



Original Article

Research and Experimental Comparison of Topographic Modeling Methods

Phan Quoc Yen*, Nguyen Thi Thu Nga, Tong Thi Hanh

Military Technical Academy, 236 Hoang Quoc Viet, Hanoi, Vietnam

Received 12 September 2019

Revised 17 November 2019; Accepted 18 November 2019

Abstract: The topography of the earth's surface can be represented in GIS by DEM data. Surface modeling is the process of determining natural or artificial surfaces using one or more mathematical equations. A general surface modeling algorithm for all applications is not available, each method of creating a topographic surface has several advantages, disadvantages and depends on its processing direction. As such, experimenting, evaluating and selecting algorithms that are appropriate to the reality of the data and research area are necessary. Research paper, experimentally evaluating 4 Spline, IDW, Kriging and Natural Neighbor algorithms to model terrain on two map pieces representing different terrain types, the research results show that: the adapting each algorithm depends heavily on the terrain of each interpolation area. Spline interpolated terrain surfaces in more detail in ravine and valley areas; Natural Neighbor excels at matching the contours of data in all regions. IDW and Kriging algorithms have similar results and have lower accuracy than the above two methods, especially it is difficult to interpolate peaks and slopes. MAE, respectively, of high and medium hills and mountains are Spline (9.7, 10.3), NN (11.8, 10.1), IDW (13.0, 10.9), Kriging (13.3, 12.2).

Keywords: Spatial interpolation, topographic modeling, DEM, DSM, accuracy.

* Corresponding author.

E-mail address: yenphanquochv@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.4445>



Nghiên cứu, thực nghiệm so sánh các phương pháp mô hình hóa địa hình

Phan Quốc Yên*, Nguyễn Thị Thu Nga, Tống Thị Hạnh

Học viện Kỹ thuật Quân Sự, 236 Hoàng Quốc Việt, Hà Nội, Việt Nam

Nhận ngày 12 tháng 9 năm 2019

Chỉnh sửa ngày 17 tháng 11 năm 2019; Chấp nhận đăng ngày 18 tháng 11 năm 2019

Tóm tắt: Địa hình bề mặt trái đất có thể được biểu diễn trong GIS bởi dữ liệu mô hình số độ cao. Mô hình hóa bề mặt là quá trình xác định bề mặt tự nhiên hoặc nhân tạo bằng cách sử dụng một hoặc nhiều phương trình toán học. Một thuật toán tổng quát để mô hình hóa bề mặt cho tất cả các ứng dụng là không có sẵn, mỗi phương pháp tạo ra bề mặt địa hình có một số ưu, nhược điểm và phụ thuộc vào hướng xử lý của nó. Như vậy, thực nghiệm, đánh giá và lựa chọn thuật toán phù hợp với thực tế của dữ liệu và khu vực nghiên cứu là cần thiết. Bài báo nghiên cứu, thực nghiệm đánh giá 4 thuật toán Spline, IDW, Kriging và Natural Neighbor để mô hình hóa địa hình trên hai mảnh bản đồ đại diện cho các dạng địa hình khác nhau, kết quả nghiên cứu cho thấy rằng: sự thích ứng từng thuật toán phụ thuộc nhiều vào dạng địa hình từng khu vực nội suy khác nhau. Spline nội suy được bề mặt địa hình chi tiết hơn ở các khu vực khe núi, thung lũng; Natural Neighbor vượt trội về khả năng khớp với đường đồng mức gốc của dữ liệu ở tất cả các khu vực. Thuật toán IDW và Kriging cho kết quả tương tự nhau và có độ chính xác thấp hơn hai phương pháp trên, đặc biệt rất khó để nội suy ra được các đỉnh và sườn núi. MAE lần lượt của đồi núi cao và đồi núi thấp trung bình là Spline (9.7, 10.3), NN (11.8, 10.1), IDW (13.0, 10.9), Kriging (13.3, 12.2).

Từ khóa: Nội suy không gian, mô hình hóa địa hình, DEM, DSM, độ chính xác.

1. Mở đầu

Mô hình hóa bề mặt là quá trình xác định bề mặt tự nhiên hoặc nhân tạo bằng cách sử dụng một hoặc nhiều phương trình toán học từ tập dữ liệu điểm mẫu [1-4]. Phương trình bề mặt 3 chiều trong không gian thường được biểu diễn mối liên hệ giữa độ cao với tọa độ x, y như hàm $z=f(x, y)$. Trong đó, hàm được lập dựa vào tập hợp điểm mẫu đại diện cho toàn bộ bề mặt, hàm này có thể

dự đoán các giá trị $z = f(x, y)$ cho các vị trí khác, gọi là hàm nội suy [5]. Có hai loại hàm nội suy, nội suy chính xác và nội suy gần đúng, chúng có thể được sử dụng cho nhiều mục đích và phụ thuộc vào từng loại công trình khác nhau. Trong thực tế, một phương thức chỉ chính xác khi biết trước biểu thức của hàm z . Một số phương pháp chính xác có thể được sử dụng với hệ số làm mịn và trong trường hợp này, nó có thể đi từ phương pháp chính xác đến phương pháp gần đúng [5].

*Tác giả liên hệ.

Địa chỉ email: yenphanquochv@gmail.com

<https://doi.org/10.25073/2588-1094/vnuees.4445>

Nội suy không gian là công cụ chủ yếu để ước tính dữ liệu không gian liên tục, các phương pháp nội suy đã được áp dụng các quy tắc khác nhau liên quan đến bề mặt trái đất. Các phương pháp nội suy được triển khai trong các ứng dụng lớn có thể được sử dụng để tạo các bề mặt như: mô hình số địa hình (DEM), mô hình số bề mặt (DSM), mô hình số độ sâu (DBM), bề mặt chuyển đổi địa lý (quasi), v.v. Chúng ta thường gọi chung là DEM, là một mảng đại diện các điểm hình vuông (Pixel) với một giá trị độ cao được liên kết với từng Pixel [6]. DEM được sử dụng từ các ứng dụng khoa học, thương mại, công nghiệp đến quân sự [3].

DEM có thể được nội suy từ các đường đồng mức, bản đồ địa hình, khảo sát thực địa, trạm đo vẽ ảnh số, Radar giao thoa, đo cao bằng laser, đo cao vệ tinh hoặc các kỹ thuật viễn thám quang học, v.v. [4, 6]. Các phương pháp nội suy không gian đã có sự phát triển nhanh chóng, với sự tăng dần số lượng phương pháp, có nhiều thuật toán nội suy đã được phát triển khác nhau. Các phương pháp nội suy thường được chia làm hai dạng nội suy toàn cục và nội suy cục bộ [4]. Các phương thức nội suy cục bộ ước tính giá trị của một điểm chưa biết dựa trên các giá trị của các điểm (pixel) lân cận, một số phương pháp nổi bật như: nghịch đảo khoảng cách (IDW), đa thức cục bộ, láng giềng tự nhiên (NN) và Hàm cơ sở xuyên tâm (RBF), v.v. Mặt khác, các phương pháp nội suy toàn cục như các hàm nội suy đa thức sử dụng tất cả các điểm mẫu có sẵn để tạo dự đoán cho một điểm cụ thể.

Các phương pháp nội suy khác nhau được áp dụng trên cùng một nguồn dữ liệu có thể có các kết quả khác nhau. Do đó, mối quan tâm chung là độ chính xác của các phương pháp. Có nhiều nghiên cứu đã công bố về so sánh độ chính xác của các phương pháp nội suy này sử dụng một dữ liệu nhất định. Thực nghiệm, đánh giá các thuật toán nội suy thường nghiên cứu về độ chính xác của các kỹ thuật nội suy liên quan đến kích thước mẫu dữ liệu, khoảng cách mẫu, loại địa hình và loại ứng dụng [1, 2, 7-12]. Kết luận của phần lớn các nghiên cứu đó là: thuật toán Kriging cho kết quả tốt hơn IDW trong bất kỳ cách lấy mẫu nào đối với toàn bộ các dạng địa hình [9,

13]. Tuy nhiên, trong nhiều nghiên cứu khác, tiếp cận các phương pháp IDW hoặc RBFs hoặc ANUDEM lại cho kết quả tốt hơn phương pháp Kriging [4, 8, 14-17].

Nhìn chung, các nghiên cứu cũng đã có những đánh giá độ chính xác của các kỹ thuật này liên quan đến các loại địa hình khác nhau. Tuy nhiên, ảnh hưởng của hình thái địa hình trong cảnh quan tự nhiên trên một phạm vi rộng lớn hiếm khi được nghiên cứu [13]. Hơn nữa, các nghiên cứu không có kết quả nhất quán về độ chính xác của các thuật toán nội suy không gian. Do đó, rất khó để chọn một phương pháp nội suy thích hợp cho một tập dữ liệu đầu vào nhất định. Vì vậy, mục tiêu chính của nghiên cứu này là đánh giá, so sánh độ chính xác của các phương pháp nội suy tạo DEM, ảnh hưởng của các kỹ thuật nội suy đến độ chính xác liên quan đến các loại địa hình. Bài báo sử dụng các dạng địa hình đồi núi cao, đồi núi thấp và trung bình để thực nghiệm, đánh giá bốn thuật toán nội suy không gian: thuật toán Spline, thuật toán láng giềng tự nhiên (NN - Natural Neighbor), nghịch đảo khoảng cách (IDW-Inverse Distance Weighting) và thuật toán Kriging. Các đỉnh đồi, đỉnh núi, khu vực chuyển tiếp, khu vực đồi thấp và thoải, khu vực sườn núi, thung lũng, v.v. được lựa chọn để kiểm tra sự sai khác và khả năng thích ứng của bình đồ nội suy được so với bình đồ gốc ban đầu.

2. Các phương pháp nội suy bề mặt địa hình

2.1. Thuật toán Spline

Phương pháp này nội suy độ cao của một điểm cụ thể bằng cách sử dụng hàm toán học để giảm thiểu độ cong bề mặt tổng thể, kết quả một bề mặt mịn, chính xác qua các điểm đầu vào [18]. Về mặt khái niệm, các điểm mẫu chứa giá trị độ cao, Spline uốn cong một mặt phẳng đi qua các điểm đầu vào, khớp một hàm toán học từ một số điểm nhất định của các điểm gần nhất trong mẫu, công thức nội suy như sau [11]:

$$S(x, y) = T(x, y) + \sum_{i=1}^n \omega_i R(r_i) \quad \text{hoặc}$$

$$Z(s_0) = \sum_{i=1}^n \omega_i R(\|s_i - s_0\|) + \omega_{n+1} \quad (1)$$

Trong đó: n là số điểm mẫu; ω_i là trọng số được tìm thấy khi giải hệ phương trình tuyến tính ($i=1$ đến $n+1$); $r_i = \|s_i - s_0\|$ là khoảng cách Euclide từ điểm chưa xác định $s_0(x, y)$ đến điểm thứ i (là s_i) trong tập mẫu; hàm $T(x, y)$ thường thiết lập bằng ω_{n+1} .

Các trọng số được xác định theo khoảng cách của các điểm đã biết, giải hệ n phương trình với n điểm đã biết thì sẽ cho một nghiệm duy nhất. Splines tạo ra kết quả tốt với các bề mặt thay đổi thoải và do đó thường không phù hợp khi có thay đổi lớn về giá trị bề mặt trong một khoảng cách ngang ngắn [12].

2.2. Thuật toán nghịch đảo khoảng cách

Phương pháp IDW xác định giá trị của các điểm chưa biết bằng cách tính trung bình trọng số khoảng cách các giá trị của các điểm đã biết trong vùng lân cận của mỗi pixel [19]. Những điểm càng cách xa điểm cần tính, càng ít ảnh hưởng đến giá trị tính toán, dẫn tới trọng số sẽ giảm. Công thức tính nội suy giá trị tại điểm chưa biết trên cơ sở các giá trị đã biết xung quanh nó như sau:

$$Z = \frac{\sum_{i=1}^n W_i Z_i}{\sum_{i=1}^n W_i} = \frac{\sum_{i=1}^n Z_i \times \frac{1}{d^k}}{\sum_{i=1}^n \frac{1}{d^k}} \quad (2)$$

Trong đó: i là các điểm dữ liệu đã biết giá trị; n là số điểm đã biết; Z_i là giá trị điểm thứ i ; d là khoảng cách đến điểm i ; k là hằng số IDW, k càng cao thì độ ảnh hưởng của các điểm ở xa càng thấp, thông thường $p = 2$.

Đặc trưng của bề mặt nội suy chịu ảnh hưởng của bán kính tìm kiếm, bán kính này giới hạn số lượng điểm mẫu được sử dụng để tính điểm nội suy. Có hai loại bán kính tìm kiếm: (1) Cố định (fixed) là bán kính tìm kiếm một khoảng cách xác định. (2) Biến đổi (variable) là bán kính tìm kiếm thay đổi để đáp ứng với số lượng các điểm mẫu theo yêu cầu.

Đầu ra phương pháp bị giới hạn phạm vi giá trị sử dụng nội suy, giá trị nội suy được không thể lớn hơn đầu vào cao nhất và nhỏ hơn đầu vào thấp nhất, vì thế nó không thể tạo ra các rặng và thung lũng nếu những điểm cực này không có

trong tập mẫu [19]. Ngoài ra, kết quả tốt nhất phương pháp khi tập mẫu liên quan đến các điểm đặc trưng địa hình có mật độ đủ dày. Nếu tập mẫu của điểm đầu vào là thưa thớt, không đều, kết quả có thể không đủ miêu tả bề mặt yêu cầu. Ngoài ra, phương pháp này không cung cấp đánh giá ngầm về chất lượng của các dự đoán [18].

2.3. Thuật toán Kriging thông thường

Theo miêu tả phương pháp Kriging thông thường, thực nghiệm biểu đồ bán phương sai được tính toán dựa vào các điểm dữ liệu nguồn và một mô hình được khớp với biểu đồ bán phương sai. Quy trình bắt đầu bằng cách tính khoảng cách giữa tất cả các điểm dữ liệu nguồn theo từng cặp. Sau đó, biểu đồ bán phương sai thu được bằng cách vẽ sự khác biệt về giá trị độ cao của các cặp điểm dữ liệu như là một hàm khoảng cách không gian tương ứng. Mô hình sử dụng trong nội suy thu được bằng cách khớp đường cong với biểu đồ bán phương sai thực nghiệm. Khớp đường cong thường được thực hiện bằng cách giảm thiểu tổng bình phương sai số giữa đường cong và các điểm biểu đồ bán phương sai. Nhiều loại mô hình đã được đề xuất, mô hình hàm mũ và mô hình Gaussian là một trong những mô hình phổ biến nhất [18]. Sau đó, đường cong đã khớp được so sánh với khoảng cách của các cặp điểm dữ liệu nguồn để tìm ra cách tính trọng số các điểm dữ liệu trong quy trình nội suy [20]. Phép nội suy được thực hiện theo phương trình sau bằng cách sử dụng các trọng số thu được như mô tả ở trên:

$$z_p = \sum_{i=1}^n \lambda_i z_i, \text{ với } \sum_{i=1}^n \lambda_i = 1 \quad (3)$$

z_p là điểm nội suy, z_i là điểm dữ liệu nguồn, và λ_i là trọng số, n là tổng số điểm dữ liệu nguồn. Nội suy Kriging cũng đưa ra ước tính lỗi nội suy. Phương sai của các điểm được nội suy có thể được tính dựa trên các trọng số và khoảng cách giữa các điểm được nội suy và các điểm dữ liệu nguồn. Phương sai sau đó có thể được sử dụng để ước tính khoảng tin cậy cho giá trị độ cao của điểm được nội suy, tức là, với xác suất nào độ cao của điểm nằm trong một phạm vi được xác định trước [20].

Một số ưu điểm của phương pháp này là giá trị của các điểm được gán không chỉ phụ thuộc vào khoảng cách mà còn phụ thuộc vào sự phân bố không gian các điểm. Điều này làm cho các giá trị nội suy mang tính tương quan không gian nhiều hơn. Phương pháp này là một quá trình gồm nhiều bước (phân tích thống kê khám phá, mô hình hóa biểu đồ phương sai, tạo bề mặt và khám phá bề mặt phương sai), vì vậy đòi hỏi nhiều thời gian tính toán, mô hình hóa, và cần nhiều dữ liệu đầu vào.

2.4. Thuật toán láng giềng tự nhiên

Phương pháp dựa trên lưới Voronoi của một tập các điểm không gian rời rạc, tạo ra một tam giác Delauney, chọn các nút gần nhất tạo thành một đường bao lồi xung quanh điểm nội suy, sau đó áp dụng các trọng số cho chúng dựa trên các khu vực tương ứng để nội suy một giá trị, hàm toán học được biểu diễn như sau [21]:

$$G(x) = \sum_{i=1}^n w_i(x) f(x_i) \quad (4)$$

Với $G(x)$ là ước tính tại x , w_i là các trọng số và $f(x_i)$ là dữ liệu đã biết tại x_i . Các thuộc tính cơ bản của nó là cục bộ, chỉ sử dụng một tập hợp các mẫu con bao quanh điểm truy vấn và độ cao nội suy được đảm bảo nằm trong phạm vi của các mẫu được sử dụng (lưới Voronoi). Thuật toán không tính được xu hướng và sẽ không tạo ra các đỉnh, hố, đường gờ hoặc thung lũng mà chưa có mẫu đại diện đầu vào. Bề mặt đi qua các mẫu đầu vào là tương đối mịn [18], các trọng số w_i được tính là $w_i(x) = \frac{A(x_i)}{A(x)}$, Trong đó $A(x)$ là diện tích của điểm Pixel mới ở giữa theo x và $A(x_i)$ là diện tích của giao điểm giữa ô mới có tâm ở x và ô cũ có tâm ở x_i [21].

Nội suy láng giềng tự nhiên có nhiều tính năng tích cực, thường hoạt động tốt với các điểm có cụm phân tán (thích hợp nhất khi các điểm dữ liệu mẫu được phân phối với mật độ không đồng đều), xử lý hiệu quả các bộ dữ liệu điểm đầu vào lớn. Đây là một kỹ thuật nội suy đa mục đích tốt và có ưu điểm là bạn không phải chỉ định các tham số như bán kính, số lượng lân cận hoặc trọng số.

3. Đánh giá độ chính xác của phương pháp

Các phương pháp đánh giá độ chính xác thực nghiệm nội suy độ cao gồm phương pháp trích điểm ngẫu nhiên, phương pháp kiểm tra khớp đường đồng mức, phương pháp mặt cắt và phương pháp phân tích trực quan địa hình.

3.1. Phương pháp trích điểm ngẫu nhiên

Theo phương pháp này, độ cao ước tính từ mô hình nội suy được so sánh tại mỗi điểm kiểm tra ngẫu nhiên từ tập điểm kiểm tra được trích xuất từ bản đồ gốc mà không tham gia vào tập điểm nội suy. Xác định các tham số: sai số tuyệt đối trung bình (MAE), sai số trung phương (RMSE) và hệ số tương quan giữa hai tập dữ liệu (R) [18].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{obs,i} - X_{model,i})^2}{n}} \quad (5)$$

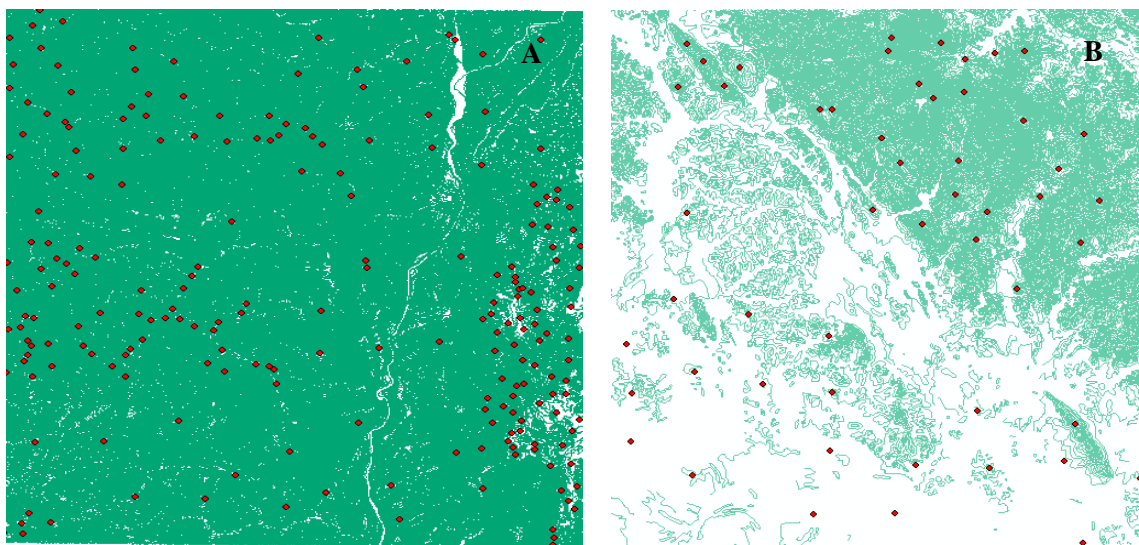
$$MAE = \frac{1}{n} \sum_{i=1}^n (|X_{obs,i} - X_{model,i}|) \quad (6)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_{model,i} - X_{obs,i})^2}{\sum_{i=1}^n (X_{model,i} - \bar{X})^2} \quad (7)$$

Trong đó: X_{obs} là giá trị dữ liệu độ cao kiểm tra, X_{model} là giá trị độ cao nội suy của vị trí thứ i , \bar{X} là giá trị độ cao trung bình của dữ liệu độ cao kiểm tra.

3.2. Phương pháp kiểm tra khớp đường đồng mức

Trong ứng dụng thực tế, để đánh giá độ chính xác tổng thể và độ tương quan giữa DEM và địa hình thực tế, chúng ta thường sử dụng phương pháp kiểm tra lại khả năng khớp đường đồng mức của DEM được nội suy với dữ liệu gốc nhằm kiểm tra lỗi về đường bình độ và độ cao toàn vùng. Phương pháp này bao gồm trích xuất lại đường bình độ gốc và trích xuất các đường đồng mức từ các DEM tính được bằng các thuật toán nội suy. Kiểm tra bằng thủ công hoặc phương pháp tự động, thống kê các lỗi không khớp về đường bình độ, đặc biệt các khu vực chuyển tiếp và khu vực đặc trưng địa hình như: đỉnh núi, thung lũng, rặng núi, khe suối, v.v.



Hình 1. Khu vực nghiên cứu và các điểm kiểm tra ngẫu nhiên.
(A) Khu vực địa hình đồi núi cao (F4839A), (B) núi trung bình và thấp (F4867A).

4. Thực nghiệm và đánh giá kết quả

4.1. Dữ liệu và khu vực thực nghiệm

Khu vực nghiên cứu gồm 2 mảnh bản đồ đại diện cho toàn bộ các dạng địa hình khác nhau. Dạng địa hình núi cao (F4839A), với độ cao cao nhất 2280m, thấp nhất 218m và trung bình 1087m, chênh cao trung bình từ 200-400m (hình 1 - tấm A). Dạng địa hình núi trung bình và đồi thấp (F4867A), với độ cao lớn nhất 1420m, thấp nhất là 5m và trung bình 114m, diện tích có một ít đồi núi ở mức độ cao trung bình, còn cơ bản là đồi thấp và đồng bằng. Hai mảnh bản đồ này thể hiện được đầy đủ sự phân bố đồng đều và đầy đủ dạng địa hình.

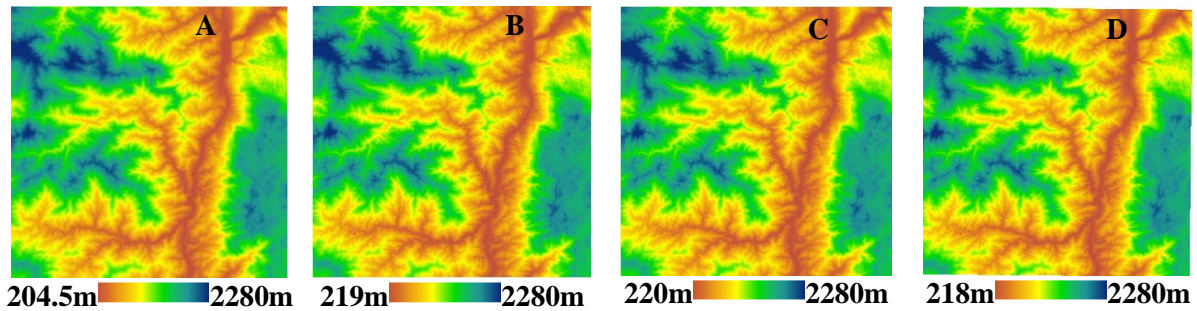
Do các DEM nội suy được tạo từ các bản đồ địa hình, độ chính xác của các DEM sẽ không cao hơn các bản đồ địa hình ban đầu. Bài báo lựa chọn ngẫu nhiên các điểm kiểm tra độ cao trong bản đồ địa hình làm bộ dữ liệu tham chiếu, được phân bố đồng đều trong khu vực nghiên cứu, và số lượng là: Khu vực núi cao 219 điểm, khu vực núi trung bình và thấp 52 điểm (hình 1).

4.2. Kết quả thực nghiệm

4.2.1. Thực nghiệm nội suy địa hình đồi núi cao

Với đầu vào là đường bình độ của bản đồ địa hình tỷ lệ 1:50.000, thực hiện các bước phân rã đường bình độ, tạo tập điểm làm đầu vào cho các thuật toán nội suy. Nội suy DEM từ tập điểm phân rã đó bằng các thuật toán đã trình bày ở trên, và trích xuất tập điểm kiểm tra lỗi ngẫu nhiên, thực hiện phân tích lỗi của các DEM tính được. DEM địa hình theo bốn phương pháp nội suy được thể hiện như Hình 2, kết quả cho thấy rằng, các phương pháp nội suy đều đạt được độ cao lớn nhất 2280m.

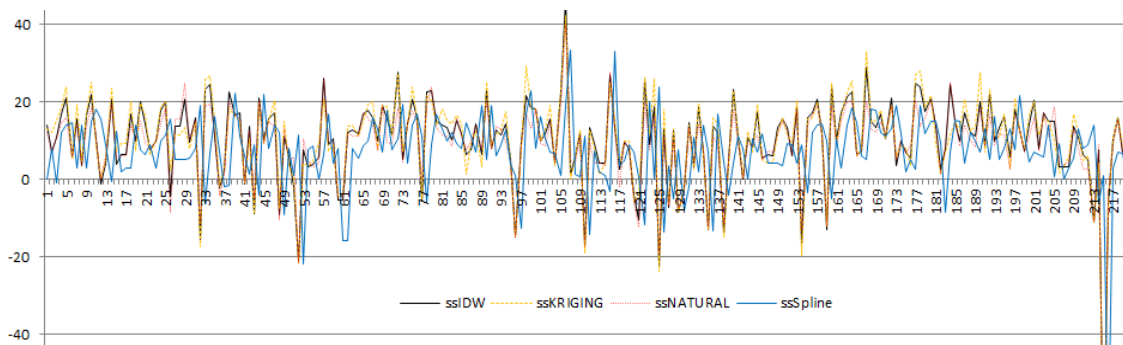
Sử dụng một số tham số về DEM như giá trị lớn nhất, nhỏ nhất, trung bình và độ lệch chuẩn (SD) để so sánh sự khác nhau của các DEM tính được cho thấy: Phương pháp Spline tính toán được với phạm vi lớn nhất (204.5-2280m), tiếp theo là phương pháp NN (218-2280m) các phương pháp khác. Độ lệch chuẩn SD và giá trị trung bình độ cao tăng dần theo thứ tự các phương pháp NN, Kriging, IDW và Spline. Tuy nhiên, sai số trung bình giữa độ cao các phương pháp chênh lệch không nhiều, vì vậy, sai số chuẩn vẫn có thể đại diện cho mức độ phân tán dữ liệu của các phương pháp và có thể so sánh được với nhau. Mức độ phân tán dữ liệu tính được theo phương pháp Spline cao nhất, phản ánh mức độ chi tiết địa hình tính được của phương pháp này (bảng 1).



Hình 2. DEM nội suy theo các thuật toán (A) Spline, (B) IDW, (C) Kriging, (D) NN.

Bảng 1. Một số tham số so sánh của 4 phương pháp nội suy trong khu vực núi cao

Thuật toán	Lỗi kiểm tra từ tập điểm ngẫu nhiên			Một số tham số thống kê kết quả			
	RMSE (m)	MAE (m)	R ²	min	Max	mean	SD
Spline	12.0073	9.7086	0.9991	204.5	2280	1091.3	459.4
IDW	15.2466	13.0156	0.9986	227.8	2280	1085.4	456.4
Kriging	15.7613	13.2968	0.9985	220	2280	1083.5	457.2
NN	13.9728	11.8408	0.9988	226.3	2280	1082.4	455.4

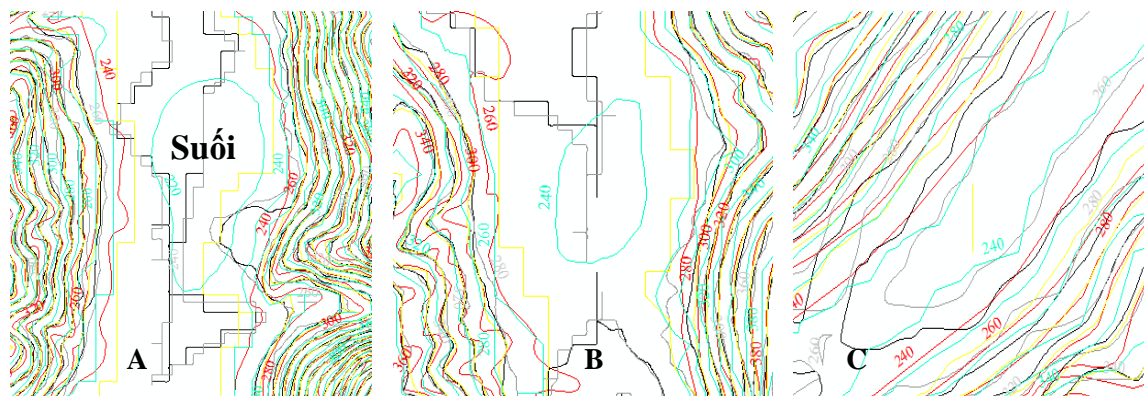


Hình 3. Đường cong lỗi của bốn phương pháp nội suy địa hình đồi núi cao.

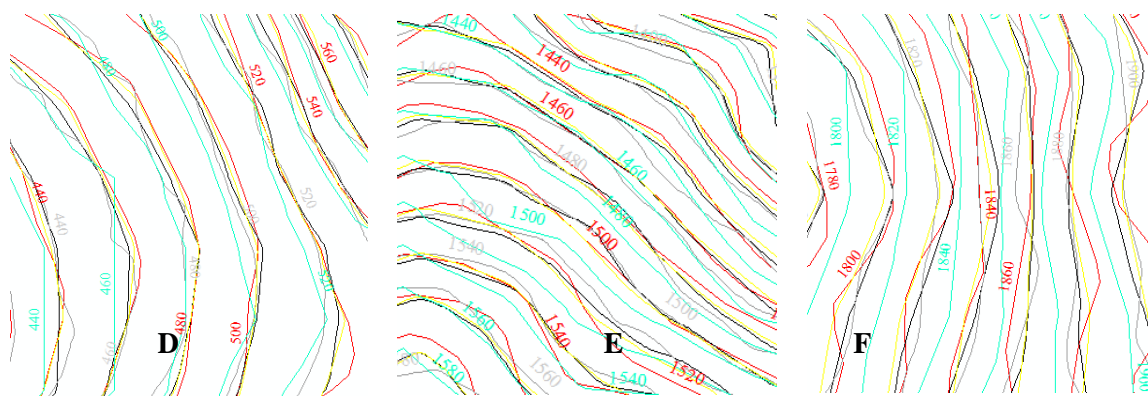
Sử dụng phương pháp kiểm tra sai số điểm ngẫu nhiên của 219 điểm, kết quả đạt được như trong bảng 1 ở trên, các giá trị RMSE của tất cả các phương pháp đều nhỏ hơn 16m. Nhìn chung, kiểm tra 3 tham số RMSE, MAE và R², thuật toán Spline có độ chính xác tốt nhất trong 3 phương pháp đối với dạng địa hình đồi núi cao (RMSE=12.0073m, R²=0.9991). Phương pháp NN cũng cho kết quả tương đối tốt, cao hơn so với hai phương pháp IDW và Kriging. Đường cong lỗi qua các điểm kiểm tra ngẫu nhiên của các DEM nội suy được hiển thị trong Hình 3, giá trị lỗi biên độ sai số của thuật toán Spline là nhỏ nhất (đường màu xanh, thấp nhất) so với các

phương pháp nội suy khác. Nhìn chung, độ chính xác được thể hiện trên đường cong giảm dần từ phương pháp Spline, NN, phương pháp IDW và Kriging cho kết quả kém nhất ở dạng địa hình này. Các phương pháp tạo ra biến thể lỗi có biên độ khác nhau, tuy nhiên hướng biến động lỗi cơ bản tương tự nhau.

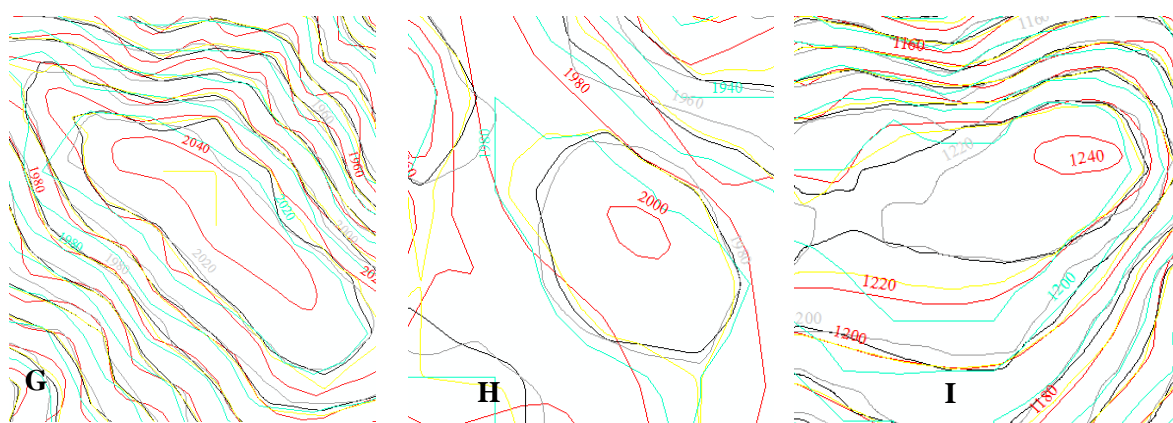
Sử dụng phương pháp chạy lại đường bình độ từ DEM tính toán được của bốn thuật toán, chồng các lớp đường bình độ này lên lớp đường bình độ gốc, tiến hành kiểm tra cho từng khu vực mảnh bản đồ F4839A, một số kết quả được thống kê như Bảng 2 và Hình vẽ 4.



Hình 4. So sánh sự khác nhau về đường đồng mức trong khu vực khe rãnh và suối (đỏ: đường bình độ gốc, vàng: NN, đen: IDW, xám: Kriging và màu xanh thuật toán Spline).



Hình 5. So sánh sự khác nhau về đường đồng mức phạm vi có độ dốc đều (đỏ: đường bình độ gốc, vàng: NN, đen: IDW, xám: Kriging và màu xanh thuật toán Spline).



Hình 6. So sánh sự khác nhau về đường đồng mức phạm vi đỉnh núi (đỏ: đường bình độ gốc, vàng: NN, đen: IDW, xám: Kriging và màu xanh thuật toán Spline).

Phân tích kết quả cho thấy rằng, đường đồng mức từ thuật toán NN (màu vàng) khớp nhất với đường đồng mức gốc, hai lớp đường đồng mức (màu đen và xám) cho kết quả tương tự nhau và độ khớp so với bình độ gốc kém hơn. Qua kiểm tra toàn bộ mảnh bản đồ, thuật toán Natural có 8 đỉnh và sườn bị lỗi, không chạy đủ đường bình độ cuối của đỉnh núi, thuật toán cho kết quả nội suy thấp hơn giá trị gốc.

Hình 4 là kiểm tra, so sánh đại diện cho khu vực khe núi và thung lũng với độ cao từ 240 đến 650m được trích xuất mẫu trên khu vực mảnh bản đồ thử nghiệm. Tấm A - Hình 4 là khu vực trích xuất từ dòng suối có độ cao thấp nhất, kết quả thấy rằng: độ cao khu vực suối đối với bản đồ gốc là không thể hiện trên bản đồ (đường bình độ màu đỏ hai bên mép suối với độ cao 240m). Trong khi đó, độ cao tính được của bốn phương pháp có sự khác biệt nhau, (1) phương pháp NN (đường màu vàng) cho kết quả tương đối giống với đường bình độ gốc, tuy nhiên đường bình độ bị dịch chuyển về phía suối. (2) phương pháp Kriging và IDW cho kết quả tương tự nhau, đường bình độ 240m bị di chuyển ra gần giữa khu vực suối. (3) Phương pháp Spline cho kết quả khớp nhất với đường bình độ gốc đối với độ cao 240m, ngoài ra, phương pháp này còn nội suy được độ cao ở khu vực lòng suối (đường bình độ 240m). Kiểm tra một số vị trí khác cũng cho kết quả tương tự (tấm B, C – Hình 4).

Thực hiện tương tự đối với khu vực núi có độ dốc đều (hình 5), kết quả kiểm tra cho thấy,

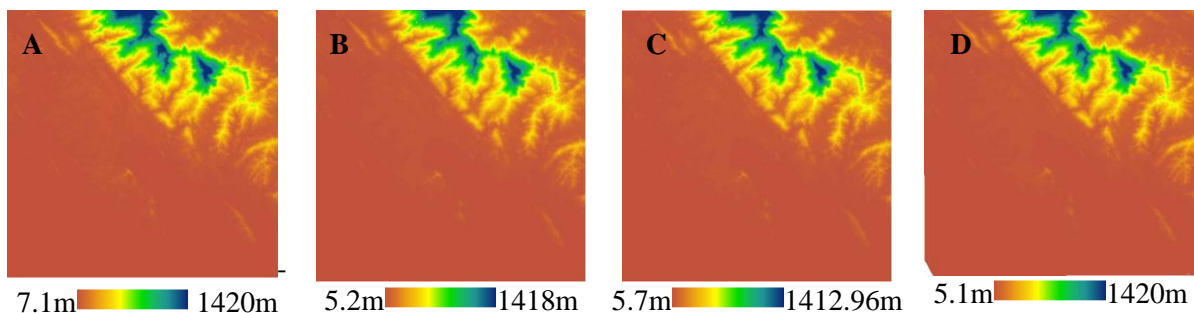
đường màu vàng (phương pháp NN) bám sát với đường bình độ gốc nhất, hai phương pháp IDW và Kriging cũng cho kết quả tương tự. Trong khu vực này, đường bình độ của phương pháp Spline cho kết quả xa nhất so với đường bình độ gốc, độ cao nội suy ở vị trí thấp hơn so với đường bình độ gốc một khoảng cách khoảng 5-10m.

Kiểm tra đối với các khu vực đỉnh núi (hình 6), cơ bản cả bốn thuật toán thường không tính đến độ cao của đỉnh núi, độ cao đỉnh núi thường bị hạ thấp hơn so với bình độ gốc, thể hiện đường bình độ của đỉnh núi thường bị thiếu.

Kiểm tra trên toàn bộ tờ bản đồ cho thấy rằng, thuật toán NN và Spline có tính được một số đỉnh núi, tuy nhiên, đường bình độ thể hiện thường không đúng với đường bình độ gốc. Thuật toán Kriging và IDW thường không tính được đỉnh núi, đường bình độ cuối cùng của đỉnh núi thường bị bỏ sót.

4.2.2. Thực nghiệm nội suy địa hình đồi núi trung bình và thấp: Thực hiện tính toán tương tự như với địa hình núi cao, địa hình núi thấp và trung bình được thực hiện ở khu vực mảnh bản đồ F4867A, kết quả DEM tính được như sau:

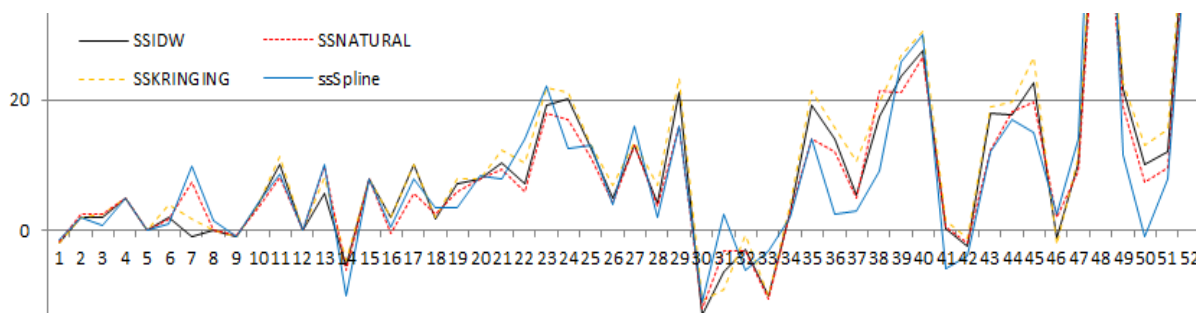
DEM tính được cũng cho thấy, phương pháp NN cho một kết quả mạnh mẽ hơn về phạm vi độ cao nội suy được. Kiểm tra tương tự như các bước đã thực hiện ở trên, địa hình núi thấp được đánh giá dựa vào 52 điểm kiểm tra ngẫu nhiên, được thể hiện qua bảng lỗi (bảng 2) và đường cong lỗi (hình 8) như sau:



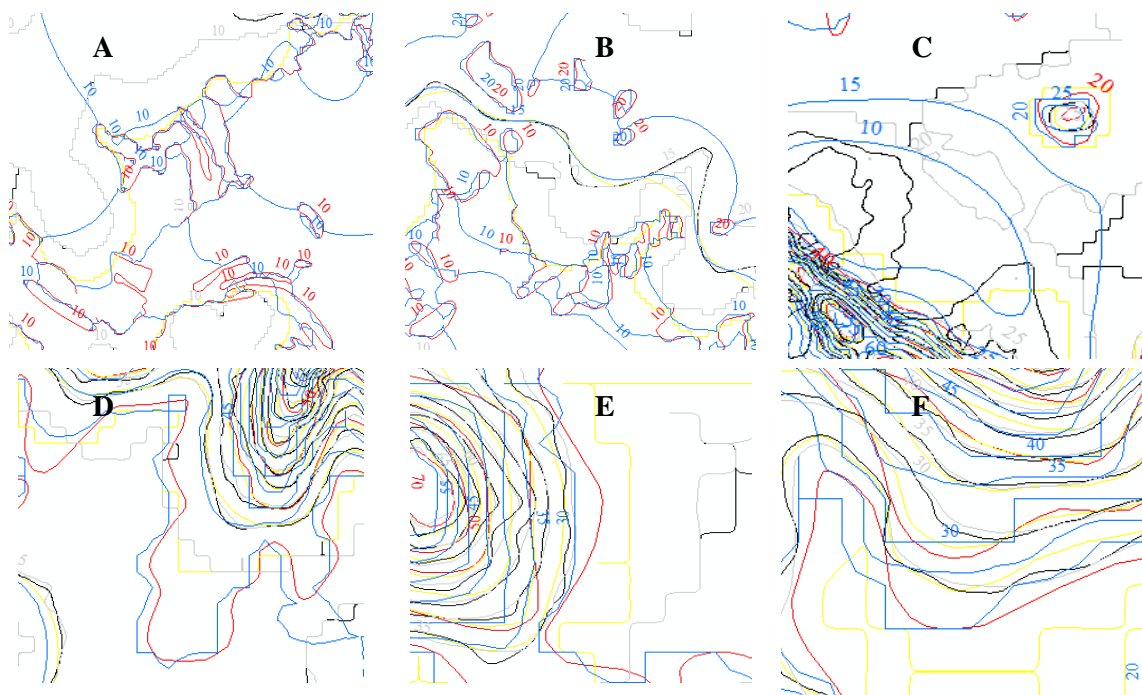
Hình 7. DEM nội suy theo các thuật toán (A) Spline, (B) IDW, (C) Kriging, (D) NN.

Bảng 2. Tham số so sánh của 4 phương pháp nội suy địa hình núi trung bình và thấp

Thuật toán	Lỗi kiểm tra từ tập điểm ngẫu nhiên			Một số tham số thống kê kết quả			
	RMSE (m)	MAE (m)	R ²	min	Max	mean	SD
Spline	17.6326	10.3046	0.9975	-7.1	1420	111	205.2
IDW	16.4754	10.9374	0.9978	10.5	1407.2	148.1	227.3
Kriging	17.9949	12.1891	0.9974	5.7	1391	111.2	205.3
NN	15.2922	10.0674	0.9981	10.6	1420	144.7	225.6



Hình 8. Đường cong lỗi của bốn phương pháp nội suy địa hình núi thấp và trung bình.



Hình 9. Khác nhau về đường đồng mức phạm vi đồi nhỏ (A, B, C) và khu vực chuyển tiếp (D, E, F) (đỏ: đường bình độ gốc, vàng: NN, đen: IDW, xám: Kriging và màu xanh thuật toán Spline).

Các tham số trong bảng 2 cho thấy rằng, các phương pháp đạt được tương tự như ở khu vực địa hình núi cao về phạm vi độ cao nội suy. Tuy nhiên, có sự khác biệt lớn về độ chính xác, độ phân tán dữ liệu. Đối với khu vực này, độ chính xác thấp hơn so với khu vực núi cao, RMSE giảm trung bình 1m so với khu vực núi cao ở tất cả các thuật toán. Kết quả cũng chỉ ra rằng, giữa số liệu ước tính được và số liệu kiểm tra có mối quan hệ tuyến tính cao với $R^2 \approx 1$. Đường cong lỗi biến thiên phức tạp, kết hợp đường cong lỗi và các tham số lỗi kiểm tra từ tập điểm ngẫu nhiên cho thấy rằng cũng thể hiện phương pháp NN cho kết quả trung bình lỗi thấp nhất (RMSE = 15.2922, $R^2 = 0.9981$) so với các phương pháp khác. Phương pháp Kriging và IDW cho kết quả thấp hơn và tương tự nhau. Phương pháp Spline cho kết quả trung bình với sai số RMSE thấp, tuy nhiên quan sát đường cong lỗi thì sai số lỗi của phương pháp Spline khu vực này không đồng đều (lúc cao nhất, có lúc lại thấp nhất).

Kiểm tra đường bình độ của các thuật toán so với đường bình độ gốc trên toàn bộ mảnh bản đồ ta thấy rằng: Các khu vực đồi núi thấp, có địa hình khá bằng phẳng (hình 9) cho thấy rằng, địa hình nội suy bằng thuật toán Spline (đường màu xanh) cho kết quả khớp với bình độ gốc và thể hiện địa hình chi tiết hơn các phương pháp khác. Ngoài ra, tấm C-Hình 9 cũng chỉ ra rằng, thuật toán NN không nội suy ra khu vực bằng phẳng giữa hai đỉnh đồi như bình độ gốc. Ngược lại, ba thuật toán còn lại đều nội suy và thể hiện đường bình độ ở khu vực này một cách chi tiết. Điều này phản ánh đúng bản chất của từng thuật toán nội suy cụ thể.

Các khu vực núi cao, thung lũng và đỉnh núi qua kiểm tra cũng cho kết quả tương tự như khu vực đồi núi cao đã thực hiện ở trên. (hình 9 - tấm D, E, F).

5. Kết luận

Kỹ thuật nội suy đóng một vai trò quan trọng trong việc đạt được độ chính xác cao của DEM. Ảnh hưởng của kỹ thuật nội suy đến độ chính xác của DEM phụ thuộc vào các loại địa hình, phân

phối các điểm mẫu và một số yếu tố khác. Kết quả nghiên cứu cho thấy như sau:

Thuật toán Spline nội suy được bề mặt địa hình chi tiết hơn ở các khu vực khe núi, thung lũng. Tạo ra được địa hình ở các khu vực sâu hơn so với vị trí xung quanh như khe rãnh, thung lũng, v.v. Phương pháp cũng được thực hiện với tốc độ nhanh, đảm bảo xử lý được với tập mẫu dữ liệu lớn. Tuy nhiên, phương pháp vẫn còn một số hạn chế ở các khu vực sườn dốc đều, thường cho kết quả với độ cao thấp hơn bình độ gốc.

Thuật toán NN vượt trội so với các kỹ thuật khác là khả năng khớp với đường đồng mức gốc của dữ liệu ở tất cả các khu vực, tốc độ xử lý nhanh, thuật toán phù hợp với các tập dữ liệu lớn. Đây là một phương pháp xác định cục bộ với độ cao nội suy được đảm bảo nằm trong phạm vi của các mẫu được sử dụng (nội suy, không ngoại suy). Tuy nhiên, thuật toán không tạo ra các đỉnh, hố, đường gờ hoặc thung lũng chưa có trong các mẫu đầu vào và điều chỉnh cục bộ theo cấu trúc của dữ liệu đầu vào.

Thuật toán IDW và Kriging cho kết quả tương tự nhau và có độ chính xác thấp hơn hai phương pháp trên, đặc biệt rất khó để nội suy ra được các đỉnh và sườn núi. Đối với một phương pháp nội suy cụ thể, địa hình nhấp nhô càng lớn, độ chính xác của DEM được tạo ra càng thấp. Thuật toán IDW có tốc độ xử lý nhanh, tuy nhiên, phương pháp Kriging, tốc độ tính toán bị ảnh hưởng bởi số điểm trong tập dữ liệu và kích thước của bề mặt nội suy. Vì vậy, thời gian thực hiện nội suy bề mặt đối với phương pháp này lâu hơn rất nhiều so với ba phương pháp còn lại.

Tài liệu tham khảo

- [1] F.J. Aguilar, et al., Effects of terrain morphology, sampling density, and interpolation methods on grid DEM Accuracy, Photogrammetric Engineering & Remote Sensing 71 (2005) 805-816.
- [2] P.A. Longley, et al., Geographic Information Systems and Science, John Wiley & Sons 3rd Edition (2010).
- [3] Q. Weng, An evaluation of spatial interpolation accuracy of elevation data, in Progress in Spatial Data Handling, Springer-Verlag, Berlin (2006) 805-824.

- [4] Pattathal Vijayakumar Arun, A comparative analysis of different DEM interpolation methods, *The Egyptian Journal of Remote Sensing and Space Science* 16.2 (2013) 133-139. <https://doi.org/10.1016/j.ejrs.2013.09.001>.
- [5] Paul Daniel Dumitru, Marin Ploeanu, Dragos Badea, Comparative study regarding the methods of interpolation, *Recent advances in geodesy and Geomatics engineering* 1 (2013) 45.
- [6] Manuel Peralvo, David Maidment, Influence of DEM interpolation methods in drainage analysis, *Gis Hydro* 4 (2004) 4-7.
- [7] A. Carrara, G. Bitelli, R. Carla, Comparison of techniques for generating digital terrain models from contour lines, *International Journal of Geographical Information Science* 11 (1997) 451-473.
- [8] J.C. Guarneri, R.C. Weih Jr, Comparing Methods for Interpolation to Improve Raster Digital Elevation Models, *Journal of the Arkansas Academy of Science* 66 (2012) 77-81. <https://scholarworks.uark.edu/jaas/vol66/iss1/16>.
- [9] G.L. Heritage, D.J. Milan, Influence of survey strategy and interpolation model on DEM quality, *Geomorphology*. 112.3 (2009) 334-344. [10.1016/j.geomorph.2009.06.024](https://doi.org/10.1016/j.geomorph.2009.06.024).
- [10] Dennis Weber, Evan Englund, Evaluation and comparison of spatial interpolators II, *Mathematical Geology* 26 (1994) 589-603.
- [11] Besim Ajvazi, Kornél Czimer, A comparative analysis of different DEM interpolation methods in GIS: case study of Rahovec, Kosovo, *Geodesy and Cartography* 45.1 (2019) 43-48. <https://doi.org/10.3846/gac.2019.7921>.
- [12] T.P. Robinson, G. Metternicht, Testing the performance of spatial interpolation techniques for mapping soil properties, *Computers and electronics in agriculture* 50.2 (2006) 97-108. [doi:10.1016/j.compag.2005.07.003](https://doi.org/10.1016/j.compag.2005.07.003).
- [13] D. Zimmerman, et al., An experimental comparison of ordinary and universal kriging and inverse distance weighting, *Mathematical Geology* 31 (1999) 375-390.
- [14] Dennis Weber, Evan Englund, Evaluation and comparison of spatial interpolators, *Mathematical Geology* 24.4 (1992) 381-391.
- [15] J Gallichand, D Marcotte, Mapping clay content for subsurface drainage in the Nile Delta, *Geoderma* 58.3-4 (1993) 165-179. [https://doi.org/10.1016/0016-7061\(93\)90040-R](https://doi.org/10.1016/0016-7061(93)90040-R).
- [16] D.J. Brus, et al., The performance of spatial interpolation methods and choropleth maps to estimate properties at points: a soil survey case study, *Environmetrics* 7.1 (1996) 1-16.
- [17] J. Fernando Aguilar, et al., Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy, *Photogrammetric Engineering & Remote Sensing* 71.7 (2005) 805-816.
- [18] Qulin Tan, Xiao Xu, Comparative analysis of spatial interpolation methods: an experimental study, *Sensors & Transducers* 165.2 (2014) 155.
- [19] David F Watson, A refinement of inverse distance weighted interpolation, *Geoprocessing* 2 (1985) 315-327.
- [20] J. Pohjola, J. Turunen, T. Lipping, Creating High-resolution Digital Elevation Model Using Thin Plate Spline Interpolation and Monte Carlo Simulation, Working Report (2009).
- [21] R. Sibson, A Brief Description of Nearest Neighbor Interpolation, *Interpolating Multivariate Data*, John Wiley & Sons, New York (1981) 21-36.