếp giữa dữ liệu vệ tinh với số liệu đo đạc thực tế ngoài hiện trường. Tại mỗi vị trí chiều cao tòa nhà được thu thập thông qua phỏng vấn ban quản lý tòa nhà, chủ sở hữu hoặc những người có hiểu biết rõ về đặc điểm và quy mô của công trình. Bộ dữ liệu thực địa này được sử dụng làm cơ sở để so sánh, đối chiếu và đánh giá sai số giữa kết quả nghiên cứu và dữ liệu thực tế.

Bộ dữ liệu thực địa bao gồm 200 công trình với chiều cao phân bố không đồng đều giữa các nhóm độ cao. Trong đó, nhóm công trình trung thấp tầng (15–30 m) chiếm ưu thế tuyệt đối với 128 mẫu, tương đương 64,0% tổng số công trình khảo sát. Nhóm công trình trung tầng (30–50 m) và cao tầng (50–100 m) lần lượt gồm 35 và 26 mẫu, phản ánh sự hiện diện đáng kể của các công trình phát triển theo chiều cao trong khu vực nghiên cứu. Trong khi đó, các công trình thấp tầng (0–15 m) và siêu cao tầng (>100 m) chỉ chiếm tỷ lệ nhỏ, lần lượt với 5 và 6 mẫu. Phân bố này cho thấy tập dữ liệu thực địa tập trung chủ yếu vào các công trình có chiều cao trung bình đến cao, phù hợp với đặc trưng phát triển đô thị nén và gia tăng mật độ xây dựng theo chiều đứng tại khu vực nội đô. Nguyên nhân dẫn đến việc thiếu tính đại diện cho đô thị là do tập mẫu được lấy theo vệt quét của hai nguồn dữ liệu viễn thám nhằm mục đích đối chiếu, so sánh và kiểm định kết quả ước tính chiều cao công trình.

### 3. Phương pháp nghiên cứu

Quy trình xử lý dữ liệu LiDAR được thực hiện như trong Hình 2:



Hình 2: Sơ đồ quy trình nghiên cứu

#### 3.1. Lọc nhiễu và chuẩn hóa dữ liệu

Trong giai đoạn tiền xử lý, thách thức chính đối với dữ liệu LiDAR thô là sự tồn tại của các tín hiệu nhiễu phát sinh từ tán xạ khí quyển, bức xạ mặt trời nền và các phản xạ không mong muốn từ bề mặt địa vật. Đối với dữ liệu ICESat-2 ATL03, tập photon thô được lọc nhiễu thông qua cờ phân loại độ tin cậy tín hiệu, nhằm loại bỏ các photon nhiễu có phân bố ngẫu nhiên và giữ lại các photon có xác suất cao đại diện cho tín hiệu phản xạ thực từ bề mặt địa vật.

Tương tự, đối với dữ liệu GEDI L2A, quá trình kiểm soát chất lượng được thực hiện thông qua trường `quality_flag = 1`, cho phép loại bỏ các vệt quét có chất lượng thấp hoặc bị ảnh hưởng bởi mây che phủ, suy hao khí quyển và nhiễu tín hiệu.

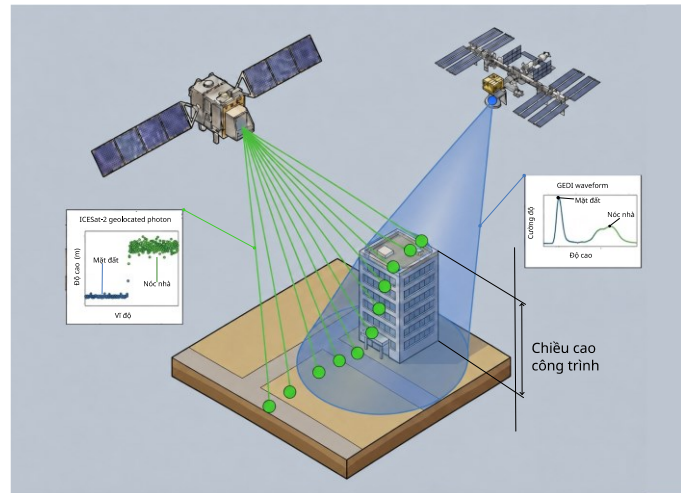
Sau bước làm sạch dữ liệu, toàn bộ đám mây điểm từ hai hệ thống được chuyển đổi và chuẩn hóa về cùng một hệ quy chiếu không gian nhằm đảm bảo tính nhất quán hình học trong quá trình phân tích. Trên cơ sở đó, dữ liệu tiếp tục được phân tách thành hai nhóm cấu trúc chính gồm lớp mặt đất và lớp phi mặt đất, phục vụ cho quá trình trích xuất độ cao công trình và mô hình hóa bề mặt đô thị.

#### 3.2. Ước tính chiều cao công trình

Để tự động xác định chiều cao của các công trình kiến trúc tại khu vực nghiên cứu, các



photon từ dữ liệu ICESat-2 và GEDI cần phải được phân loại thành photon bề mặt công trình và photon mặt đất. Nhằm tối ưu hóa độ chính xác và độ tin cậy của quá trình phân loại, đồng thời nâng cao độ chính xác khi xác định chiều cao công trình, dữ liệu được xử lý theo các bước chính. Đầu tiên, tập hợp các điểm tín hiệu thuộc lớp mặt đất sẽ được đưa vào thuật toán nội suy nhằm kiến tạo bề mặt Mô hình độ cao kỹ thuật số (DTM) cục bộ. Bề mặt DTM này đóng vai trò là mặt phẳng tham chiếu cơ sở cho từng khu vực cụ thể. Các tín hiệu LiDAR phản xạ từ nóc tòa nhà sẽ được tham chiếu xuống bề mặt DTM tương ứng. Sau khi áp dụng quy trình lọc nhiễu và phân loại tín hiệu, bộ dữ liệu được lọc để loại bỏ các điểm ngoại lai. Chỉ những điểm tín hiệu có độ tin cậy cao, xác định rõ ràng là phản xạ từ nóc các tòa nhà tại khu vực trung tâm mới được giữ lại để tính toán chiều cao. Chiều cao thực của khối công trình chính là khoảng cách không gian từ mặt phẳng địa hình cơ sở lên đến điểm tiếp nhận tín hiệu cao nhất - có thể là đỉnh cụm photon đối với ICESat-2 hoặc đỉnh năng lượng của đội sóng đầu tiên đối với GEDI (Hình 3)



Hình 3: Nguyên lý của ICESat-2 và GEDI trong ước tính chiều cao công trình đô thị

### 3.3. Đánh giá độ chính xác

Nhằm đánh giá một cách khách quan độ tin cậy và nhận diện các giới hạn kỹ thuật của từng hệ thống cảm biến, chiều cao công trình trích xuất từ mô hình được đối sánh trực tiếp với tập dữ liệu kiểm chứng thực địa (ground-truth).

Độ chính xác ước tính chiều cao của hai nguồn dữ liệu vệ tinh. Kết quả được đánh giá thông qua ba chỉ số thống kê phổ biến, bao gồm hệ số xác định ( $R^2$ ), sai số toàn phương trung bình (RMSE) và sai số tuyệt đối trung bình (MAE). Trong đó, hệ số xác định ( $R^2$ ) phản ánh mức độ tương quan và khả năng giải thích biến thiên hình học giữa dữ liệu ước tính và dữ liệu thực địa, được xác định theo công thức:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

Trong khi đó, RMSE và MAE được sử dụng như các thước đo trực tiếp nhằm định lượng mức độ sai lệch tuyệt đối về cao độ giữa dữ liệu viễn thám và giá trị thực tế:

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

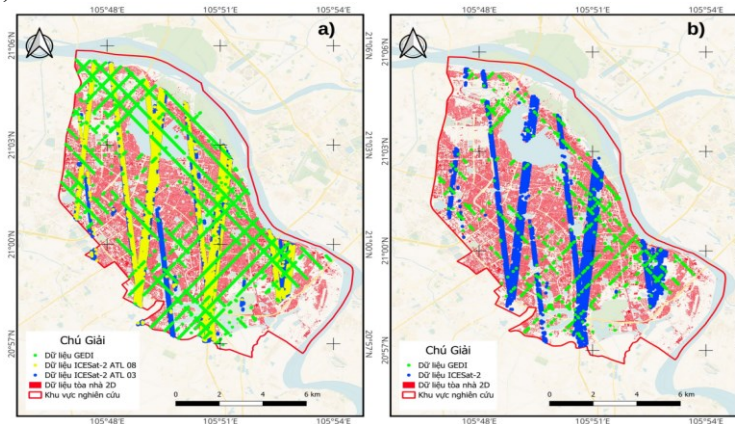
$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

## 4. Kết quả nghiên cứu

### 4.1. Kết quả ước tính chiều cao công trình từ dữ liệu ICESat 2 và GEDI

Dữ liệu LiDAR không gian ban đầu thu nhận được dưới dạng các vệt quét rời rạc bao phủ toàn bộ thành phố. Ở trạng thái thô, mật độ dữ liệu rất dày đặc nhưng chứa lượng nhiễu lớn từ bức xạ mặt trời và các đối tượng địa vật không mong muốn. Sau khi áp dụng quy trình lọc nhiễu và phân loại tín hiệu, bộ dữ liệu đã được tinh lọc đáng kể. Chỉ những điểm tín hiệu có độ tin cậy cao, xác định rõ ràng là phản xạ từ nóc các tòa nhà tại khu vực trung tâm mới được giữ lại để tính toán chiều cao (Hình 4). Việc tinh lọc này giúp loại bỏ các giá trị sai lệch (outliers), tạo cơ sở dữ liệu sạch cho bước đánh giá độ chính xác tiếp theo. Dữ liệu LiDAR không gian ban đầu được thu nhận dưới dạng các dải quét rời rạc phân bố trên toàn bộ khu vực nghiên cứu. Ở trạng thái thô, tập dữ liệu có mật độ điểm rất cao nhưng đồng thời chứa nhiều nhiễu phát sinh từ bức xạ mặt trời, phản

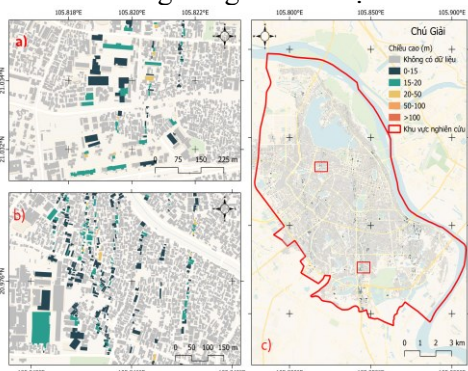
xạ khí quyển và các đối tượng địa vật không liên quan (Hình a). Để nâng cao chất lượng dữ liệu, nghiên cứu đã áp dụng quy trình lọc tín hiệu loại bỏ các điểm nhiễu và các phản xạ không mong muốn (Hình b)



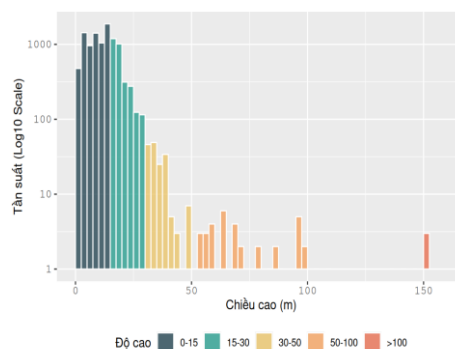
Hình 4: Mật độ dữ liệu ICESat-2 và GEDI trước và sau khi lọc

Hình 5 **Error! Reference source not found.** c Thể hiện dữ liệu chiều cao được phân loại thành các khoảng 0–15 m, 15–20 m, 20–50 m, 50–100 m và trên 100 m nhằm phản ánh cấu trúc phát triển theo chiều đứng của khu vực nghiên cứu. Kết quả chi tiết tại hai khu vực mẫu (Hình 5a và Hình 5b) cho thấy mô hình có khả năng tái hiện tương đối tốt sự biến thiên chiều cao ở quy mô công trình riêng lẻ. Các công trình có diện tích mái lớn và hình học đơn giản được xác định ổn định với vết quét đồng nhất và ranh giới rõ ràng. Tại khu vực có mật độ xây dựng dày đặc, mô hình vẫn duy trì khả năng phân biệt giữa các nhóm chiều cao khác nhau, mặc dù xuất hiện hiện tượng phân mảnh cục bộ đối với các công trình nhỏ, hẹp hoặc nằm sát nhau. Ngược lại, ở các khu vực có cấu trúc đô thị thưa hơn, kết quả ước tính cho thấy tính liên tục và độ ổn định cao hơn trong biểu diễn chiều cao công trình.

Hình 6 thể hiện phân bố chiều cao công trình thể hiện dạng lệch phải rõ rệt, với phần lớn công trình có chiều cao dưới 30 m. Tần suất xuất hiện giảm nhanh khi chiều cao tăng, tạo nên phân bố đuôi dài. Các công trình cao trên 50 m xuất hiện tương đối ít, trong khi chỉ có một số ít ngoại lệ vượt quá 100 m, đạt tới khoảng 150 m. Kết quả này phản ánh sự chiếm ưu thế của các công trình thấp và trung tầng trong khu vực nghiên cứu, đồng thời cho thấy tiềm năng mở rộng theo chiều thẳng đứng đối với nội đô Hà Nội là rất lớn.



Hình 5: Kết quả ước tính chiều cao công trình tại Hà Nội

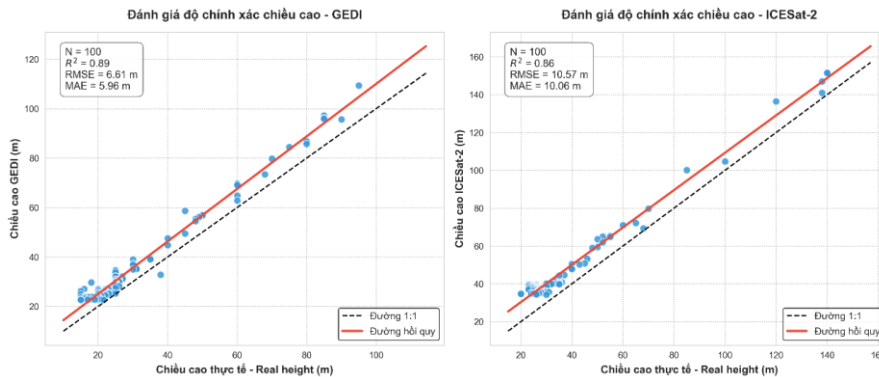


Hình 8: Dữ liệu phân bố thống kê chiều cao công trình tại một số quận trung tâm Hà Nội được trích xuất từ dữ liệu vệ tinh

#### 4.2. Đánh giá độ chính xác

Để làm rõ biên độ sai số cũng như xu hướng ước tính của từng thiết bị, kết quả đối chiếu giữa mô hình và thực địa được trực quan hóa bằng sơ đồ phân tán (Hình ). Với lượng mẫu kiểm chứng

$N = 200$ , các chỉ số thống kê thực nghiệm chỉ ra sự phân hóa rõ rệt về năng lực đo cao giữa hai công nghệ LiDAR đếm photon của ICESat-2 và LiDAR dạng sóng đầy đủ của GEDI. Cảm biến GEDI thể hiện ưu thế vượt trội trong môi trường đô thị với hệ số xác định  $R^2 = 0.89$ , đồng thời sai số  $RMSE = 6.61$  m và  $MAE = 5.96$  m. Trong khi đó, các chỉ số tương ứng của cảm biến ICESat-2 cho thấy mức độ chính xác thấp hơn, khi hệ số  $R^2$  giảm xuống 0.86, còn sai số  $RMSE$  và  $MAE$  tăng cao, lần lượt là 10.57 m và 10.06 m.



Hình 6: Biểu đồ phân tán so sánh dữ liệu thực và chiều cao ước tính từ hai vệ tinh ICESat-2 và GEDI

Biểu đồ phân tán của GEDI chỉ ra rằng các điểm dữ liệu bám rất sát đường lý tưởng 1:1, và đường hồi quy tuyến tính gần như trùng với đường chuẩn trong khoảng chiều cao từ 20 m đến hơn 100 m. Điều này chứng minh công nghệ LIDAR full-waveform của GEDI, với kích thước vệt quét khoảng 25 m, có khả năng tích phân và ghi nhận năng lượng phản xạ trung bình từ toàn bộ bề mặt mái công trình một cách ổn định, ít bị tác động bởi các yếu tố gây nhiễu đơn lẻ.

Ngược lại, hệ thống biểu đồ của ICESat-2 hiển thị độ phân tán lớn hơn, đồng thời đường hồi quy có xu hướng lệch dần xuống phía dưới đường lý tưởng 1:1 khi chiều cao công trình tăng lên. Xu hướng đánh giá thấp này là hệ quả trực tiếp của lỗi rìa công trình kết hợp với nhiều nền đô thị phức tạp. Do ICESat-2 sử dụng công nghệ LiDAR đếm photon với kích thước footprint nhỏ (xấp xỉ 11 m), các photon phát ra có xu hướng dễ bị ảnh hưởng bởi cấu trúc hình học phức tạp của môi trường đô thị. Trong nhiều trường hợp, photon không trực tiếp phản xạ từ đỉnh mái cao nhất mà có thể rơi vào phần rìa mái, bề mặt nghiêng hoặc khoảng trống giữa các công trình cao tầng liền kề. Ngoài ra, tín hiệu phản hồi cũng có thể bị tác động bởi tán cây hoặc các đối tượng che phủ xung quanh công trình, làm gia tăng sai lệch trong quá trình xác định chính xác độ cao mái nhà. Mặc dù có biên độ sai số lớn hơn, ICESat-2 vẫn đóng vai trò là nguồn dữ liệu bổ trợ không thể thiếu nhờ lợi thế về mật độ điểm dày và độ phân giải không gian bình diện cao, mở ra triển vọng lớn nếu được tối ưu hóa bằng các thuật toán học máy nhằm hiệu chỉnh sai số hình học rìa công trình trong các nghiên cứu tiếp theo.

#### 4.2. Xây dựng mô hình 3D

Dựa trên tập dữ liệu độ cao công trình đã được trích xuất từ nguồn dữ liệu LiDAR vệ tinh và kiểm định độ chính xác thông qua tập dữ liệu thực địa độc lập, nghiên cứu tiếp tục tích hợp các giá trị độ cao này với lớp dữ liệu không gian 2D dạng building footprints nhằm xây dựng mô hình đô thị 3D cho toàn bộ khu vực trung tâm Hà Nội ở mức độ chi tiết LOD1 (Level of Detail 1). Trong mô hình này, mỗi công trình được biểu diễn dưới dạng khối hộp ba chiều với chiều cao tương ứng với giá trị độ cao đã được hiệu chỉnh, qua đó cho phép tái hiện tổng quan hình thái không gian đô thị trên phạm vi rộng. Việc kết hợp giữa dữ liệu độ cao và lớp bản đồ nền 2D không chỉ giúp trực quan hóa cấu trúc phân bố công trình trong môi trường ba chiều mà còn tạo tiền đề cho các ứng dụng chuyên sâu như phân tích mật độ xây dựng, đánh giá cảnh quan đô thị, mô phỏng che phủ bóng đổ, quy hoạch không gian và quản lý phát triển đô thị thông minh.

Mô hình 3D được xây dựng cho thấy khả năng phản ánh rõ nét sự khác biệt về chiều cao, mật độ và đặc trưng hình thái giữa các khu vực chức năng trong nội đô Hà Nội, từ các khu hành chính – chính trị với mật độ xây dựng thấp và công trình quy mô lớn, đến các khu dân cư có cấu trúc xây dựng dày đặc hoặc các cụm cao tầng hiện đại với biến thiên độ cao phức tạp. Đồng thời,



việc trực quan hóa dữ liệu dưới dạng không gian ba chiều giúp nâng cao khả năng nhận diện các quy luật phát triển đô thị theo chiều đứng, vốn khó quan sát khi chỉ sử dụng dữ liệu bản đồ hai chiều truyền thống.

Để đánh giá toàn diện khả năng tái tạo hình thái kiến trúc và mức độ đáp ứng của mô hình đối với các dạng cấu trúc đô thị khác nhau, nghiên cứu tập trung phân tích chi tiết tại ba khu vực điển hình đại diện cho các kiểu không gian kiến trúc đặc thù của Hà Nội. Các khu vực này bao gồm: (i) khu vực công trình hành chính – chính trị với không gian mở và hình thái kiến trúc đồng nhất, (ii) khu vực đô thị mật độ cao với cấu trúc xây dựng phức tạp và xen kẽ nhiều loại hình công trình, và (iii) khu vực tập trung các tòa nhà cao tầng hiện đại có sự biến thiên lớn về chiều cao và hình khối kiến trúc. Thông qua việc so sánh trực quan giữa mô hình 3D và đặc điểm thực tế của từng khu vực, nghiên cứu đánh giá khả năng thu nhận tín hiệu, độ ổn định của dữ liệu độ cao cũng như mức độ chính xác trong quá trình tái tạo không gian đô thị từ dữ liệu LiDAR vệ tinh.



Hình 7: Kết quả chiều cao và mô hình 3D LID tại khu vực Lăng Chủ tịch Hồ Chí Minh



Hình 8: Kết quả chiều cao và mô hình 3D LID tại khu chung cư HH Linh Đàm

Hình thể hiện khu vực Lăng Chủ tịch Hồ Chí Minh, đây là công trình kiến trúc mang tính biểu tượng với khối tích lớn, mái bằng và không gian xung quanh tương đối quang đãng. Việc không có các công trình cao tầng lân cận che khuất giúp dữ liệu LiDAR (đặc biệt là GEDI) bắt được tín hiệu nóc công trình cực kỳ ổn định. Mô hình 3D tái tạo chính xác tỷ lệ, hình khối và độ bề thế của công trình so với cảnh quan quảng trường xung quanh. Bản đồ độ cao thể hiện sự đồng nhất cao về màu sắc trên toàn bộ diện tích mái.

Hình minh họa khu chung cư HH Linh Đàm ở trung tâm. Đây một trong những cụm dân cư có mật độ xây dựng cao nhất Hà Nội, với các tòa tháp cao xấp xỉ 40 tầng mọc san sát nhau. Kết quả trực quan hóa cho thấy mô hình vẫn phân tách thành công các khối nhà siêu cao tầng khỏi các công trình dịch vụ thấp tầng lân cận. Dù vậy, khoảng cách quá hẹp giữa các tháp nhà cũng bộc lộ rõ thách thức đối với vết quét của ICESat-2 khi rất dễ xảy ra hiện tượng nhiễu tín hiệu chéo giữa các tòa nhà.



Hình minh họa mô hình 3D của khu đô thị Royal City. Khu vực này mang đặc trưng của một tổ hợp đô thị hiện đại, kết hợp giữa khối đế thương mại rộng lớn và các tháp chung cư cao tầng bố trí theo hình vòng cung. Mô hình 3D tái hiện rất sắc nét sự giật cấp về chiều cao giữa khối đế và tháp ở trên. Diện tích bề mặt mái lớn của các khối nhà giúp các vết quét của vệ tinh dễ dàng xác định được đỉnh sóng phản xạ, mang lại kết quả mô phỏng không gian 3D sát với thực trạng.

## 5. Kết luận

Nghiên cứu đã thực hiện đánh giá thực nghiệm và kiểm chứng độ chính xác của dữ liệu LiDAR không gian trong việc trích xuất chiều cao công trình tại khu vực nội thành Hà Nội. Dựa trên 200 điểm thực địa nghiệm chứng độc lập, kết quả khẳng định: Cả hai vệ tinh đều có khả năng



nhận diện hình thái đô thị, nhưng dữ liệu dạng sóng của GEDI mang lại độ tin cậy về cao độ tuyệt đối vượt trội ( $R^2 = 0.89$ ,  $RMSE = 6.61$  m) so với công nghệ đếm photon của ICESat-2 ( $R^2 = 0.86$ ,  $RMSE = 10.57$  m).

Kết quả này cung cấp cơ sở khoa học quan trọng cho việc ưu tiên lựa chọn nguồn dữ liệu GEDI trong các dự án xây dựng cơ sở dữ liệu không gian 3D cấp độ thành phố. Mặc dù vẫn còn tồn tại giới hạn về độ bao phủ không gian theo vệt bay, LiDAR không gian đã chứng minh được tiềm năng là nguồn dữ liệu hỗ trợ hiệu quả, có khả năng cập nhật liên tục và tiết kiệm chi phí cho công tác quản lý quy hoạch đô thị tại Việt Nam.

### Lời cảm ơn

Nghiên cứu này được tài trợ bởi đề tài mã số TNK.26.12, của Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội

### Cam kết của các tác giả

Tất cả các tác giả có tên trong bài báo cam kết sự đồng thuận và không có xung đột lợi ích trong công bố khoa học tại bài báo này.

### Tài liệu tham khảo

- [1] S. Chakraborty và c.s., “In pursuit of sustainability – Spatio-temporal pathways of urban growth patterns in the world’s largest megacities”, *Cities*, tập 131, tr 103919, 2022, <https://doi.org/10.1016/j.cities.2022.103919>.
- [2] R. Neuville, J. Pouliot, F. Poux, và R. Billen, “3D Viewpoint Management and Navigation in Urban Planning: Application to the Exploratory Phase”, *Remote Sensing*, tập 11, số 3, 2019, <https://doi.org/10.3390/rs11030236>.
- [3] F. Schug, D. Frantz, S. van der Linden, và P. Hostert, “Gridded population mapping for Germany based on building density, height and type from Earth Observation data using census disaggregation and bottom-up estimates”, *PLOS ONE*, tập 16, số 3, tr 1–23, thg. 3 2021, <https://doi.org/10.1371/journal.pone.0249044>.
- [4] A. Ozturk, M. E. Atik, M. M. Koşucu, và S. O. Atik, “A Framework for 3D Flood Analysis Using an Open-Source Game Engine and Geospatial Data: A Case Study of the Bozkurt District of Kastamonu, Türkiye”, *Geomatics*, tập 5, số 3, 2025, <https://doi.org/10.3390/geomatics5030046>.
- [5] Z. Zhao, H. Li, và S. Wang, “The mutual impacts of individual building design and local microclimate in high-density cities and the major influential parameters”, *Building Simulation*, tập 18, số 1, tr 99–121, thg. 1 2025, <https://doi.org/10.1007/s12273-024-1195-5>.
- [6] T. Markus và c.s., “The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation”, *Remote Sensing of Environment*, tập 190, tr 260–273, 2017, <https://doi.org/10.1016/j.rse.2016.12.029>.
- [7] R. Dubayah và c.s., “The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography”, *Science of Remote Sensing*, tập 1, tr 100002, 2020, <https://doi.org/10.1016/j.srs.2020.100002>.
- [8] Y. Kaya, “Automated Estimation of Building Heights with ICESat-2 and GEDI LiDAR Altimeter and Building Footprints: The Case of New York City and Los Angeles”, *Buildings*, tập 14, số 11, 2024, <https://doi.org/10.3390/buildings14113571>.
- [9] B. Li và Q. Weng, “Estimation of Urban Building Heights From ICESat-2 Laser Altimeter Data and Quality Control”, *IEEE Transactions on Geoscience and Remote Sensing*, tập 64, tr 5702118–5702118, 2026, <https://doi.org/10.1109/TGRS.2026.3687994>.
- [10] X. Ma và c.s., “Mapping fine-scale building heights in urban agglomeration with spaceborne lidar”, *Remote Sensing of Environment*, tập 285, tr 113392, 2023, <https://doi.org/10.1016/j.rse.2022.113392>.
- [11] X. Ma và c.s., “A global product of 150-m urban building height based on spaceborne lidar”, *Scientific Data*, tập 11, số 1, tr 1387, thg. 12 2024, <https://doi.org/10.1038/s41597-024-04237-5>.
- [12] X. Tang và c.s., “A flexible framework for built-up height mapping using ICESat-2 photons and multisource satellite observations”, *Remote Sensing of Environment*, tập 318, tr 114572, 2025, <https://doi.org/10.1016/j.rse.2024.114572>.
- [13] J. Liu, J. Liu, H. Xie, D. Ye, và P. Li, “A Multi-Level Auto-Adaptive Noise-Filtering Algorithm for Land ICESat-2 Photon-Counting Data”, *Remote Sensing*, tập 15, số 21, 2023, <https://doi.org/10.3390/rs15215176>.
- [14] T. Neumann và c.s., “ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 6”. NASA

- National Snow and Ice Data Center Distributed Active Archive Center, 2023. <https://doi.org/10.5067/ATLAS/ATL03.006>.
- [15] A. Neuenschwander và c.s., “ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 7”. NASA National Snow and Ice Data Center Distributed Active Archive Center, 2025. <https://doi.org/10.5067/ATLAS/ATL08.007>.
- [16] S. Hancock và c.s., “The GEDI Simulator: A Large-Footprint Waveform Lidar Simulator for Calibration and Validation of Spaceborne Missions”, *Earth and Space Science*, tập 6, số 2, tr 294–310, 2019, <https://doi.org/10.1029/2018EA000506>.
- [17] W. Sirko và c.s., “Continental-Scale Building Detection from High Resolution Satellite Imagery”, *arXiv e-prints*, tr arXiv:2107.12283, thg. 7 2021, <https://doi.org/10.48550/arXiv.2107.12283>.

Article © 2026 by Magazine of Geodesy - Cartography is licensed under CC BY 4.0

