BỘ GIÁO DỤC VÀ ĐÀO TẠO TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

NGUYỄN THỊ THU HƯỜNG

NGHIÊN CỨU NÂNG CAO ĐỘ CHÍNH XÁC CỦA MÔ HÌNH SỐ ĐỘ CAO DẠNG GRID BẰNG MẠNG NEURON HOPFIELD

LUẬN ÁN TIẾN SĨ KỸ THUẬT

HÀ NỘI - 2021

BỘ GIÁO DỤC VÀ ĐÀO TẠO TRƯỜNG ĐẠI HỌC MỎ - ĐỊA CHẤT

NGUYỄN THỊ THU HƯỜNG

NGHIÊN CỨU NÂNG CAO ĐỘ CHÍNH XÁC CỦA MÔ HÌNH SỐ ĐỘ CAO DẠNG GRID BẰNG MẠNG NEURON HOPFIELD

Ngành : Kỹ thuật Trắc địa - Bản đồ Mã số : 9 52 05 03

LUẬN ÁN TIẾN SĨ KỸ THUẬT

NGƯỜI HƯỚNG DẪN KHOA HỌC

PGS.TS NGUYÊN QUANG MINH

HÀ NỘI - 2021

Lời cam đoan

Tôi xin cam đoan đây là công trình nghiên cứu của riêng tôi. Các số liệu được sử dụng và kết quả phân tích, trình bày trong luận án là trung thực và chưa từng được công bố trong bất cứ công trình nào.

Tác giả luận án

Nguyễn Thị Thu Hương

MỤC LỤC

Lời cam đo	oani
MỤC LỤC	Zii
DANH MU	ỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT TIẾNG VIỆT VÀ TIẾNG ANHv
DANH MU	ỤC CÁC BẢNG, BIỂU ĐÔvi
DANH MU	ỤC CÁC HÌNH VĨ viii
MỞ ĐẦU.	
Chương 1	TỔNG QUAN CÁC VẤN ĐỀ NGHIÊN CỨU VỀ MÔ HÌNH SỐ ĐỘ
CAO, MA	ANG NEURON VÀ ỨNG DỤNG MẠNG NEURON HOPFIELD
TRONG C	CÁC BÀI TOÁN TỐI ƯU HÓA8
1.1	Tổng quan về mô hình số độ cao8
1.1.1 (Các khái niệm và định nghĩa về mô hình số độ cao8
1.1.2 0	Các cấu trúc của mô hình số độ cao (DEM)10
1.1.3 (Các phương pháp thành lập mô hình số độ cao (DEM)18
1.1.4 H	Độ chính xác bề mặt mô hình số địa hình (DEM)20
1.1.5 (Các ứng dụng của mô hình số độ cao24
1.1.6 (Công tác thành lập DEM ở trong và ngoài nước26
1.1.7 N	Một số nghiên cứu về cải thiện và đánh giá độ chính xác DEM30
1.2	Tổng quan về mạng neuron37
1.2.1	Cấu tạo của một neuron sinh học38
1.2.2 N	Nguyên lý hoạt động của các neuron38
1.2.3 I	Khái niệm và cấu trúc của mạng neuron nhân tạo
1.2.4 I	Phân loại mạng neuron41
1.2.5 H	Đặc điểm của mạng neuron nhân tạo43
1.2.6 U	Úng dụng của mạng neuron nhân tạo44
1.2.7 N	Mạng neuron Hopfield45
1.2.8 U	Úng dụng mạng neuron Hopfield trong các bài toán tối ưu hóa46
1.3 I	Luận giải về tăng độ phân giải không gian grid DEM47

1.4	Một số nghiên cứu tiêu biểu về tăng độ phân giải không gian và tăng độ
chính xác	20 DEM
1.5	Kết luân chương 1
Chương 2	2 KHẢO SÁT VỀ KHẢ NĂNG TĂNG ĐÔ CHÍNH XÁC CỦA GRID
DEM BĂ	NG CÁC THUÂT TOÁN TÁI CHIA MẫU52
2.1	Các phương pháp đánh giá đô chính xác của grid DEM52
2.1.1	Phương pháp đánh giá trực quan
2.1.2	Phương pháp đánh giá đinh lương
2.2	Môt số thuật toán tái chia mẫu (Resampling) phổ biến nhằm tăng đô phân
giải khôn	g gian cho DEM dang grid
2.2.1	Phương pháp tái chia mẫu Bilinear (song tuyến)
2.2.2	Phương pháp nội suy dựa vào điểm lận cân gần nhất (Nearest Neighbor)61
2.2.3	Phương pháp tái chia mẫu Bi-cubic
2.2.4	Phương pháp nôi suy Kriging 63
2.3	Thực nghiêm tăng đô phân giải không gian của mô hình số đô cao DEM
dang grid	bằng các thuật toán tái chia mẫu phổ biến
2.3.1	Dĩr liêu và thực nghiêm 65
2.3.2	Phân tích về đô chính xác 75
2.4	Kết luận chương 2 97
Chirong ?	NGHIÊN CI'TU NÂNG CAO ĐÔ CHÍNH XÁC CỦA MÔ HÌNH SỐ
	DANG GRID BẰNG PHƯƠNG PHÁP SỬ DUNG MẠNG NEURON
HODELEI	
3 1	Ca sả khoa học của việc ứng dụng mạng neuron Honfield để tặng độ nhận
J.I giải khôn	có số khôa học của việc ủng dụng mặng neuron rippicit de tang độ phản a gian và đô chính vác của mô hình số đô cao dạng grid
	Mang nouron Honfield ứng dụng cho giêu nhận giải bản đề (super
J.2	mang neuron riopried ung dung cho sieu phan giai ban do (super-
2 2 1	Vây dựng mô bình
3.2.1	Thiết lập các hàm mục tiêu và điều kiên
3.2.2	Thiết lập các hàm mục tiêu và điều kiện102

3.3	Xây dựng thuật toán nâng cao độ chính xác của mô hình số độ cao dạng
grid	
3.3.1	Xây dựng mô hình mạng neuron Hopfield nhằm tăng độ phân giải không
gian củ	a mô hình số độ cao DEM dạng grid105
3.3.2	Sơ đồ khối của thuật toán109
3.3.3	Thiết kế chương trình tăng độ phân giải không gian và nâng cao độ chính
xác của	a grid DEM sử dụng mạng neuron Hopfiled110
3.4	Thực nghiệm tăng độ phân giải không gian và độ chính xác của mô hình
số độ ca	o dạng grid bằng phương pháp sử dụng mạng neuron Hopfield (mô hình
HNN đã	biến đổi)114
3.4.1	Dữ liệu thực nghiệm114
3.4.2	Kết quả thực nghiệm, phân tích độ chính xác119
3.4.3	Đánh giá trực quan122
3.4.4	Đánh giá định lượng132
3.5	So sánh độ chính xác về độ cao giữa các DEM sau khi tăng độ phân giải
bằng thu	ật toán mạng neuron Hopfiled và các phương pháp tái chia mẫu với các
điểm độ	cao kiểm tra được đo bằng máy toàn đạc điện tử136
3.6	Kết luận chương 3138
KÊT LU	ÂN VÀ KIẾN NGHỊ141
DANH N	MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA TÁC GIẢ143
TÀI LIỆ	U THAM KHẢO145
PHỤ LỤ	C 1152
PHỤ LỤ	C 2158
PHỤ LỤ	TC 3

DANH MỤC CÁC KÝ HIỆU, CHỮ VIẾT TẮT TIẾNG VIỆT VÀ TIẾNG ANH

DEM	Digital Elevation Model:	Mô hình số độ cao
DSM	Digital Surface Model:	Mô hình số bề mặt
DTM	Digital Terrain Model:	Mô hình số địa hình
HNN	Hopfield Neuron Network:	Mang neuron Hopfiled
ME	Mean Error:	Sai số trung bình
LiDAR	Light Detection And Ranging:	Công nghệ phát hiện và đo khoảng
cách bằng ch	ùm tia laser	
RMSE	Root Mean Square Error:	Sai số trung phương
SRTM	Shuttle Radar Topography Miss	ion: Công nghệ radar sử dụng cảm
biến đặt trên	tàu con thoi để tạo DEM	
UAV	Unmanned Aircraft Vehicle:	Thiết bị máy bay không người lái
2D	Two Dimension:	Hai chiều
3D	Three Dimension:	Ba chiều

DANH MỤC CÁC BẢNG, BIẾU ĐỒ

Bảng 1-1. Bảng phân cấp độ chính xác dữ liệu độ cao số theo Hiệp hội Đo ảnh và
Viễn thám Mỹ [23]23
Bảng 1-2. So sánh độ chính xác trên dữ liệu mô hình số độ cao và độ chính xác tính
theo khoảng cao đều đường đồng mức tương ứng [23]23
Bảng 1-3. Một số các thông số được tính toán từ DEM và các ứng dụng của chúng
Bảng 1-4. Một số loại DEM ở nước ta hiện nay
Bảng 2-1. Đánh giá độ chính xác của DEM dựa trên tiêu chuẩn ASPRS cho dữ liệu
không gian địa lý kỹ thuật số72
Bảng 2-2. Sai số trung phương (RMSE) của các phương pháp tái chia mẫu sử dụng
mô hình song tuyến (Bilinear), Bi-cubic và Kriging cho DEM 20m khu vực Nghệ
An
Bảng 2-3. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng
mô hình song tuyến (Bilinear), tái lấy mẫu Bi-cubic và theo thuật toán nội suy
Kriging cho DEM SRTM 30m khu vực Nghệ An
Bảng 2-4. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng
mô hình song tuyến (Bilinear), Bi-cubic và thuật toán Kriging cho DEM 5m khu
vực Lạng Sơn90
Bảng 2-5. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng
phương pháp tái chia mẫu song tuyến (Bilinear), Bi-cubic và thuật toán Kriging cho
DEM 30m khu vực Đắc Hà91
Bảng 2-6. Các hệ số hồi quy tuyến tính cho các bộ dữ liệu DEM sau khi tái chia
mẫu độ phân giải 20m khu vực Nghệ An và DEM độ phân giải 30m khu vực Nghệ
An và các bộ dữ liệu DEM lấy mẫu độ phân giải 5m khu vực Lang Sơn và DEM độ
phân giải 30m khu vực Đắc Hà95
Bảng 3-1. Sai số trung phương của các phương pháp tái chia mẫu song tuyến, Bi-
cubic, nội suy Kriging và phương pháp dùng mô hình mạng neuron Hopfiled HNN
Bảng 3-2. Các hệ số hồi quy tuyến tính cho cả bốn bộ dữ liệu D1, D2, S1, S2121

Bảng 3-3. Bảng thống kê các chênh lệch độ cao, các sai số giữa các DEM tăng	độ
phân giải và các điểm đo bằng toàn đạc điện tử1	37
Bảng PL2-1. Bảng so sánh chênh cao giữa DEM sau khi tăng độ phân giải bằ	ng
thuật toán mạng neuron Hopfiled (DEM HNN), DEM đầu vào (DEM Input) và c	các
DEM sau khi tái chia mẫu song tuyến (DEM Bilinear), Bi-cubic (DEM Bi-cubi	c),
Kriging (DEM Kriging) – đối với bộ dữ liệu DEM 5m Lạng Sơn với 236 điểm	độ
cao được đo bằng máy toàn đạc điện tử ở cùng khu vực1	58

DANH MỤC CÁC HÌNH VĨ

Hình 2-8. Quá trình thực nghiệm sử dụng dữ liệu DEM giảm độ phân giải66
Hình 2-9. Các bước thực nghiệm với dữ liệu grid DEM thực67
Hình 2-10. Tăng độ phân giải không gian của DEM khu vực Yên Thành, Nghệ An
từ độ phân giải 60m đến 20m (bộ dữ liệu D1)69
Hình 2-11. Tăng độ phân giải không gian của DEM SRTM khu vực Yên Thành,
Nghệ An) từ độ phân giải 90m đến 30m (bộ dữ liệu D2)71
Hình 2-12. Dữ liệu DEM khu vực Mai Pha, Lạng Sơn sau khi tăng độ phân giải
không gian từ 20m lên 5m (bộ dữ liệu S1)73
Hình 2-13. Dữ liệu DEM khu vực Đắc Hà, Kon Tum sau khi tăng độ phân giải
không gian từ 90m lên 30m (Bộ dữ liệu S2)74
Hình 2-14. Sơ đồ vị trí các mặt cắt của 4 khu vực thực nghiệm76
Hình 2-15. Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM giảm độ phân
giải 20m khu vực Nghệ An77
Hình 2 16 Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM giảm độ nhận
Thin 2-10. Một số hật cát độc và hật cát ngàng của bộ du hệu DEM giam độ phản
giải 30m ở Nghệ An
 giải 30m ở Nghệ An
 rinn 2-10. Một số nặt cát dọc và mặt cát ngang của bộ dữ liệu DEM gian dộ phản giải 30m ở Nghệ An
 rinn 2-10. Một số mặt cát dọc và mặt cát ngang của bộ dữ liệu DEM gian dộ phản giải 30m ở Nghệ An
 rinn 2-10. Một số mặt cát dọc và mặt cát ngang của bộ dũ liệu DEM gian dộ phản giải 30m ở Nghệ An
 rinn 2-10. Một số mặt cát dọc và mặt cát ngang của bộ dữ liệu DEM gian dộ phản giải 30m ở Nghệ An
 rinn 2-10. Một số nặt cát dọc và mặt cát ngang của bộ dũ liệu DEM gian độ phản giải 30m ở Nghệ An
giải 30m ở Nghệ An
 rinin 2-10. Một số hiệt cát dọc và hiệt cát ngang của bộ dữ liệu DEM giản độ phản giải 30m ở Nghệ An
 rinni 2-10. Một số mặt cát dọc và mặt cát ngang của bộ dữ liệu DEM giản dộ phản giải 30m ở Nghệ An
rinni 2-10. Một số hiệt cát dọc và mặt cát ngàng của bộ dữ liệu DEM giản độ piản giải 30m ở Nghệ An
rinn 2-10. Một số hạt cát dọc và mặt cát ngang của bộ dữ liệu DEM giản dộ phản giải 30m ở Nghệ An
 rinn 2-10. Một số mặt cát dọc và mặt cát ngang của bộ dữ liệu DEM giản dộ phản giải 30m ở Nghệ An

Hình 2-22. Các biểu đồ tán xạ của các DEM tham chiếu độ phân giải không gian cao so sánh với các DEM được tăng độ phân giải - đối với bộ dữ liệu DEM mẫu độ phân giải 30m tai Đắc Hà......85 Hình 3-1. Mạng neuron Hopfield 5 nút [64].....101 Hình 3-2. Mô hình HNN sử dụng cho tăng độ phân giải của DEM dạng grid107 Hình 3-3. Sơ đồ khối thuật toán tăng độ phân giải không gian của DEM sử dụng mô hình mạng neuron Hopfield (ví dụ minh họa về tăng độ phân giải không gian của Hình 3-4. Cửa sổ chạy chương trình tăng độ phân giải không gian và nâng cao độ chính xác của mô hình số độ cao dạng grid sử dụng mạng neuron Hopfield......113 Hình 3-5. Tăng độ phân giải không gian của DEM từ độ phân giải 60m lên 20m.115 Hình 3-6. Tăng độ phân giải không gian của DEM SRTM khu vực Yên Thành, Nghệ An từ độ phân giải 90m lên 30m (tập dữ liệu D2)116 Hình 3-7. Tăng độ phân giải không gian của DEM cho tập dữ liệu S1......117 Hình 3-8. Nâng cao độ phân giải của DEM từ 90m lên 30m cho bộ dữ liệu S2118 Hình 3-9. Vị trí của các mặt cắt để đánh giá độ chính xác của DEM......124 Hình 3-10. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyển (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bô dữ liêu DEM Nghê An 20m (D1).....125 Hình 3-11. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tặng đô phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu DEM Nghệ An 30m (D2).....126 Hình 3-12. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic),

DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu
DEM Lạng Sơn 5m (S1)127
Hình 3-13. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải
bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia
mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic),
DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu
DEM Đắc Hà 30m (S2)128
Hình 3-14. Các biểu đồ tán xạ của DEM tham chiếu độ phân giải không gian cao so
với các DEM tăng độ phân giải của bộ dữ liệu DEM Nghệ An 20m (D1)130
Hình 3-15. Các biểu đồ tán xạ của DEM tham chiếu độ phân giải không gian cao so
với các DEM tăng độ phân giải của bộ dữ liệu DEM Lạng Sơn 5m (S1)131
Hình 3-16. Rải các điểm đo thực lên DEM 5m Lạng Sơn137

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Mô hình số độ cao – Digital Elevation Model (DEM) là phương pháp mô hình hóa bề mặt địa hình cũng như cho phép hiển thị bề mặt địa hình trong hệ tọa độ không gian ba chiều [63]. DEM thu hút được sự chú ý ngay từ khi nó mới bắt đầu được đưa vào sử dụng từ cuối những năm 1950s. Ngày nay, DEM được ứng dụng khá rộng rãi trong nhiều lĩnh vực như: đo đạc bản đồ, lập các mô hình về nguy cơ xói mòn đất, đánh giá thủy văn và lũ lụt, các mô hình phát tán ô nhiễm, quản lý đa dạng sinh học, trong quy hoạch, quân sự...Bề mặt DEM có thể được xây dựng từ mô hình các tam giác không đều TIN, dạng grid hoặc dưới dạng các mô hình toán học. Trong đó, các mô hình DEM ở dạng grid được sử dụng rộng rãi vì có cấu trúc đơn giản và dễ sử dụng để phân tích thông tin bề [18]. Dữ liệu về bề mặt địa hình phổ biến rộng rãi ở dạng này. Đặc biệt khi xét về các khía cạnh như lưu trữ cập nhật, tích hợp và truy cập dữ liệu của một hệ thống DEM phủ trùm quốc gia hay vùng lãnh thổ thì cấu trúc DEM dạng grid được ưa chuộng hơn so với cấu trúc dạng TIN.

Ở nước ta, từ năm 1995 đến năm 1998, Bộ KHCNMT đã có một dự án xây dựng Hệ thống thông tin địa lý phục vụ quản lý tài nguyên thiên nhiên và bảo vệ môi trường. Dự án này là một dự án với quy mô lớn đầu tiên ở nước ta để xây dựng một cơ sở dữ liệu không gian về điều kiện tự nhiên, kinh tế - xã hội với hai cấp: cấp toàn quốc với 7 cơ sở dữ liệu (CSDL) của 7 ngành và cấp tỉnh với 40 CSDL của 40 tỉnh. Các CSDL này phần lớn đều được xây dựng trên bản đồ địa hình tỷ lệ 1:50.000 nhưng trong đó không có tỉnh nào xây dựng DEM. Kết quả của dự án cũng đã xây dựng được DEM dạng Grid cho toàn quốc dựa trên CSDL địa hình tỷ lệ 1:100.000 nhưng nó chỉ như là một sản phẩm trình diễn, không ứng dụng được vào một mục đích cụ thể nào [87]. Hiện nay, DEM được thành lập từ bản đồ địa hình tỷ lệ 1:100.000 và 1: 50.000 là phổ biến nhất. Một số các ứng dụng cần DEM được thành lập từ các bản đồ địa hình tỷ lệ lớn hơn, từ 1:10.000 đến 1: 25.000.

Trong thực tế, DEM đóng vai trò khá quan trọng trong nhiều ngành, nhiều lĩnh vực có liên quan điều tra cơ bản, điều tra tài nguyên thiên nhiên, khí tượng thủy văn và biến đổi khí hậu, xây dựng và hạ tầng, nghiên cứu phòng chống tai biến địa chất như trượt lở, lũ quét, v.v. Trong các ngành ứng dụng trên, DEM đóng vai trò là dữ liệu đầu vào từ đó sử dụng các phép phân tích bề mặt địa hình để làm cơ sở cho các mô hình phân tích và dự báo. Do vậy độ phân giải và độ chính xác của DEM đóng vai trò khá quan trọng. DEM có độ phân giải và độ chính xác càng cao thì càng thể hiện được chi tiết bề mặt địa hình, từ đó các kết quả phân tích từ DEM sẽ cho độ chính xác cao hơn.

DEM có độ phân giải cao hiện nay có thể có được từ nhiều công nghệ, trong đó phải kể đến như sử dụng UAV ở khoảng cách gần, quét LiDAR từ trên máy bay và mặt đất. DEM có được từ các công nghệ này cho độ phân giải và độ chính xác rất cao. Tuy nhiên, do việc thực hiện ở tỷ lệ rất lớn nên việc xây dựng DEM cho các khu vực rộng lớn, xa xôi sẽ rất khó khăn và tốn kém và đối với một số lĩnh vực là không khả thi trong điều kiện hiện nay. Ngược lại, DEM có thể được cung cấp từ dữ liệu vệ tinh như ASTER hay STRM có độ phân giải trung bình (30m – 90m) nhưng diện tích phủ trùm rất cao. Nếu có thể nâng cao độ chính xác cho các DEM từ các nguồn này có thể cung cấp dữ liệu đầu vào tốt hơn cho các ngành có sử dụng DEM đã đề cập ở trên.

Mô hình số độ cao dạng grid có thể được làm trơn (smoothing) để tăng độ chính xác. Một loạt nghiên cứu của các tác giả khác nhau đã tiến hành đánh giá độ chính xác của các sản phẩm từ DEM dạng grid sau khi làm trơn như [18], [59], [75], [47] và [81]. Trong các nghiên cứu này, các tác giả đã xác định tác động của các phương pháp làm trơn của mô hình số độ cao DEM đến độ chính xác của các sản phẩm phái sinh như bản đồ độ dốc, bản đồ độ dốc theo hướng, tính toán khối lượng đào đắp, v.v. Mặt khác, trong các nghiên cứu này, việc đánh giá độ chính xác mới được tiến hành dựa trên một số tiêu chí đánh giá độ chính xác cơ bản như sai số trung phương (root mean square error – RMSE), sai số tuyệt đối. Trên cơ sở những nghiên cứu này, một số tác giả đã đề xuất phương pháp nâng cao độ chính xác và

mức độ chi tiết của mô hình số độ cao DEM như [78] sử dụng các phương pháp địa thống kê để nâng cao độ chính xác của DEM có độ phân giải thấp.

Mạng neuron Hopfield đã được sử dụng để nâng cao độ chính xác của lớp phủ mặt đất dựa trên kết quả phân loại mềm. Các nghiên cứu thực hiện bởi các nhóm nghiên cứu như [73], [77] và [43]. Trong nghiên cứu của Nguyễn Quang Minh và *nnk* trong [50] đã cho thấy mạng neuron Hopfield có thể tăng độ chính xác của lớp phủ mặt đất có được từ phân loại mềm. Từ những nghiên cứu trên, Nguyễn Quang Minh và *nnk* đã phát triển các phương pháp tăng độ phân giải của ảnh viễn thám bằng mô hình mạng neuron Hopfield và đã cho kết quả tốt [52]. Cũng trên cơ sở của mạng neuron Hopfield, Nguyễn Quang Minh đã sử dụng hàm mục tiêu là hàm có *semi-variogram 0* để làm mượt ảnh viễn thám đa phổ và kết quả cho thấy thuật toán trên giúp cho ảnh viễn thám đa phổ với độ phân giải cao được tạo ra có sai số trung phương RMSE (Root Mean Square Error) nhỏ hơn so với ảnh đa phổ gốc khi so sánh với ảnh tham chiếu [24].

Đánh giá tổng quan các nghiên cứu đã thực hiện trong và ngoài nước, có thể thấy rằng việc nghiên cứu các phương pháp nâng cao độ nâng cao độ chính xác của mô hình DEM bằng các thuật toán còn một số hạn chế nhất định. Các nghiên cứu về phương pháp làm trơn bề mặt địa hình đang được sử dụng như các phương pháp sử dụng hàm spline, v.v... có đặc điểm là không có các điều kiện để hạn chế sự biến đổi giá trị của các bề mặt địa hình. Ngoài ra, các phương pháp này không cho phép tăng độ phân giải của mô hình DEM dạng grid do kích thước của từng grid vẫn giữ nguyên. Cơ chế của mô hình mạng neuron Hopfield cho phép sử dụng hàm điều kiện để khống chế sự thay đổi giá trị của các grid trên bề mặt DEM đồng thời chia nhỏ các grid của DEM đang có và xác định giá trị độ cao cho các grid được chia nhỏ này. Bằng cách đó, độ phân giải không gian của grid DEM đã được tăng lên rõ rệt và làm nâng cao được độ chính xác của DEM, nhất là để có thể nâng cao độ chính xác cho các lĩnh vực phân tích DEM.

Xuất phát từ những nhu cầu thực tế trên, đề tài "Nghiên cứu nâng cao độ chính xác của mô hình số độ cao dạng grid bằng mạng neuron Hopfield" nhằm giải quyết và tiếp cận với các vấn đề trên.

2. Mục tiêu nghiên cứu

- Đánh giá được các phương pháp đánh giá độ chính xác khác nhau để đánh giá sự cải thiện về độ chính xác của DEM dạng grid khi tái chia mẫu bằng các phương pháp tái chia mẫu đang được sử dụng nhiều hiện nay;
- Xây dựng được thuật toán và chương trình tăng độ phân giải không gian và độ chính xác của mô hình số độ cao (DEM) dạng grid phù hợp với một số dạng địa hình ở Việt Nam.

3. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu chính gồm: dữ liệu DEM dạng grid được xây dựng từ các phương pháp khác nhau như: LiDar DEM, đường bình độ và đo đạc thực địa.

Phạm vi nghiên cứu của luận án bao gồm độ phân giải không gian và độ chính xác của các grid DEM nói trên.

4. Nội dung nghiên cứu

- Tổng quan về mô hình số độ cao và mạng neuron trong công tác Trắc địa -Bản đồ;
- Sử dụng các tiêu chí về độ chính xác khác nhau để đánh giá mức độ cải thiện về độ chính xác khi tái chia mẫu grid DEM bằng các phương pháp tái chia mẫu phổ biến Nearest Neighbor, Bilinear, Bi-cubic và Kriging;
- Xây dựng thuật toán nâng cao độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield;
- Xây dựng các chương trình để kiểm chứng thuật toán nâng cao độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield và kiểm chứng nhận định có thể cải thiện độ chính xác của DEM dạng grid khi tái chia mẫu bằng các phương pháp tái chia mẫu phổ biến.

5. Phương pháp nghiên cứu

 Phương pháp tiếp cận lý thuyết, phân tích và tổng hợp các dữ liệu, tài liệu liên quan;

- Phương pháp thực nghiệm: Thử nghiệm các dữ liệu thực tế để sáng tỏ cơ sở lý thuyết cho các hướng nghiên cứu đưa ra;
- Phương pháp mô hình hoá: Các hướng nghiên cứu trong luận án được mô hình hóa giúp dễ hiểu, dễ sử dụng trong quá trình xử lý dữ liệu;
- Phương pháp so sánh: Đối chiếu các kết quả nghiên cứu theo các hướng tiếp cận khác nhau để đưa ra nhận định và chứng minh tính đúng đắn trong các đề xuất mới của luận án;
- Phương pháp chuyên gia: Tiếp thu ý kiến của người hướng dẫn, tham khảo ý kiến các nhà khoa học và đồng nghiệp về các vấn đề trong nội dung luận án.

6. Ý nghĩa khoa học và thực tiễn của luận án

Ý nghĩa khoa học: Luận án đã phân tích, đề xuất và khẳng định tính đúng đắn của việc đưa ra thuật toán nâng cao độ chính xác của mô hình số độ cao (DEM) dạng grid bằng phương pháp sử dụng mạng neuron Hopfield. Xác lập tính khoa học trong mỗi hướng nghiên cứu, đề xuất trong luận án, mở ra hướng tiếp cận mới trong việc nâng cao độ chính xác của các DEM dạng grid.

Ý nghĩa thực tiễn: Bằng cách thử nghiệm các dữ liệu thực tế để khẳng định mỗi nghiên cứu, đề xuất trong luận án hoàn toàn có thể ứng dụng trong thực tiễn, góp phần giảm công sức và chi phí trong công tác xây dựng các grid DEM có độ phân giải và độ chính xác cao; đưa ra các sản phẩm có tính ứng dụng tốt nhất phục vụ cho các lĩnh vực khác nhau trong đời sống, đặc biệt là trong các công tác phân tích địa hình, địa mạo, quản lý tài nguyên thiên nhiên.

7. Các luận điểm bảo vệ

- Luận điểm 1: Các phương pháp tái chia mẫu phổ biến (phương pháp Bilinear, Bi-cubic, Kriging) cải thiện được độ chính xác của mô hình số độ cao dạng grid khi đánh giá độ chính xác của DEM bằng các tiêu chí định tính và định lượng;
- Luận điểm 2: Thuật toán nâng cao độ chính xác của mô hình số độ cao (DEM) dạng grid bằng phương pháp sử dụng mạng neuron Hopfield cho phép tăng độ phân giải không gian và độ chính xác của mô hình grid DEM.

8. Các điểm mới của luận án

- Thử nghiệm và đánh giá được rằng các phương pháp tái chia mẫu phổ biến như Bilinear, Bi-cubic và Kriging cải thiện được độ chính xác của mô hình số độ cao dạng grid;
- Đề xuất được thuật toán nâng cao độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield;
- Xây dựng được chương trình tăng độ phân giải không gian và độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield.

9. Cấu trúc và khối lượng luận án

Luận án có các phần như sau:

Mở đầu

- Chương 1 Tổng quan các vấn để nghiên cứu về mô hình số độ cao, mạng neuron và ứng dụng mạng neuron Hopfield trong các bài toán tối ưu hóa;
- Chương 2 Khảo sát về khả năng tăng độ chính xác của grid DEM bằng các thuật toán tái chia mẫu;
- Chương 3 Nghiên cứu nâng cao độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield;

Kết luận và kiến nghị.

10. Lời cảm ơn

Trước tiên, tôi xin bày tỏ sự kính trọng và lòng biết ơn sâu sắc tới PGS.TS Nguyễn Quang Minh, người Thầy đã rất tận tình hướng dẫn, động viên, giúp đỡ tôi trong suốt quá trình nghiên cứu và viết luận án này. Những nhận xét, đánh giá và đặc biệt là những gọi ý của Thầy về hướng giải quyết các vấn đề thực sự là những bài học vô cùng quý giá đối với tôi trong suốt quá trình nghiên cứu và viết luận án.

Tôi xin gửi lời cảm ơn chân thành tới các Thầy Cô trong Bộ môn Trắc địa phổ thông và Sai số, tập thể giảng viên Khoa Trắc địa - Bản đồ và Quản lý đất đai,

Ban Giám hiệu trường Đại học Mỏ - Địa chất, nơi tôi đang công tác, đã tạo điều kiện thuận lợi giúp đỡ tôi trong quá trình học tập và nghiên cứu.

Tôi xin gửi lời cảm ơn chân thành đến TS. Lê Đại Ngọc, TS. Lã Phú Hiến đã chia sẻ, cung cấp các số liệu, tài liệu để giúp tôi hoàn thành các nghiên cứu của mình.

Cuối cùng, tôi xin cảm ơn tới gia đình và người thân đã luôn bên cạnh động viên, cổ vũ tinh thần để giúp tôi hoàn thành luận án này.

Nghiên cứu sinh

Nguyễn Thị Thu Hương

Chương 1 TỔNG QUAN CÁC VẤN ĐỀ NGHIÊN CỨU VỀ MÔ HÌNH SỐ ĐỘ CAO, MẠNG NEURON VÀ ỨNG DỤNG MẠNG NEURON HOPFIELD TRONG CÁC BÀI TOÁN TỐI ƯU HÓA

1.1 Tổng quan về mô hình số độ cao

1.1.1 Các khái niệm và định nghĩa về mô hình số độ cao

Mô hình số độ cao – Digital Elevation Model (DEM) thể hiện bề mặt địa hình dưới dạng 3D theo các định dạng số. Bề mặt địa hình 3D được mô hình hóa bằng một hàm có dạng z = f(x, y) trong đó mỗi điểm (x, y) trong mặt phẳng D được gắn với một giá trị độ cao f(x, y). Theo quan điểm này, bề mặt địa hình là đồ thị biểu thị hàm số *f* theo các biến số là giá trị thuộc *D* [48].

Trên thực tế, bề mặt địa hình sẽ được đại diện bằng một số lượng hữu hạn các điểm trên bề mặt D theo các dạng như các ô vuông đều, các điểm rải rác hoặc các điểm gắn với đường bình độ (là mặt cắt của bề mặt địa hình với các mặt phẳng nằm ngang).

Tuy về mặt định nghĩa thì khái niệm bề mặt địa hình dạng số được định nghĩa khá rõ ràng, trong các tài liệu hiện nay vẫn có thể có các biến thể khác nhau bao gồm: mô hình số độ cao (Digital Elevation Model - DEM), mô hình số địa hình (Digital Terain Model - DTM) và mô hình số bề mặt (Digital Surface Model -DSM). Vì vậy trước tiên cần phải hiểu rõ và phân biệt được các khái niệm DEM, DTM và DSM.

Mô hình số bề mặt (DSM) là một mô hình số độ cao miêu tả bề mặt mặt đất, bao gồm cả các đối tượng vật thể trên đó như cây, rừng, nhà cửa...[45], [44] và [26].



Hình 1-1. Mô hình số bề mặt (DSM) và mô hình số địa hình (DTM) [13]

Mô hình số địa hình (DTM) và mô hình số độ cao (DEM) cũng là các mô hình số miêu tả bề mặt mặt đất nhưng không bao gồm các đối tượng vật thể trên đó. Giữa mô hình số địa hình (DTM) và mô hình số độ cao (DEM) thì sự phân biệt không thật sự rõ ràng. Theo Ackermann [27], tại châu Âu, không có gì khác biệt giữa mô hình số địa hình (DTM) và mô hình số độ cao (DEM). Theo Maune [70] và Ackermann [27], tại Mỹ, DTM được coi gần giống với DEM nhưng có kèm theo độ cao của các yếu tố địa hình nổi bật và của các điểm mà tại đó độ dốc địa hình thay đổi đột ngột (breaklines).

Mặc dù lúc đầu mô hình số độ cao (DEM) được xây dựng với mục đích là để mô hình hóa bề mặt địa hình nhưng chúng cũng có thể được dùng để mô hình hóa các thuộc tính Z khác (nhiệt độ bề mặt nước biển, nhiệt độ, mật độ ô nhiễm không khí...) trên một bề mặt hai chiều.



DEM

DTM (Source: www.zackenberg.dk)

Đã có nhiều nghiên cứu nói đến các khái niệm, định nghĩa về DEM [28], [27], [70], [45],... Chúng ta có thể tóm tắt như sau: Khi nói tới DEM là muốn chỉ một tập hợp của các điểm độ cao, còn khi nói về DTM thì ngoài các điểm độ cao còn có thêm thông tin về độ cao của các yếu tố địa hình nổi bật (các đường breaklines).

Một cách khái quát, chúng ta có thể định nghĩa về DEM như sau: mô hình số độ cao – Digital Elevation Model (DEM) là phương pháp mô hình hóa bề mặt địa hình cũng như cho phép hiển thị bề mặt địa hình trong hệ tọa độ không gian ba chiều [63]. DEM miêu tả bề mặt địa hình bằng các hàm đơn trị Z = F(X, Y). Với bất cứ giá trị (X, Y) nào chỉ có một giá trị độ cao Z được xác định trong một DEM [87]. Định nghĩa này giúp chúng ta phân biệt được giữa DEM mô tả bề mặt địa hình và các mô hình 3D. Các bề mặt hoặc các vật thể (ví dụ như các mô hình nhà, các chi tiết máy,...) được mô tả trong mô hình 3D, với một vị trí (X, Y) có thể có nhiều hơn một giá trị Z. Còn trong mô hình số độ cao DEM, khi mô tả bề mặt địa hình, trừ trường hợp địa hình đặc biệt như các vách đá dựng đứng hay các đứt gãy khác, còn lại bề mặt địa hình đều được mô tả bởi các hàm đơn trị nói trên.

1.1.2 Các cấu trúc của mô hình số độ cao (DEM)

Cấu trúc cơ bản của DEM xuất phát từ mô hình dữ liệu được sử dụng để đại diện cho nó. Có nhiều phương thức khác nhau để tạo ra bề mặt DEM như sử dụng mô hình DEM dạng grid, mô hình các tam giác không đều TIN (Triangulation Irregular Network) hoặc sử dụng mô hình toán học. Trong các công tác thu thập, lưu trữ và thể hiện, các dữ liệu độ cao số thường được chia thành 4 loại cơ bản là: các đường bình độ, các mặt cắt, mạng lưới đều (GRID) và mạng lưới tam giác không đều (TIN) [87]. Trong các phương pháp trên, mô hình DEM dạng grid được sử dụng nhiều vì có dạng thức đơn giản và dễ dàng sử dụng để phân tích thông tin bề mặt [18].

Trên Hình 1-3(a), dữ liệu độ cao số được mô tả bằng một tập hợp các điểm độ cao cách đều nhau, hình thành nên một mạng lưới đều (Grid DEM). Hình 1-3(b) minh họa các đường bình độ được tạo nên bởi một số các điểm rời rạc. Hình 1-3(c)

biểu diễn các điểm độ cao thuộc một mặt cắt. Còn trên Hình 1-3(d), minh họa một mạng lưới tam giác không đều (TIN).



Hình 1-3. Các phương pháp biểu diễn dữ liệu độ cao số [87]

1.1.2.1 Mô hình số độ cao dạng GRID

Ở DEM dạng GRID, các giá trị Z tại vị trí của mỗi pixel là giá trị của độ cao tuyệt đối. Khi đó, bề mặt đất có thể được xem như bị chia ra thành các ô vuông nhỏ bởi có một mạng lưới phủ đều lên. Khoảng cách giữa các mắt lưới là khoảng cách giữa hai điểm nút kế tiếp nhau. Khi khoảng cách giữa các mắt lưới đã được xác định thì mỗi ô vuông trong mạng lưới có thể được xác định vị trí bằng các tọa độ hàng/cột, có thể lấy góc cao bên trái của mạng lưới làm gốc tọa độ - tương tự như cấu trúc của ảnh số. Khi khoảng cách giữa các mắt lưới và số hàng/số cột được xác định thì các điểm nút được lưu trữ bằng cách tăng thứ tự của hàng và cột [87]. Grid DEM là cấu trúc ma trận ngầm ghi lại quan hệ Topo giữa các điểm dữ liệu và giá trị độ cao được lưu trữ tại mỗi điểm nút nói trên. Trong Grid DEM, các mắt lưới được



Hình 1-5. DEM dạng grid biểu thị bề mặt địa hình bằng các ô vuông, mỗi ô vuông có môt giá trị độ cao đại diện cho độ cao của các điểm nằm trong ô vuông đó và DEM dạng grid biểu thị bề mặt địa hình bằng mạng lưới nối các điểm độ cao được phân bố đều theo hàng dọc và ngang [61]



Hình 1-4. DEM dạng GRID có cấu trúc như một ảnh dạng raster trong đó bề mặt chia thành các ô vuông và mỗi ô vuông có một giá trị độ cao [34]

thể hiện theo hai hình thức: hoặc là lưu trữ theo điểm (các điểm độ cao) như trong DEM của Mỹ, hoặc là cả một pixel với kích thước là khoảng cách mắt lưới như DEM 9" của Úc. Theo Singare và Kale [42], Grid là cấu trúc ma trận ngầm ghi lại quan hệ Topo giữa các điểm dữ liệu. Cấu trúc dữ liệu DEM dạng Grid tương tự như cấu trúc Raster của file ảnh số. Trong đó, Raster là một grid dạng số, thống nhất, hình vuông, bao phủ một diện tích nhất định trên bề mặt trái đất. Mỗi pixel của DEM dạng grid có một giá trị độ cao nhất định là giá trị độ cao trung bình đại diện cho toàn bộ các điểm mà pixel đó bao phủ được trên bề mặt trái đất [89]. Các Grid DEM có nhiều loại kích thước nhưng phổ biến nhất là từ 5m đến 30m.

Nếu thể hiện bề mặt địa hình dưới dạng 3D có thể thấy bề mặt địa hình sẽ được sấp sỉ hóa bằng các hình trụ vuông mà chiều cao của hình trụ chính là độ cao so với bề mặt chuẩn (có thể là mặt geoid) như Hình 1-5. Theo cách biểu thị này, bề mặt địa hình sẽ càng được mô phỏng chính xác khi kích thước các ô vuông (mắt lưới) càng nhỏ. Và như vậy một bề mặt địa hình sẽ được biểu diễn bằng một hoặc nhiều ảnh có cấu trúc dữ liệu dạng raster với mỗi điểm ảnh (pixel) tương ứng với một ô vuông trong lưới ô vuông GRID và giá trị *Z* của mỗi điểm ảnh chính là giá trị độ cao của ô vuông trong lưới ô vuông GRID [34].

Một cách biểu thị GRID DEM khác là việc sử dụng một ma trận các điểm tạo thành lưới ô vuông đều, trong đó mỗi điểm đỉnh lưới ô vuông là một điểm độ cao. Nếu biểu diễn DEM dạng GRID theo cách này thì bề mặt địa hình 3D sẽ được biểu diễn như hình vẽ bên trái trong Hình 1-5 trong đó bề mặt địa hình tương đối trơn và giống với bề mặt địa hình thực. Khác với GRID DEM biểu thị theo cách đầu tiên, dữ liệu GRID DEM lưu trữ theo cách này sẽ là dữ liệu dạng vector, trong đó mỗi điểm mắt lưới sẽ là một đối tượng dữ liệu dạng điểm với thuộc tính là Z. Cách biểu thị GRID DEM này có ưu điểm là mô hình bề mặt 3D khá trơn và mượt, có cảm giác giống bề mặt địa hình thực, tuy nhiên dữ liệu lưu trữ theo cấu trúc dữ liệu vector và về bản chất cũng tương tư như mô hình số đô cao dang mang lưới tam giác không đều (TIN) do vây có những nhược điểm của mô hình số đô cao dang TIN như khó thực hiên phân tích không gian bề mặt địa hình, bề mặt mô hình số địa hình là một bề mặt rời rạc, mỗi khi cần xác định độ cao của một điểm địa hình bất kỳ sẽ phải nội suy lại dữ liệu để xác định giá trị độ cao cho điểm này. Mặt khác, dữ liệu gốc để xây dựng DEM có độ phân giải trung bình, có quy mô lớn và phổ biến hiện nay chủ yếu là dữ liệu ảnh, do đó dữ liệu GRID DEM theo định dạng raster được sử dụng nhiều và rất phổ biến.

1.1.2.2 Mô hình số độ cao dạng TIN

Mô hình số độ cao ở dạng TIN (Triangulated Irregular Networks) đã được phát triển như là một cách đơn giản để xây dựng một bề mặt từ một tập hợp các điểm có sự phân bố không đều. Cấu trúc dữ liệu dạng TIN dựa trên các điểm, đường, vùng có phân bố không đồng đều và thường được chia ra thành các đám điểm (mass points) và các đường breaklines [27]. Theo Ackermann [27], TIN còn được gọi là mạng lưới tam giác không quy chuẩn, là tập hợp của các tam giác liền kề, không chồng đè, không có tam giác đảo (tam giác nằm bên trong một tam giác khác). Mạng TIN có thể được xây dựng từ nhiều nguồn dữ liệu: các đường bình độ được số hóa, các điểm đặc trưng địa hình hoặc từ Grid DEM. Trong cấu trúc này, thông thường các điểm không phản ánh các đặc trưng địa hình sẽ bị loại bỏ và chỉ giữ lại các đỉnh tam giác với giãn cách lớn nhất có thể, tùy theo sự thay đổi của địa hình (tam giác Delaunay). Theo phân loại mô hình dữ liệu, mô hình dữ liệu TIN là kiểu dữ liệu véc tơ.

1.1.2.3 Đường bình độ

Cấu trúc đường bình độ được mô hình hóa bằng một tập hợp các điểm nằm đủ gần nhau để tái tạo đường cong với đủ độ chính xác cần thiết bằng cách nối hai điểm cạnh nhau bởi một đoạn thẳng. Mỗi đường cong được thể hiện bằng một tập hợp các đoạn thẳng này. Vì thế nên đường cong dạng số không thể mô tả một cách chính xác một đường cong thực tế vì nó được tạo bởi một tập hợp các điểm rời rạc. Do đó, trên thực tế luôn cần phải áp dụng thêm các kỹ thuật làm trơn đường cong, lấy bỏ tổng hợp hóa hoặc chêm dày điểm để cải thiện hình dạng của đường bình độ, tùy thuộc vào tỷ lệ bản đồ cần thành lập. Dưới dạng số, một đường bình độ được xác định bằng: độ cao H của nó và tọa độ mặt bằng của tất cả các điểm. Còn các mặt cắt (Profiles) là tập hợp của các điểm độ cao dọc theo một hướng nhất định, trong đó các điểm độ cao được đo với độ giãn cách thay đổi, tùy theo độ dốc của địa hình. Trong dạng này, các điểm độ cao được thể hiện dọc theo các đường song song với nhau. Các giá trị độ cao thường được đo tại các điểm có độ dốc thay đổi. Dưới dạng số, mặt cắt được lưu trữ bằng cách mã hóa một trong hai tọa độ mặt bằng và độ cao nên sẽ chiếm nhiều bộ nhớ hơn dạng lưới Grid [87].

1.1.2.4 So sánh giữa các cấu trúc dữ liệu DEM

Các dữ liệu đường bình độ cho phép miêu tả chính xác độ cao dọc theo các đường này nhưng giữa các đường bình độ thì chỉ cung cấp các thông tin tương đối chính xác về độ cao và độ dốc. Hay nói cách khác, nó thường chỉ cung cấp đầy đủ các thông tin về độ cao dọc theo các đường bình độ. Còn giữa các đường bình độ hoặc theo hướng vuông góc với các đường bình độ thì không được đầy đủ, thậm chí là thiếu các thông tin về độ cao. Còn các mặt cắt thì chỉ đủ để cung cấp các thông tin khá là rời rạc về địa hình dọc theo lát cắt. Do đó, trong các dạng cấu trúc trên, cấu trúc dữ liệu DEM dưới dạng mạng lưới đều (Grid) và dạng TIN là thông dụng hơn so với các đường bình độ và các mặt cắt, trong đó mô hình DEM dạng Grid được sử dụng nhiều vì có dạng thức đơn giản và dễ dàng sử dụng để phân tích thông tin bề mặt [18].

Cấu trúc dữ liệu DEM dạng GRID khá giống với dữ liệu ảnh số. Nếu đặt số thứ tự *i*, *j* tương đương với số hàng, số cột trong ảnh số và độ cao *Z* của DEM có thể được liên hệ với giá trị độ xám của pixel. Với cấu trúc dữ liệu này, tọa độ mặt phẳng của các điểm độ cao có thể được lược bỏ hay không cần phải biểu thị ra trực tiếp mà thông qua một phép tính đơn giản, tương tự như đếm số hàng, số cột trong ảnh số. Điều này hoàn toàn khác so với mô hình TIN vì trong mô hình TIN, mỗi đỉnh của một tam giác phải được lưu trữ rõ ràng và đầy đủ cả *X*, *Y*, *Z*. Hơn nữa, sau khi tạo tam giác xong còn phải thiết lập thêm các mối quan hệ liền kề Topology giữa các tam giác với nhau. Trong cấu trúc Grid DEM, mặc dù số điểm mắt lưới có thể lớn hơn số điểm độ cao trong mạng TIN nhiều lần nhưng dung lượng tệp tin lại thường nhỏ hơn do có cấu trúc và cách lưu trữ đơn giản hơn [87].

Trong một số nghiên cứu trong [27], [63], [28] và [42] đã so sánh Grid DEM với TIN DEM nhưng đã rút ra một kết luận: không có cấu trúc nào trong hai cấu trúc này tỏ ra là hoàn toàn ưu việt cho tất cả các ứng dụng liên quan tới mô hình hóa bề mặt địa hình nên việc chọn cấu trúc nào còn phải tùy thuộc vào dạng phân tích dữ liệu cần thiết (các phép phân tích không gian) trong các ứng dụng cụ thể. Ví dụ khi cần tính toán lưu vực sông, phân tích thủy văn, phân tích vùng đệm, phân tích kề cận, v.v. thì người ta thường sử dụng Grid DEM. Còn trong các ứng dụng mà các thông tin vi địa hình (micro-relief information) được coi là quan trọng thì mô hình TIN tỏ ra có ưu việt hơn. Hai cấu trúc DEM dạng Grid và DEM dạng TIN có thể chuyển đổi cho nhau.

Theo Ackermann [27] và GS. Trương Anh Kiệt [86], trên thực tế, trong mô hình TIN, phép nội suy trong các tam giác là nội suy tuyến tính. Trong phép nội suy này, mặt phẳng được xác định bởi 3 điểm của tam giác được coi là mặt địa hình và giá trị độ cao của một điểm bất kỳ được nội suy bởi tam giác chứa điểm đó. Mỗi tam giác là một mặt phẳng trong không gian ba chiều và độ cao của một điểm nằm bên trong tam giác được xác định bởi tiếp điểm của mặt phẳng tam giác với đường thẳng đứng (theo hướng dây dọi) đi qua điểm đó. Vì vậy trong phép nội suy này, độ cong của địa hình trong nội bộ mỗi tam giác đã không được xem xét tới. Do đó, theo Ackermann [27], dạng TIN hầu như không được đề xuất cho DEM phủ trùm toàn quốc với hàng triệu điểm. Ở các nước như Mỹ, Úc, Anh, Đức, Nhật, Trung Quốc và nhiều nước khác đều xây dựng DEM phủ trùm toàn quốc theo cấu trúc dạng Grid. Và hiện nay có nhiều thuật toán để xử lý DEM dạng Grid hơn dạng TIN.

1.1.2.5 Phân tích về cấu trúc DEM dạng Grid

Dưới đây, tác giả sẽ phân tích thêm về cấu trúc DEM dạng Grid. Đây là một cấu trúc phù hợp cho hệ thống DEM phủ trùm quốc gia với mục đích phục vụ quản lý tài nguyên thiên nhiên và cũng là đối tượng chính mà tác giả đã lựa chọn để nghiên cứu trong luận án.

Theo [27] và [69], DEM ở dạng Grid còn được gọi là DEM dạng lưới ô vuông quy chuẩn hay ma trận độ cao. Các điểm độ cao trong DEM dạng này được bố trí theo khoảng cách đều theo hai hướng tọa độ X, Y để biểu diễn địa hình. Trong đó tọa độ mặt phẳng của một điểm bất kỳ có độ cao Z được xác định theo số thứ tự (i, j) của ô lưới theo hai hướng trên:

 $X_{i} = X_{o} + i. \Delta_{X} \text{ với (} i = 0, 1, 2, ..., n_{X-1} \text{)}$

 $Y_{i} = Y_{o} + j. \Delta_{Y} \quad v \acute{\sigma}i (j = 0, 1, 2, ..., n_{Y-1})$ (1.1)

Trong đó:

 $X_{\rm o}$, $Y_{\rm o}$ là tọa độ của điểm gốc lưới ô vuông;

 Δ_X , Δ_Y là khoảng cách của mắt lưới trên các hướng X, Y;

 n_X , n_Y là số ô lưới trên hai hướng tọa độ X, Y của mô hình số độ cao.

Trong công thức trên, hướng tọa độ *X*, *Y* có thể là các tọa độ vuông góc theo lưới chiếu bản đồ hoặc cũng có thể là hệ tọa độ địa lý. Ví dụ như các DEM theo cấu trúc lưới UTM của Mỹ với $\Delta_X = \Delta_Y = 30$ mét hay 10mét, còn nếu được thành lập theo hệ tọa độ địa lý thì Δ_X , Δ_Y được thay bằng Δ_{φ} , Δ_{λ} và thường được tính theo đơn vị giây cung của kinh tuyến và vĩ tuyến (như DEM 9" của Úc hay DEM 3" hoặc 30" của Mỹ).



Hình 1-6. DEM theo lưới UTM của Mỹ với $\Delta_X = \Delta_Y = 30 mét$ (USGS, 1993) [28]

Trong Hình 1-6, bốn góc của mảnh DEM được chia theo tọa độ địa lý nhưng các điểm lưới trong DEM lại được chia theo lưới chiếu UTM nên có độ không song song giữa chúng (giữa hệ tọa độ địa lý và hệ tọa độ theo lưới chiếu UTM). Nhưng khoảng cách giữa các mắt lưới hoàn toàn đều đặn. Còn trong Cấu trúc dữ liệu DEM dạng GRID khá giống với dữ liệu ảnh số. Nếu đặt số thứ tự *i*, *j* tương đương với số

hàng, số cột trong ảnh số và độ cao Z của DEM có thể được liên hệ với giá trị độ xám của pixel. Với cấu trúc dữ liệu này, tọa độ mặt phẳng của các điểm độ cao có thể được lược bỏ hay không cần phải biểu thị ra trực tiếp mà thông qua một phép tính đơn giản, tương tự như đếm số hàng, số cột trong ảnh số. Điều này hoàn toàn khác so với mô hình TIN vì trong mô hình TIN, mỗi đỉnh của một tam giác phải được lưu trữ rõ ràng và đầy đủ cả *X*, *Y*, *Z*. Hơn nữa, sau khi tạo tamDEM ở dạng này không gặp phải vấn đề không song song nêu trên nhưng vì khoảng cách mắt lưới được tính theo chiều dài cung tròn của các đường kinh, vĩ tuyến nên giá trị độ dài của Δ_X thay đổi theo vĩ độ. Vì vậy khi sử dụng cơ sở dữ liệu địa hình, lấy hệ tọa độ theo lưới chiếu bản đồ làm cơ sở có nhiều ưu điểm hơn cả.

1.1.3 Các phương pháp thành lập mô hình số độ cao (DEM)

Theo Florinsky và Nelson và *nnk [20]*, DEM có thể được tạo ra từ nhiều nguồn khác nhau. Hiện nay, DEM chủ yếu được tạo ra từ các nguồn dữ liệu sau: từ các kết quả đo đạc thực địa, từ các dữ liệu được số hóa trên các bản đồ đã có sẵn (như số hóa các đường bình độ và các điểm ghi chú độ cao), từ kết quả đo vẽ ảnh hàng không và ảnh vệ tinh, từ các dữ liệu đo Radar độ mở tổng hợp giao thoa và laser đặt trên máy bay, từ dữ liệu đo UAV.

Từ dữ liệu đo đạc thô, các điểm đo, hoặc các đường bình độ, để có dữ liệu DEM là bề mặt liên tục cần sử dụng các phương pháp nội suy. Các phương pháp nội suy như: tuyến tính, song tuyến, đa thức bậc 3, trung bình trọng số, nội suy dựa vào hàm Spline và phương pháp nội suy Kriging thường được sử dụng phổ biến trong các phần mềm để thành lập DEM.

Dưới đây là một số phương pháp thành lập mô hình số độ cao phổ biến :

1.1.3.1 Đo đạc địa hình thông thường

Chúng được thực hiện bằng các máy đo quang học, máy đo laser tacheometers và các máy thủy bình. Các DEM được thành lập bằng phương pháp này thường rất chi tiết và có tỷ lệ lớn, thường được thành lập ở các khu vực có diện tích tương đối nhỏ. Các DEM loại này thường được ứng dụng trong nghiên cứu đất đai và địa chất.

1.1.3.2 Phương pháp đo GPS động

Phương pháp này sử dụng hệ thống vệ tinh định vị toàn cầu và các máy thu GPS được gắn trên một chiếc xe di chuyển trên bề mặt địa hình để thu thập số liệu để thành lập ra DEM. Đây là phương pháp thành lập các DEM tỷ lệ lớn và chi tiết với thời gian nhanh và hiệu quả cao.

1.1.3.3 Phương pháp ảnh tương tự và ảnh số

Phương pháp ảnh tương tự và ảnh số. Chúng được sử dụng để thành lập các DEM từ các cặp ảnh (ảnh hàng không, ảnh vệ tinh). Người ta có thể sử dụng ảnh vệ tinh từ nhiều nguồn khác nhau như ảnh SPOT, Landsat MSS, NOAA AVHRR, ASTER, Ikonos, QuickBird,... Đo ảnh số đã được sử dụng để sản xuất ra hai loại DEM tỷ lệ trung bình là ASTER DEM cho hầu hết các phần của bề mặt trái đất và SPOT DEM cho các phần của diện tích Á-Âu, châu Phi và Trung Mỹ.

Trong các nghiên cứu "siêu thực" (superdetailed) độ chi tiết rất cao ví dụ như DEM trong các mô hình vật lý trong phòng thí nghiệm để nghiên cứu các quá trình kiến tạo như gấp khúc và đứt gãy, người ta đã sản xuất DEM với độ phân giải vài milimet từ các cặp ảnh lập thể [87].

1.1.3.4 Kỹ thuật Radar

Được sử dụng phổ biến nhất. Kỹ thuật này với Radar độ mở tổng hợp (InSAR) để sản xuất DEM với ba cách tiếp cận: phân cực, radargrammetry sử dụng cặp ảnh radar lập thể và giao thoa kế. Trong các phương pháp đó, Radar độ mở tổng hợp giao thoa là được sử dụng phổ biến nhất. Đặc biệt, phương pháp này đã được thực hiện để tạo ra DEM độ phân giải thấp toàn cầu Venus, SRTM3 và DEM độ phân giải trung bình cho một phần lớn bề mặt trái đất [35].

DEM độ phân giải thấp toàn cầu được sử dụng để mô tả địa hình (như trong DEM SRTM3_PLUS và DEM toàn cầu) [36]. SARs trên các vệ tinh ERS-1 và Radarsat đã được sử dụng để sản xuất DEM của các khối băng ở Greenland và Nam Cực [82],... DEM InSAR được sử dụng trong nghiên cứu các bề mặt đất vừa và nhỏ, trong nghiên cứu, khảo sát địa chất địa cầu [79]. Đôi khi các kỹ thuật Radar mặt đất được sử dụng để tạo ra DEM độ phân giải cao. Các dữ liệu từ SAR trên vệ tinh Cassini đã được sử dụng để sản xuất các mô hình số vùng bờ của hố sâu hồ hydrocacbon Ontario trên Titan.

1.1.3.5 Phương pháp đo laser

Phương pháp này sử dụng các xung laser để xác định khoảng cách giữa mục tiêu và bộ cảm biến (sensor). Trong phương pháp sử dụng công nghệ LiDAR, các phép đo đạc trên không đã được sử dụng để tạo ra các DEM tỷ lệ lớn và chi tiết của cả bề mặt đất và mặt biển. Mặc dù một vài vấn đề kỹ thuật vẫn chưa được giải quyết (ví dụ: hiệu quả của việc lọc nhiễu), kỹ thuật LiDAR đo đạc trên không vẫn được đề xuất để lập các DTM độ phân giải cơ sở cho trong nghiên cứu khoa học đất, địa chất trên quy mô lớn [67]. Sử dụng laser vệ tinh độ cao, hai DEM toàn cầu của Mặt Trăng đã được tạo ra: DEM quy mô nhỏ và DEM quy mô trung bình, một loạt các DEM toàn cầu quy mô vừa và nhỏ của Sao Hỏa và các DEM cho các khối băng ở Nam Cực và Greenland [57]. LiDAR mặt đất đã được sử dụng để tạo ra các DEM độ phân giải cao để từ đó tạo ra được các mô hình địa chất ba chiều.

1.1.4 Độ chính xác bề mặt mô hình số địa hình (DEM)

Độ chính xác của DEM được xác định bằng độ giống nhau giữa độ cao xác định trên bề mặt DEM của một điểm và giá trị độ cao thực tế [58]. Để xác định được mức độ giống nhau này, cần sử dụng các điểm kiểm định đã biết cả tọa độ và độ cao, sau đó dựa vào tọa độ của điểm xác định độ cao tương ứng trên bề mặt DEM. Có hai đại lượng có thể đặc trưng cho độ chính xác về độ cao của bề mặt DEM được sử dụng nhiều trong các nghiên cứu trước đây là sai số trung phương (RMSE) và sai số trung bình (ME) [88].

Sai số trung phương được xác định bằng công thức (1.1):

$$RMSE_{Z} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (Z_{di} - Z_{ri})^{2}}$$
(1.1)

Trong đó $RMSE_Z$ là giá trị sai số trung phương; Z_{di} là giá trị độ cao thứ i trên bề mặt DEM kết quả của phương pháp tái chia mẫu; Z_{ri} là giá trị độ cao thứ i trên bề mặt DEM tham khảo; *n* là số lượng điểm độ cao kiểm tra. Sai số trung phương là

đại lượng được sử dụng nhiều nhất để đặc trưng cho độ chính xác của DEM vì dựa vào sai số trung phương có thể xác định độ lớn của sai số ngẫu nhiên tồn tại trên bề mặt DEM.

Sai số trung bình được xác định theo công thức:

$$ME_{Z} = \left[\frac{1}{n}\sum_{i=1}^{n} (Z_{di} - Z_{ri})\right]$$
(1.2)

Trong đó, ME_Z là sai số trung bình còn các giá trị n, Z_{di} , và Z_{ri} là các thành phần giống như trong Công thức (1.1).

Khác với sai số trung phương, sai số trung bình sẽ cho biết mức độ tồn tại của sai số hệ thống trên bề mặt DEM. Nếu sai số trung bình > 0, DEM được xác định là cao hơn bề mặt thực tế. Nếu sai số trung bình < 0 thì DEM sẽ có xu thế thấp hơn bề mặt thực tế.

Để xác đinh xu thế chung của bề mặt DEM so với thực tế, một số tác giả cũng sử dụng phép hồi quy tuyến tính, trong đó xây dựng hàm hồi quy tuyến tính giữa các điểm độ cao trên DEM và thực tế ở cùng một vị trí và tìm hệ số tương quan giữa hai tập dữ liệu độ cao DEM và thực tế [88]. Việc sử dụng các đại lượng này sẽ cho cái nhìn trực quan hơn về sự phân bố của sai số ở các mức độ cao khác nhau trên bề mặt địa hình.

Trong một nghiên cứu về độ chính xác của DEM của F. J. Aguilar và cộng sự [8], các tác giả đã tìm hiểu 3 cách tiếp cận để ước tính và đánh giá độ chính xác của mô hình số độ cao (DEM) và sử dụng các điểm kiểm tra để đánh giá độ chính xác. Cách tiếp cận đầu tiên là sử dụng sai số trung phương (RMSE) để tính toán độ chính xác của DEM và xác định một số tham số thống kê như Chi-squared và Asymptotic Student-T để xem xét sự phân phối sai số. Cách tiếp cận thứ ba được phát triển trong bài viết này là một phương pháp mới dựa trên lý thuyết về các hàm ước tính có thể được coi là tổng quát hơn nhiều so với hai trường hợp trước. Nó dựa trên một cách tiếp cận phi tham số trong đó không có phân phối cụ thể nào được giả định. Do đó, có thể tránh được những sai số về tính quy tắc phân phối trong phần lớn các tiêu chuẩn hiện tại về độ chính xác vị trí. Ba phương pháp đã được thử nghiệm bằng cách sử dụng mô phỏng Monte Carlo cho một số dữ liệu đo thừa được

tạo ra từ dữ liệu được lấy mẫu ban đầu. Các grid DEM gốc được coi là dữ liệu mặt đất, được thu thập bằng phương pháp quang điện tử kỹ thuật số từ 7 khu vực với các hình thái địa hình khác nhau, sử dụng khoảng cách lấy mẫu 2m x 2m. Các grid DEM ban đầu được ghép lại để tạo các DEM mới có độ phân giải thấp hơn. Mỗi DEM mới này sau đó được nội suy để lấy độ phân giải ban đầu của nó bằng hai thuật toán khác nhau. Sự khác biệt về độ cao giữa các grid DEM gốc và DEM nội suy đã được tính toán để thu được các quần thể còn lại. Từ nghiên cứu này các tác giả đã xác định số lượng mẫu phù hợp để đánh giá độ chính xác cho DEM. Trong thực tế, cần ít nhất 64 điểm kiểm tra là cần thiết để xây dựng một MSE với khoảng tin cậy chính xác 95%. Tuy nhiên, khi hình thái địa hình có xu hướng bình thường, có thể đạt được xấp xỉ tốt chỉ với 16 điểm kiểm tra. Do đó, cách tiếp cận này sẽ tiết kiệm thời gian và tiền bạc trong việc đánh giá độ chính xác của DEM so với phương pháp chỉ sử dụng tiêu chuẩn RMSE.

Sai số độ cao trên bề mặt DEM ngoài sai số do quá trình đo đạc và sai số nội suy bề mặt thì các yếu tố khác như phương thức thể hiện DEM (grid, TIN) hay độ phân giải (kích thước của grid) cũng gây ra một số nguồn sai số nhất định [22]. Chẳng hạn kích thước các grid quá lớn sẽ gây ra các sai số khi yêu cầu mức độ chi tiết cao [37].

Để xác định các cấp hạng cho dữ liệu DEM và dữ liệu độ cao địa không gian, có thể xây dựng các quy chuẩn về độ chính xác dữ liệu trong đó xác định rõ sai số trung phương về độ cao và mặt bằng của mỗi cấp hạng, như tiêu chuẩn về độ chính xác dữ liệu địa không gian của Hiệp hội Đo ảnh và viễn thám Mỹ [23], hay chuẩn quốc gia về độ chính xác của dữ liệu địa không gian do Ủy ban thông tin địa lý Liên Bang Mỹ [31]. Các tiêu chuẩn này xác định rõ yêu cầu về độ chính xác dữ liệu độ cao địa không gian được xác định bằng sai số trung phương và có so sánh tương đương với các tiêu chuẩn trước đây về độ chính xác địa hình dựa vào khoảng cao đều đường đồng mức. Chẳng hạn theo tiêu chuẩn về độ chính xác dữ liệu không gian của Hiệp hội Đo ảnh và viễn thám Mỹ (Bảng 1-1và

Bảng 1-2) thì mô hình số độ cao ở có sai số trung phương xác định độ cao là 66.7 cm sẽ tương ứng với địa hình được biểu thị bằng đường đồng mức với khoảng cao đều 2m (theo hạng 1 về độ chính xác) và 1m (theo hạng 2 về độ chính xác) hoặc 2.19 m theo tiêu chuẩn NMAS.

Bảng 1-1. Bảng phân cấp độ chính xác dữ liệu độ cao số theo Hiệp hội Đo ảnh và Viễn thám Mỹ [23]

	Độ chính xác tuyệt đối			Độ chính xác tương đối		
Cấp hạng	RMSEz Không có cây (cm)	Sai số giới hạn lấy khoảng tin cậy 95% (cm)	VVA at 95th Percent ile (cm)	Khu vực có bề mặt phản xạ tốt (cm)	Khu vực có cây (RMSDz) (cm)	Swath-to- Swath Non-Veg Terrain (Max Diff) (cm)
1-cm	1.0	2.0	3	0.6	0.8	1.6
2.5-cm	2.5	4.9	7.5	1.5	2	4
5-cm	5.0	9.8	15	3	4	8
10-cm	10.0	19.6	30	б	8	16
15-cm	15.0	29.4	45	9	12	24
20-cm	20.0	39.2	60	12	16	32
33.3-cm	33.3	65.3	100	20	26.7	53.3
66.7-cm	66.7	130.7	200	40	53.3	106.7
100-cm	100.0	196.0	300	60	80	160
333.3-cm	333.3	653.3	1000	200	266.7	533.3

Bảng 1-2. So sánh độ chính xác trên dữ liệu mô hình số độ cao và độ chính xác tính theo khoảng cao đều đường đồng mức tương ứng [23]

Cấp độ chính xác	RMSEz Không có thực phủ (cm)	Tương ứng với khoảng cao đều đường đồng mức Hạng 1 (ASPRS 1990) (cm)	Tương ứng với khoảng cao đều đường đồng mức Hạng 2 (ASPRS 1990) (cm)	Tương ứng với khoảng cao đều đường đồng mức theo chuẩn NMAS (cm)
1-cm	1.0	3.0	1.5	3.29
2.5-cm	2.5	7.5	3.8	8.22
5-cm	5.0	15.0	7.5	16.45
10-cm	10.0	30.0	15.0	32.90
15-cm	15.0	45.0	22.5	49.35
Cấp độ chính xác	RMSEz Không có thực phủ (cm)	Tương ứng với khoảng cao đều đường đồng mức Hạng 1 (ASPRS 1990) (cm)	Tương ứng với khoảng cao đều đường đồng mức Hạng 2 (ASPRS 1990) (cm)	Tương ứng với khoảng cao đều đường đồng mức theo chuẩn NMAS (cm)
---------------------	------------------------------------	---	---	--
20-cm	20.0	60.0	30.0	65.80
33.3-cm	33.3	99.9	50.0	109.55
66.7-cm	66.7	200.1	100.1	219.43
100-cm	100.0	300.0	150.0	328.98
333.3-cm	333.3	999.9	500.0	1096.49

1.1.5 Các ứng dụng của mô hình số độ cao

Hiện nay, DEM có rất nhiều ứng dụng trong các lĩnh vực của đời sống kinh tế, xã hội như trong quản lý thiên tai, giao thông, trong thông tin liên lạc, dẫn đường, trong xây dựng các công trình dân dụng, trong thiết kế và xây dựng cơ sở hạ tầng, trong quân sự,... Trong đó, DEM có vai trò to lớn trong việc phân tích kết quả, ra quyết định và phát triển sản phẩm.

Một số các thông số được tính toán từ DEM và các ứng dụng của chúng (Bảng 1-3) [87]:

Thông số	Úng dụng			
Độ cao	Xác định tiềm lực năng lượng, các biến số của khí hậu, áp suất, nhiệt			
	độ, đặc tính của đất và cây trồng, tính toán khối lượng đào đắp.			
Độ dốc	Xác định độ dốc địa hình, dòng chảy bề mặt và dưới đất; phân loại			
	đất, thực phủ; hiệu chỉnh ảnh viễn thám.			
Hướng dốc	Nghiên cứu bức xạ mặt trời, sự bốc hơi nước, các thuộc tính thực phủ;			
	hiệu chỉnh ảnh viễn thám.			
Độ cong dọc theo	Xác định gia tốc của dòng chảy, các vùng gia tăng xói mòn hoặc bồi			
hướng dốc	đắp, các chỉ sô đánh giá đất và thổ nhưỡng.			
Độ cong vuông góc với	Nghiên cứu các dòng chảy hội tụ, phân tán; phân tích tính chất của			
hướng dốc	đất.			
Hướng dòng chảy cục	Tính toán các thuộc tính của vùng lưu vực như là một hàm số của thủy			
bộ	hệ; đánh giá sự vận chuyển của vật chất trong mạng thủy hệ cục bộ.			
Vùng lưu vực	Phân tích lưu vực, khối lượng vật chất chảy ra khỏi khu vực.			
Chỉ số năng lượng dòng	Nghiên cứu khả năng gây xói mòn của dòng chảy bề mặt.			
chảy				

Bảng 1-3. Một số các thông số được tính toán từ DEM và các ứng dụng của chúng

Thông số	Úng dụng			
Chỉ số vận chuyển trầm	Nghiên cứu đặc tính của các quá trình xói mòn và bồi lắng.			
tích				
Chỉ số địa hình	Xác định chỉ số duy trì độ ẩm			
Tầm nhìn	Phân tích vùng thông hướng nhìn của các trạm tiếp song, tháp canh,			
	các ứng dụng quân sự, các công trình cao tầng.			
Bức xạ	Nghiên cứu thổ nhưỡng, cây trồng, sự bốc hơi, vị trí xây dựng các			
	công trình tiết kiệm năng lượng.			

DEM cũng là một thành phần cơ bản trong cơ sở dữ liệu địa hình và nó có một vai trò to lớn trong các ngành khoa học nghiên cứu về trái đất. Có thể kể đến trong một số ngành khoa học sau:

- Trong phân tích địa mạo, cảnh quan;
- Trong các nghiên cứu phục vụ mục đích thành lập bản đồ;
- Trong mô hình hóa địa chất và thủy học;
- Trong quản lý nguồn tài nguyên nước;
- Trong nghiên cứu ảnh hưởng của khí hậu;
- Trong công nghệ thông tin địa lý;
- Trong các nghiên cứu phục vụ giáo dục.

Ngoài ra, DEM và các sản phẩm dẫn xuất từ DEM còn được ứng dụng trong các lĩnh vực sau:

- Khảo sát, thăm dò địa chất;
- Thiết kế và xây dựng;
- Các dịch vụ khí tượng thủy văn;
- Dẫn đường hàng không;
- Thông tin viễn thông;
- Sử dụng trong các ứng dụng đa phương tiện và trò chơi điện tử.

Trong đo đạc bản đồ, các chức năng về phân tích DEM là rất ít mà chủ yếu là các ứng dụng: thành lập DEM, đánh giá chất lượng DEM, thành lập bản đồ ảnh trực giao và bản đồ địa hình, hiện chỉnh bản đồ [87].

1.1.6 Công tác thành lập DEM ở trong và ngoài nước

1.1.6.1 Thành lập DEM ở nước ngoài

Những kỹ thuật mới trong công nghệ viễn thám đã làm tăng khả năng thành lập các DEM với độ phủ trùm lớn và độ phân giải cao. Theo Dell's Acqua and Gamba [38], dự án X-SAR/SRTM đo địa hình bằng thiết bị radar độ mở tổng hợp giao thoa (băng X- với bước sóng $\lambda = 3.1$ cm, đặt trên tàu vũ trụ con thoi Endeavour được phóng vào tháng 2 năm 2000), đã cho phép xây dựng được DEM với độ phủ trùm 80% bề mặt trái đất (từ vĩ độ 60° Bắc đến vĩ độ 56° Nam), độ phân giải mặt bằng 30m, ước tính độ chính xác tương đối về độ cao các điểm của DEM là 6m và độ chính xác tuyệt đối là 16m [87].

Công ty ISTAR của Pháp là nhà phân phối DEM lớn nhất toàn cầu, tiếp đó là công ty Atlantis Scientific Inc., của Mỹ. Đến 2004, ISTAR đã sử dụng các ảnh SPOT là nguồn dữ liệu chính và sau đó bổ sung các dữ liệu từ ERS và RADASAT. Một nguồn dữ liệu DEM khác được xây dựng ở quy mô toàn cầu là dữ liệu DEM từ chương trình Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) trên tàu vũ trụ NASA cung cấp dữ liệu từ cặp ảnh lập thể với độ phân giải 15m cho phép thành lập DEM bằng cách tạo cặp ảnh lập thể tự động và cho DEM với độ phân giải từ 30m đến 150m với độ chính xác độ cao (tính bằng sai số trung phương) trong khoảng từ ±7 and ±15m [60]. Từ năm 2011 đến 2015, Trung tâm Hàng không vũ trụ Đức - German Aerospace Center (DLR), và các công ty EADS Astrium GmbH and Infoterra GmbH cũng đã quan trắc trái đất rất nhiều lần và xây dựng được bộ dữ liệu DEM toàn cầu với độ phân giải 90m [29]. Theo một số nghiên cứu thì thị trường DEM được thành lập từ nhiều nguồn khác nhau với độ chính xác từ 1m đến 100m sẽ ngày càng tăng trưởng mạnh.

Công nghệ đo laser đặt trên máy bay LIDAR (Light Detection And Ranging) để thành lập DEM với độ chính xác cao. Theo Dell's Acqua and Gamba [38] DEM được thành lập bằng công nghệ LIDAR đã được Cơ quan quản lý tình trạng khẩn cấp của Mỹ (FEMA) sử dụng để nghiên cứu bảo hiểm lũ lụt và thành lập bản đồ đánh giá tình trạng lũ lụt tại một số bang như Carolina, Louisiana và Texas. Trong đó, DEM độ chính xác cao được sử dụng kết hợp với các dữ liệu khác để thành lập bản đồ thủy văn và dự đoán diện tích ngập nước. DEM từ nguồn dữ liệu LIDAR cũng được sử dụng ở nhiều nước châu Âu như: Anh, Pháp, Đức, Bỉ, Hà Lan, ... trong công tác phòng chống lũ lụt và dự đoán thiên tai [76].

Thời gian gần đây, cùng với sự phát triển của công nghệ định vị vệ tinh, công nghệ máy bay không người lái (Unmanned Aerial Vehicles-UAV), các thuật toán tự động xử lý ảnh, khôi phục mô hình 3 chiều (Structure-from-Motion, SfM), giá thành trang thiết bị cho công nghệ đo vẽ sử dụng UAV đã giảm xuống rất mạnh nên công nghệ này đang được ứng dụng thành công và rộng rãi trong nhiều lĩnh vực như công tác đo đạc thành lập bản đồ, giao thông, sản xuất nông nghiệp, nghiên cứu địa chất, và nghiên cứu môi trường. Điểm cần bổ sung hiện nay của công nghệ này, là quy trình tính toán, công nghệ thiết kế bay chụp UAV sao cho đạt được độ chính xác mong muốn của các sản phẩm bản đồ, cho từng mục đích cụ thể [89].

1.1.6.2 Tình hình thành lập DEM ở Việt Nam

Từ năm 1995 đến năm 1998, Bộ KHCNMT đã có một dự án xây dựng Hệ thống thông tin địa lý phục vụ quản lý tài nguyên thiên nhiên và bảo vệ môi trường. Dự án này là một dự án với quy mô lớn đầu tiên ở nước ta để xây dựng một cơ sở dữ liệu không gian về điều kiện tự nhiên, kinh tế - xã hội với hai cấp: cấp toàn quốc với 7 cơ sở dữ liệu (CSDL) của 7 ngành và cấp tỉnh với 40 CSDL của 40 tỉnh. Các CSDL này phần lớn đều được xây dựng trên bản đồ địa hình tỷ lê 1:50.000 nhưng trong đó không có tỉnh nào xây dựng DEM. Kết quả của dự án cũng đã xây dựng được DEM dạng Grid cho toàn quốc dựa trên CSDL địa hình tỷ lệ 1:100.000 nhưng nó chỉ như là một sản phẩm trình diễn, không ứng dụng được vào một mục đích cụ thể nào [87].

Các DEM được xây dựng bằng các phần mềm SURFER, PCI đã được Viện Địa lý – thuộc Trung tâm Khoa học Tự nhiên và Công nghệ quốc gia, nghiên cứu và sử dụng trong các bài toán phân tích, đánh giá về điều kiện tự nhiên của các vùng lãnh thổ. Và cơ quan này đã xây dựng được DEM phủ trùm toàn quốc từ bản đồ địa hình tỷ lệ 1: 1.000.000, với khoảng cách mắt lưới là 250m. Loại DEM này được sử dụng trong công tác đánh giá tiềm năng xói mòn, hỗ trợ công tác phân loại lớp phủ bằng tư liệu viễn thám độ phân giải trung bình. Bên canh đó, Viện Địa lý cũng đã thành lập được DEM độ phân giải 30m, từ các bản đồ địa hình tỷ lệ lớn, cho các khu vực: thành phố Hạ Long, Huế, Tánh Linh, Bình Thuận. Các DEM này phục vụ cho công tác quy hoạch môi trường và vùng lãnh thổ.

Ở Trung tâm viễn thám đã thành lập được các loại DEM sau: dùng ảnh SPOT để hiện chỉnh bản đồ địa hình tỷ lệ 1 : 50.000, thành lập được DEM phủ trùm từ bản đồ địa hình tỷ lệ 1 : 250.000; các DEM được thành lập từ các bản đồ địa hình tỷ lệ 1 : 100.000 và 1 : 50.000; thành lập được DEM từ các cặp ảnh SPOT lập thể, từ cặp ảnh SPOT 4 với độ chính xác về độ cao khoảng 7 – 8 m, từ SPOT 5 với độ chính xác khoảng 4-5 m.

Viện Khoa học Đo đạc Bản đồ đã xây dựng được các quy định kỹ thuật trong việc xây dựng DEM phục vụ cho công tác nắn ảnh trực giao, thành lập bản đồ và đã đề xuất coi DEM là một sản phẩm độc lập. Viện cũng đã ứng dụng DEM trong công tác thành lập bản đồ địa hình, bản đồ địa chính. Bên cạnh đó, DEM cũng đã được sử dụng để tính số cải chính địa hình trong mô hình Geoid.

DEM cũng đã được sử dụng từ rất sớm tại Viện Điều tra quy hoạch rừng trong các nghiên cứu, điều tra về tài nguyên rừng và trong các bài toán phân tích không gian phục vụ quản lý và phát triển rừng. Tại cơ quan này, DEM được xây dựng chủ yếu bằng phương pháp số hóa từ bản đồ địa hình rồi dựa trên các đường bình độ đã số hóa để nội suy DEM, chủ yếu dùng phần mềm Ilwis.

Đối với DEM độ phân giải cao, công nghệ LiDAR đang được ứng dụng khá nhiều tại Việt Nam. Từ năm 2007, Hệ thống Lidar tích hợp ảnh kỹ thuật số Harrier 56/G4 kết hợp máy quét Lidar LMS-Q560 (Riegl) với máy ảnh Rollei AIC P45 và phần mềm xử lý dữ liệu Lidar TOPPIT được đưa vào sản xuất thành lập bản đồ và CSDL ở Việt Nam từ năm 2007 trong Dự án "Thành lập cơ sở dữ liệu nền thông tin địa lý tỉ lệ 1:2.000, 1:5.000 các khu vực đô thị, khu vực công nghiệp, khu kinh tế trọng điểm" và từ năm 2012 hệ thống này tham gia thực hiện dự án "Xây dựng mô hình số độ cao độ chính xác cao khu vực đồng bằng và ven biển phục vụ công tác nghiên cứu, đánh giá tác động của biến đổi khí hậu, nước biển dâng". Năm 2018 bắt đầu giai đoạn tiếp theo của công nghệ Lidar tích hợp chụp ảnh hàng không ở Việt nam với việc Tổng Công ty Tài nguyên và Môi trường Việt Nam đầu tư hệ thống Citymapper của hãng Leica (Thụy sỹ) [85]. Hệ thống này đã được sử dụng cho Dự án "Xây dựng mô hình số độ cao độ chính xác cao khu vực đồng bằng và ven biển phục vụ công tác nghiên cứu, đánh giá tác động của biến đổi khí hậu, nước biển dâng" với các sản phẩm đám mây điểm Lidar, mô hình số bề mặt DSM, mô hình số địa hình DTM và ảnh nắn trực giao (orthophoto).

Tuy đã có một số khu vực trọng điểm được thành lập DEM ở độ phân giải cao, hiện nay, ở Việt Nam, DEM được thành lập từ bản đồ địa hình tỷ lệ 1:100.000 và 1: 50.000 là phổ biến nhất. Một số các ứng dụng cần DEM được thành lập từ các bản đồ địa hình tỷ lệ lớn hơn, từ 1:10.000 đến 1: 25.000 [87].

Một số loại DEM và các ứng dụng của nó ở nước ta hiện nay (Bảng 1-4):

ТТ	Độ phân	Ứng dụng	Khu vực	Phần mềm	Dữ liệu gốc	Cơ quan
	giải			ứng		xây dựng
	DEM			dụng		
1	20m	Đánh giá tiềm	Thanh Ba, Hà		Landsat TM, độ	Trung tâm
		năng xói mòn	Hòa, Phú Thọ;	ILWIS	phân giải 30m;	viễn thám
		theo mô hình	Khu vực núi Ba		bản đồ địa hình tỷ	và
		DEM	Vì		lệ 1:50.000	Geomatics
						(VTGEO)
2	Vùng	Thành lập bản	Thừa Thiên Huế,	Phần	Bản đồ địa hình	Trung tâm
	đồng	đồ ngập lụt	Quảng Nam, Đà	mềm tự	tỷ lệ 25.000 và	khí tượng
	bằng 5m,	theo mô hình	Nẵng	xây	50.000	Thủy văn
	vùng núi	thủy lực, thủy		dựng		quốc gia
	10m	văn				
3	20m	Thành lập bản	Nam Thanh, Hải	ILWIS	Bản đồ địa hình	VTGEO
		đồ nguy cơ	Dương		tỷ lệ 50.000	
		ngập lụt				
4	30m	Quy hoạch sử	Vườn Quốc gia	ILWIS	Landsat TM; bån	Viện Khoa
		dụng tài	Ba Bể; Tiền Hải,		đồ 50.000	học Thủy
		nguyên, đánh	Thái Bình; Hàm			lợi
		giá thích nghi	Tân, Bình Thuận			
		cây trồng				

Bảng 1-4. Một số loại DEM ở nước ta hiện nay

TT	Độ phân giải DEM	Ứng dụng	Khu vực	Phần mềm ứng dụng	Dữ liệu gốc	Cơ quan xây dựng
5	20m	Xây dựng bản	Đồng bằng sông	ILWIS	Landsat TM; bån	VTGEO
		đồ địa mạo	Hồng		đồ 10.000	
6	5m	Đánh giá biến	Khu công nghiệp	ILWIS	Bản đồ 10.000	VTGEO
		động khu công	Dung Quất			
		nghiệp				
7	15m	Nắn trực chiếu	Hạ Hòa, Phú Thọ	PCI	RADARSAT,	VTGEO
		ånh Radar			bản đồ 1: 25.000	
					và 1: 50.000	

1.1.7 Một số nghiên cứu về cải thiện và đánh giá độ chính xác DEM

Ngoài việc cải thiện độ chính xác của DEM bằng các phương pháp và công nghệ đo đạc, có nhiều nghiên cứu sử dụng các thuật toán để nâng cao độ chính xác cho DEM, đặc biệt là các thuật toán này sử dụng phối hợp nhiều nguồn thông tin khác nhau để nâng cao độ chính xác cho DEM.

1.1.7.1 Cải thiện độ chính xác của DEM bằng các dữ liệu bổ sung

Trong một nghiên cứu của Carlos A. Felgueiras và cộng sự [11] đã sử dụng các điểm độ cao có độ chính xác cao để hiệu chỉnh lại mô hình số độ cao. Trong nghiên cứu này, một tập các điểm độ cao có độ chính xác cao hơn độ cao của các điểm trên DEM đã được sử dụng trong mô hình tính toán kết hợp theo thuật toán địa thống kê KED (Kriging with an External Drift). Phần thực nghiệm đã thực hiện tính toán kết hợp giữa dữ liệu DEM SRTM và một tập hợp các điểm độ cao mẫu thu được từ vùng Campinas, thành phố San Paulo ở Brazil đã cho thấy sự cải thiện về độ chính xác về độ cao của DEM SRTM khi được kết hợp sử dụng cùng 505 điểm độ cao mẫu.

Phương pháp địa thống kê đã được sử dụng trong một nghiên cứu khác của C. A. Felgueiras và cộng sự [11] cho phép kết hợp dữ liệu độ cao thu được từ các nguồn khác nhau, cấu trúc dữ liệu và độ chính xác khác nhau để cải thiện độ chính xác độ cao của mô hình số độ cao (DEM). Nhóm nghiên cứu đã sử dụng các kỹ thuật như: OCK (Ordinary Cokriging), KED (Kriging with an External Drift),

Kriging hồi quy (RK) để mô hình hóa dữ liệu hợp nhất được thực hiện từ các DEM hiện có, chủ yếu có sẵn miễn phí trên internet và bổ sung tập hợp các điểm mẫu 3D có độ chính xác cao. Kết quả định lượng của nghiên cứu đã cho thấy những cải thiện về độ chính xác của các sản phẩm DEM hợp nhất (DEM kết quả) cho khu vực nghiên cứu. Trong đó phương pháp KED và RK mang lại kết quả định lượng tốt hơn phương pháp CoK. Do đó, bài báo đã chứng minh rằng có thể cải thiện độ chính xác của các DEM hiện có bằng các công cụ địa thống kê được xem xét trong nghiên cứu này.

Trong một nghiên cứu của Yunwei Tang và cộng sự, năm 2014 [10], nhóm nghiên cứu đã đề xuất một phương pháp sử dụng kỹ thuật thống kê địa lý đa điểm (MPG) dựa trên mẫu cải tiến để kết hợp các mô hình số đô cao (DEM) với mục đích là hợp nhất thông tin không gian địa lý từ các nguồn khác nhau để nâng cao độ chính xác cho DEM. Trong đó, mối tương quan không gian được đặc trưng bởi thống kê nhiều điểm. Trong MPG, dựa trên mẫu, hai bộ dữ liệu có thể được tích hợp bằng cách sử dụng dữ liệu thứ cấp làm giá trị trung bình thay đổi cục bộ (LVM). Trong cách tiếp cận này, một phương pháp mới để hình thành lớp nguyên mẫu được áp dụng, dựa trên bề mặt còn đã có, phép đo độ gồ ghề của địa hình bằng các véctơ (VRM) và độ phức tạp tại những vùng địa hình đặc trưng (RVC) của dữ liệu địa hình. Phương pháp này đã được thử nghiệm trên dữ liệu SRTM và GMTED2010. Dữ liệu SRTM ở độ phân giải không gian 3 cung giây được mô phỏng bằng cách kết hợp dữ liệu điểm độ cao rải rác và dữ liệu GMTED2010 ở độ phân giải không gian thấp hơn là 7,5 cung giây. Phương pháp MPG đề xuất được so sánh với mô phỏng MPG dưa trên mẫu truyền thống. Một số công cu dư báo kriging đã được áp dụng để cung cấp LVM cho mô phỏng MPG. Kết quả cho thấy rằng phương pháp mới có thể đạt được dự đoán chính xác hơn và giữ lại nhiều chi tiết không gian hơn so với các điểm chuẩn.

1.1.7.2 Ảnh hưởng của các phương pháp nội suy đến độ chính xác của DEM

Một trong những thách thức khoa học quan trọng của mô hình số độ cao là việc biểu diễn bề mặt địa hình của các khu vực rộng lớn với độ phân giải cao. Mặc

dù đã có nhiều nghiên cứu về độ chính xác của các kỹ thuật nội suy để tạo ra các

mô hình số độ cao liên quan đến các dạng địa mạo và số lượng hoặc mật độ dữ liệu, vẫn cần phải đánh giá hiệu suất của các kỹ thuật này trên các cảnh quan tự nhiên khác nhau, hình thái và trên một loạt các quy mô khác nhau. Để thực hiện việc đánh giá như vậy, trong một nghiên cứu của Vincent Chaplot và cộng sự [12], đã nghiên cứu khảo sát tại 6 khu vực, 3 khu vực ở miền núi phía Bắc Lào và 3 khu vực ở vùng đồi núi phía Tây nước Pháp, với nhiều diện tích bề mặt khác nhau từ khoảnh nhỏ, vùng đồi núi và các lưu vực. Các kỹ thuật nội suy: IDW (Inverse Distance Weighting), OK (Ordinary Kriging), UK (Universal Kriging), **MRBF** (Multiquadratic Radial Basis Function) và RST (Regularized Spline with Tension) đã được sử dung để nôi suy các dữ liêu điểm đô cao với các giá tri mật đô từ 4 đến 109 điểm / km2. Kết quả nghiên cứu cho thấy, tại các khu vực có hệ số biến thiên (CV-coefficient of variation) về độ cao từ 12% đến 78%, cấu trúc không gian đẳng hướng hay dị hướng với các mức độ mạnh, yếu khác nhau (yếu với tỷ lệ nugget/sill là 0,8, mạnh (0,01)) thì sự thay đổi về phân bố độ cao có rất ít sự khác biệt giữa các phương pháp nội suy nếu mật độ lấy mẫu cao. Mặc dù kỹ thuật nội suy MRBF hoạt động tốt hơn khi mật độ lấy mẫu thấp hơn, kỹ thuật Kriging là phương pháp nội suy tốt nhất cho địa hình có cấu trúc không gian mạnh (có CV thấp và dị hướng thấp), trong khi RST lại là phương pháp nội suy tốt nhất cho địa hình có CV thấp và cấu trúc không gian yếu. Trong điều kiện CV cao, cấu trúc không gian manh và tính di hướng manh, phương pháp nôi suy IDW hoat đông tốt hơn một chút so với các phương pháp còn lai. Các kết quả nghiên cứu này cũng chỉ ra rằng, đô chính xác của các kỹ thuật nôi suy để tao DEM cần được đánh giá theo cả các dạng địa mạo và mật độ dữ liệu.

Sự thay đổi từ bản đồ giấy sang GIS, trong các loại ứng dụng và phân tích dữ liệu địa lý khác nhau, đã làm cho việc sử dụng dữ liệu không gian các ứng dụng khác nhau và việc kết hợp một số các lớp thành các mô hình không gian khá phức tạp, bao gồm cả không gian ba chiều tham chiếu (bề mặt) trong mô hình số bề mặt (DTM). Trong một nghiên cứu của Gabriela Droj và cộng sự [1], nhóm nghiên cứu đã so sánh các thuật toán phổ biến nhất liên quan đến việc tạo ra một DEM, để thiết lập các yếu tố chính ảnh hưởng đến độ chính xác của DEM và để cải thiện chất lượng của DEM được tạo ra. Có một số phương pháp nội suy thường được sử dụng trong GIS và so sánh 8 phương pháp được sử dụng rộng rãi Inverse distance weighted (IDW), Spline Biquadratic interpolation, Spline Bicubic interpolation, Bspline interpolation, Nearest Neighbor - Voronoi diagram, Delaunay Triangulation, Quadratic Shepard interpolation, Kriging interpolation. Hiệu suất của 8 phương pháp, được đánh giá dựa trên độ chính xác của bề mặt được tạo ra. Các kết quả đưa ra bằng các phương pháp tính toán khác nhau bị ảnh hưởng bởi nhiều yếu tố như: đặc điểm bề mặt địa hình, mật độ của các điểm, chất lượng của các giá trị đã biết và thuật toán đã sử dụng. Kết quả thực nghiệm cho thấy rằng: dữ liệu đầu vào là cơ sở để tính toán DEM, mật độ của các điểm đã biết là một yếu tố quan trọng trong việc tăng chất lượng DEM và không có thuật toán tối ưu cho tất cả.

Sẽ không đúng khi khái quát rằng một phương pháp nội suy nào đó (ví dụ: Kriging, IDW, Spline,...) tốt hơn các phương pháp khác mà không tính đến cấu trúc, loại và bản chất của tập dữ liệu và các yếu tố liên quan. Vì thế, trong một nghiên cứu của Maduako Nnamdi Ikechukwu và cộng sự [9], các tác giả đã đánh giá về mặt lý thuyết, toán học và thực nghiệm hiệu suất của các phương pháp nội suy Kriging, IDW và Spline tương ứng trong việc ước tính các giá trị độ cao và lập mô hình địa hình. Nghiên cứu này thực hiện một phân tích so sánh dựa trên tiêu chuẩn sai số trung phương của các phương pháp. Kết quả thử nghiệm cho mỗi phương pháp trên cho thấy phương pháp Spline cho kết quả nội suy tốt hơn và chính xác hơn đối với dữ liệu độ cao được thu thập trực tiếp từ khảo sát thực địa không ngẫu nhiên đồng nhất và không chuẩn hóa so với phương pháp IDW và Kriging. Từ kết quả nội suy cũng cho thấy trong phương pháp Spline, các dự đoán bằng RBF không bị giới hạn trong phạm vi giá trị đo được, tức là giá trị dự đoán có thể cao hơn giá trị đo tối đa hoặc nhỏ nhất.

1.1.7.3 Đánh giá độ chính xác của mô hình DEM

Trong một nghiên cứu của Annamaria Castrignanò và cộng sự năm 2006 [72], các tác giả đã trình bày một cách sử dụng địa thống kê để đánh giá độ chính xác của

DEM. Các điểm độ cao chuẩn được phân bố rải rác được kết hợp với độ cao ước tính của DEM để tính toán sai số cho một khu vực nhỏ. Sau đó sai số của tập dữ liệu này được sử dụng để ước tính các sai số cho mô hình tại mỗi nút DEM. Phương pháp mô phỏng thống kê này được tạo ra bởi quy trình kiểu Monte Carlo cho các mô hình cục bộ này và được sử dụng để xây dựng các phép đo không gian về độ chính xác của DEM, ví dụ: bản đồ xác suất của các sai số dương hoặc âm, nghĩa là cho điểm độ cao nội suy cao hơn hay thấp hơn địa hình thật. Cách tốt nhất để kiểm tra chất lượng DEM là đánh giá kết quả dựa trên xác suất thống kê, ước tính dựa trên mô hình sai số được phân phối theo không gian sử dụng mô phỏng ngẫu nhiên có điều kiện. Phương pháp được sử dụng có thể được thực hiện trong một số các số liệu thống kê địa lý được kết hợp với GIS; bằng cách này, nó có thể được sử dụng ngay để trở thành một phần của bộ công cụ GIS tiêu chuẩn.

Trong một nghiên cứu của Høhle Joachim và Höhle Michael năm 2009 [5], đã trình bày các biện pháp đánh giá độ chính xác của mô hình số độ cao (DEM) và các đặc điểm của DEM thu được từ quá trình quét laze và phép đo ảnh tự động. Nhóm nghiên cứu đã đề xuất về các yêu cầu đối với dữ liệu tham chiếu về độ chính xác và đề xuất các phương pháp thống kê xác định độ chính xác của DEM. Sau đó, nhóm tác giả đã kiểm chứng kết quả bằng 4 bộ số liệu DEM và đã đưa ra kết luận rằng các biện pháp như sai số trung phương, độ lệch chuẩn và các phương pháp thống kê nên được sử dụng để đánh giá độ chính xác của các DEM có nguồn gốc từ đo ảnh kỹ thuật số và quét laze. Phương pháp đánh giá độ chính xác trong nghiên cứu này có thể được tóm tắt như sau: Tính toán sai số dọc với tất cả các điểm trong mẫu. Sau đó, tạo biểu đồ và biểu đồ Q – Q để hình dung lỗi phân phối và để đánh giá tính không chuẩn mực. Sau đó, tính toán sai số trung phương và độ lệch chuẩn cùng với khoảng tin cậy. Các tiêu chuẩn để đánh giá độ chính xác DEM phải tính đến các yếu tố này và các phương pháp phân phối, phi tham số phải được tính toán trong đánh giá độ chính xác của DEM.

Trong một nghiên cứu khác của Miles và cộng sự [2], nhóm nghiên cứu đã so sánh các giá trị độ cao, độ dốc và hướng dốc từ các DEM ASTER (một cảm biến thụ động được đặt trên vệ tinh Terra sử dụng chế độ lập thể các kỹ thuật tương quan để tạo DEM) và DEM SRTM tới dữ liệu thu thập được trên mặt đất ở một số vườn quốc gia và khu bảo tồn ở Nepal. Mục tiêu chính là xác định khả năng sử dụng của các phép đo cảm biến từ xa của các đặc tính địa hình (độ dốc và hướng dốc) thông qua đánh giá sai số và những hạn chế liên quan. Điều này được thực hiện bằng cách đo độ lớn của sai số giữa dữ liệu mặt đất và dữ liệu viễn thám, sau đó phân phối độ lớn của các sai số thành các lớp. Dữ liệu không gian địa lý từ các khu vực nghiên cứu được phân tích bằng cách sử dụng Hệ thống thông tin Địa lý (GIS), phân tích viễn thám và các khống chế mặt đất (GCP). Một bộ điều khiển từ xa phân tích cảm biến được sử dụng để phân tích sự khác biệt sai số giữa các dữ liệu viễn thám và dữ liệu GCP trong 6 khu vực của một khu bảo tồn ở Nepal. Sự khác biệt về sai số giữa GCP và DEM được sử dụng để đánh giá độ chính xác của DEM.

Kết quả cho thấy những hạn chế của DEM được tạo ra từ viễn thám ở những vùng núi hiểm trở. Kích thước và độ thô của pixel là một lý do có thể tạo ra sai số. Vì SRTM và ASTER sử dụng pixel 90m và 30m tương ứng trên một địa hình không đồng nhất và có sự thay đổi nên độ chính xác của dữ liệu địa hình rất đa dạng tùy thuộc vào độ gồ ghề của nó. Đô phân giải không gian càng thô (kích thước pixel càng lớn), bề mặt địa hình càng mịn, nhưng các chi tiết của cảnh quan thì mất đi. SRTM có tổng hợp dữ liệu về địa hình lớn hơn nhiều so với ASTER, nhưng đáng ngạc nhiên là nó không ảnh hưởng đáng kế đến độ chính xác. Siêu dữ liệu thông báo rõ ràng cho người dùng về những hạn chế này. Xem xét địa hình của Makalu Barun, nơi có vách đá dưng đứng và các hẻm núi, pham vi sai số tiêu chuẩn trong việc thu thập dữ liệu mặt đất cao hơn nhiều so với các công viện khác. Điển hình là cái khác công viên có sai số tiêu chuẩn GPS dao đông trong khoảng 4-5m. Điểm có vách đá dựng đứng có sai số khoảng 9-10m. Lưu ý: đối với toàn bộ dự án, bất kỳ vị trí điểm nào có lỗi tiêu chuẩn lớn hơn 10m đã bị loại bỏ. Đánh giá độ dốc và hướng dốc trên mặt đất cũng là một vấn đề. Địa hình của những khu vực này rất hiểm trở và nhiều khu vực không dễ tiếp cận. Một yếu tố tiềm năng khác để tạo ra sai số là thuật toán đã tạo ra độ dốc và các bề mặt.

Trong một nghiên cứu của Fernando J. Aguilar và cộng sự [74] các tác giả đã nghiên cứu sự ảnh hưởng của hình thái địa hình, mật độ lấy mẫu và phương pháp

nôi suy đối với dữ liêu mẫu phân tán về đô chính xác của đô cao nôi suy trong Mô hình số độ cao dạng grid (grid DEM). Dữ liệu mẫu được thu thập với khoảng cách lấy mẫu 2 x 2 mét theo 7 phương pháp khác nhau, áp dụng phương pháp đo ảnh kỹ thuật số cho ảnh hàng không tỷ lệ lớn (1: 5000). Phân tích này đưa ra một số kết luận như sau: Độ chính xác của DEM (tiêu chuẩn RMSE) bị ảnh hưởng đáng kể bởi các biến được nghiên cứu trong bài báo này theo phương pháp "nội suy mật độ lấy mẫu hình thái học". Trong đó, phương pháp RBF (Multiquadric Radial Basis Function) được đánh giá là phương pháp nôi suy tốt nhất, mặc dù Multilog RBF hoat đông tương tư đối với hầu hết các hình thái. Phần còn lai của các phép nôi suy RBF được thử nghiêm (Đường khối khối từ nhiên, Đường đa dang nghịch đảo và Đường trục mảng mỏng) cho thấy sự không ổn định với hệ số làm trơn (smoothing) thấp. Phương pháp IDW (Inverse Distance Weighted) có trọng số hoạt động kém hơn RBF Multiquadric hoặc RBF Multilog. Ngoài ra, người ta thấy rằng mối quan hệ giữa sai số trung phương (RMSE) và mật độ lấy mẫu N được xác định bằng hàm số giảm dần, có thể được biểu thị là RMSE/Sdz 0,1906 (N/M) 0,5684 (R2 0,8578). Trong đó, Sdz độ lệch chuẩn của độ cao của M điểm kiểm tra được sử dụng để ước tính độ chính xác và N số lượng điểm lấy mẫu được sử dụng để tạo DEM. Kết quả thu được trong nghiên cứu này cho phép thiết lập các mối quan hệ thực nghiệm giữa sai số trung phương (RMSE) dự kiến trong phép nội suy DEM dạng lưới và các biến như độ gồ ghề của địa hình, mật độ lấy mẫu và phương pháp nội suy, trong số những biến khác có thể được thêm vào. Do đó, có thể thiết lập trước kích thước lưới tối ưu cần thiết để tao ra hoặc lưu trữ DEM với đô chính xác cu thể, với tính kinh tế về thời gian tính toán và kích thước têp. Điều này sẽ rất hữu ích để thiết lập các chiến lược lấy mẫu thích ứng liên quan đến hình thái của bề mặt địa hình đang được lập mô hình.

Trong nghiên cứu về sự ảnh hưởng của độ chính xác của mô hình số độ cao đối với thủy văn và địa mạo của Walker và cộng sự [4], đã so sánh giữa các mô hình số độ cao (DEM) có các khoảng cách mắt lưới (grid) khác nhau trong kỹ thuật đo bản đồ và đo ảnh với bộ dữ liệu điểm độ cao mẫu trên mặt đất (thu được bằng đo đạc ngoại nghiệp trên mặt đất), và nghiên cứu tác động của những khác biệt này đối với thủy văn bằng các số liệu thống kê. Kết quả cho thấy đã có sự khác biệt đáng kế giữa diện tích lưu vực và hướng dòng chảy được nội suy tính toán từ các DEM với các yếu tố đó ngoài thực tế trong hầu hết các trường hợp. Hơn nữa, kết quả tương quan được xác định từ các DEM được nâng cao độ chính xác luôn nằm ngoài giới hạn về độ tin cậy từ 60% đến 90%, cho thấy rằng các đặc tính thủy văn này được các định kém chính xác từ các DEM đã nâng cao độ chính xác. Tuy nhiên, các tính toán nội suy về diện tích, độ dốc được xác định từ các DEM dẫn xuất được cho là ít nhạy cảm hơn hình dạng, kích thước và mạng lưới dòng chảy. Nghiên cứu này cũng đã khảo sát về mối quan hệ giữa độ phân giải mặt đất của DEM đầu vào với độ chính xác theo chiều dọc của các điểm kiểm tra và dự đoán độ phân giải mặt bằng khoảng 10m là phù hợp với các DEM được thử nghiệm. Nghiên cứu cũng chỉ ra rằng khoảng cách mắt lưới của DEM có ảnh hưởng đáng kể đến độ chính xác của DEM trong việc xác định đô dốc, diện tích lưu vực và hướng dòng chảy một cách tổng thể.

1.2 Tổng quan về mạng neuron

Mô phỏng sinh học, đặc biệt là mô phỏng các chức năng hoạt động của bộ não người là mơ ước từ lâu của loài người. Với số lượng lớn các neuron và các mối liên kết giữa chúng làm cho bộ não người có cấu trúc cực kỳ phức tạp. Điều này cho phép bộ não người có khả năng tư duy, nghi nhớ những sự kiện quá khứ, dự đoán, tổng quát hóa, nhận dạng, phân loại, điều khiển. Việc nghiên cứu bộ não người theo khía cạnh giải phẫu học, tâm lý học, thần kinh học để hiểu biết các nguyên tắc hoạt động của bộ não là rất cần thiết. Từ đó cho phép chúng ta tạo ra được những hệ thống thông minh có thể giải quyết nhiều vấn đề phức tạp trong đo lường, điều khiển tự động, hệ thống chuyên gia, công nghệ robot.v.v...

Công trình nghiên cứu về não đầu tiên thuộc về Ramond- Cajál (1911), ông cho rằng hệ thần kinh được cấu tạo từ các neuron. Bộ não con người chứa khoảng 10^{11} các phần tử liên kết chặt chẽ với nhau (khoảng 10^4 liên kết đối với mỗi phần tử) gọi là các neuron.

1.2.1 Cấu tạo của một neuron sinh học

Các thành phần chính của một neuron sinh học:

- *Khớp kết nối (Synapse):* là phần nối giữa dây thần kinh ra (axon) của neuron này với dây thần kinh vào (dendrite) của neuron khác.

- *Xúc tu (Dendrites):* là các dây thần kinh vào có nhiệm vụ thu nhận thông tin về nhân qua khớp. Dây thần kinh vào có độ dài khoảng 200-300μm. Mỗi tế bào thần kinh nhận nhiều đầu vào (khoảng 10⁴) và sau các quá trình xử lý sẽ tạo ra một tín hiệu đầu ra truyền dọc theo dây thần kinh ra (axon).

Thân tế bào (Cell body): có đường kính khoảng 30μm. Trong thân tế bào là nhân tế bào thần kinh, có chức năng tổng hợp các tín hiệu và khi đủ mạnh thì có tín hiệu ra ở trục cảm ứng.

- Trục cảm ứng (Axon): là dây thần kinh ra (hay còn gọi là sợi trục thần kinh), có độ dài từ 50µm cho đến vài mét. Thông tin tạo ra bởi tế bào thần kinh được truyền dọc theo dây thần kinh ra. Trung bình có khoảng 10.000 khớp thần kinh nối với mỗi dây thần kinh ra, đưa tín hiệu ra và truyền tới các neuron khác qua các khớp



Hình 1-7. Cấu trúc của một neuron sinh học [16]

kết nối [21].

1.2.2 Nguyên lý hoạt động của các neuron

Các tế bào thần kinh tự nhiên nhận được tín hiệu thông qua các khớp thần kinh (synapse) nằm trên dây thần kinh vào (dendrites) hoặc màng tế bào thần kinh

tới tế bào thân, tế bào thân sẽ thực hiện gộp (sum) và phân ngưỡng (threshold) các tín hiệu đến. Khi các tín hiệu nhận được đủ mạnh (vượt qua ngưỡng nhất định), neuron sẽ được kích hoạt và phát ra một tín hiệu tới sợi trục. Sợi trục thần kinh làm nhiệm vụ đưa tín hiệu từ tế bào thân ra ngoài. Tín hiệu này có thể được gửi đến một khóp thần kinh khác, và có thể kích hoạt các neuron khác.

Sự sắp xếp của các neuron và mức độ mạnh yếu của các khớp thần kinh được quyết định bởi các quá trình hoá học phức tạp, sẽ thiết lập chức năng của mạng neuron.

Một vài neuron có sẵn từ khi sinh ra, các phần khác được phát triển thông qua việc học, ở đó có sự thiết lập các liên kết mới và loại bỏ các liên kết cũ. Các khớp khi mới sinh ít kết nối với nhau, chúng được kết nối nhờ quá trình học.

1.2.3 Khái niệm và cấu trúc của mạng neuron nhân tạo

Mạng neuron nhân tạo (Artificial Neural Networks – ANNs) hay thường được gọi ngắn gọn là mạng neuron là một phương pháp tính toán mới có nền tảng từ sinh học nhằm mô phỏng một số chức năng của bộ não con người, dựa trên quan điểm cho rằng bộ não người là bộ điều khiển.

Hai thành phần chính cấu tạo nên mạng neuron là các neuron nhân tạo (mô phỏng các tế bào thần kinh) và các synapse (mô phỏng các khớp nối thần kinh). Trong kiến trúc của một mô hình kết nối, các neuron chính là các nút mạng, được liên kết với nhau thông qua các synapse, là các cung mạng. Các nút mạng được nối với nhau và xử lý thông tin bằng cách truyền theo các kết nối và tính giá trị mới tại các nút.

Các neuron chính là các đơn vị xử lý thông tin cơ sở của mạng neuron. Mỗi neuron là một đơn vị tính toán có nhiều đầu vào và một đầu ra, mỗi đầu vào đến từ một synapse. Trong đó, synapse là một thành phần liên kết giữa các neuron, nối đầu ra của neuron này với đầu vào của neuron khác. Trên cơ sở mạng neuron sinh học, mạng neuron nhân tạo ở dạng đơn giản nhất gồm chỉ có một neuron (còn gọi là dạng mang Perceptron) được cấu tạo như Hình 1-9 [16]. Cấu trúc của mạng này bao gồm các "đầu vào" là X là một tập của các giá trị $(x_1, x_n, ..., x_n)$. Giống như mạng neuron sinh học, mỗi giá trị này cũng được nối với một neuron bằng một liên kết có mang trọng số tương ứng với các giá trị x_i là $(w_1, w, ..., w_n)$. Giá trị đầu vào chung cho neuron sẽ là tổng các giá trị đầu vào nhân với trọng số tương ứng:

$$Sum = \sum_{1}^{n} x_i w_i + b \tag{1.1}$$

Trong đó b được gọi là giá trị bias.

Sau đó, giá trị đầu ra của neuron sẽ bằng:

$$Output = f(Sum) \tag{1.2}$$

Trong đó, *f* được gọi là hàm kích hoạt. Tùy thuộc vào mạng lưới neuron với mục đích sử dụng khác nhau mà có thể có các hàm kích hoạt khác nhau như hàm *Sigmoid* có dạng:



(b) Artificial neural network

Hình 1-8. Nguyên lý hoạt động của một neuron sinh học mô phỏng bằng các neuron nhân

tạo [66]

Hay hàm *tanh* có dạng:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$
(1.4)



Hình 1-9. Mô hình mạng neuron nhân tạo một nút – Perceptron [71]

Kết quả đầu ra của nơ ron sẽ thay đổi tùy thuộc vào trọng số. Bằng cách điều chỉnh trọng số của một neuron nhân tạo, chúng ta có thể có được đầu ra chúng ta muốn cho đầu vào cụ thể. Từ một neuron đơn giản như Hình 1-9, có thể phát triển thành mạng neuron thực hiện những vấn đề phức tạp hơn dựa vào việc kết nối nhiều neuron thành một mạng lưới.

Một mạng lưới có thể có nhiều lớp, mỗi lớp có một ma trận W, một ngưỡng *b* và một véctơ đầu ra *a*. Thông thường giá trị ra của một lớp là giá trị vào của lớp tiếp theo. Mỗi lớp trong mạng đảm nhiệm vai trò khác nhau, lớp cho kết quả ở đầu ra của mạng được gọi là lớp ngõ ra *_ output layer*, tất cả các lớp còn lại được gọi là lớp ẩn *_ hidden layers*. Mạng đa lớp có khả năng xử lý rất lớn. Cấu trúc một mạng có thể mô tả bằng số lớp, số neuron trong một lớp, hàm truyền của mỗi lớp và kết nối giữa các lớp. Tùy thuộc vấn đề mạng cần giải quyết mà có cấu trúc khác nhau.

1.2.4 Phân loại mạng neuron

Có nhiều loại mạng khác nhau và cũng có nhiều cách để phân loại mạng neuron [19].

- Dựa vào số lớp có trong mạng neuron ta có thể phân loại thành: mạng neuron một lớp, mạng neuron nhiều lớp.

- Dựa vào đường truyền tín hiệu trong mạng neuron ta phân loại thành: mạng neuron truyền thẳng, mạng neuron phản hồi, mạng neuron tự tổ chức.

Mạng neuron truyền thẳng: Mạng neuron truyền thẳng là mạng hai hay nhiều lớp mà tín hiệu truyền theo một hướng từ đầu vào đến đầu ra.



Hình 1-10. Mô hình mạng neuron nhiều lớp [83]

Mạng neuron phản hồi: Mạng neuron phản hồi là mạng mà trong đó một hoặc nhiều đầu ra của các phần tử lớp sau truyền tín hiệu ngược lại tới đầu vào của lớp trước. Một kiểu phân loại điển hình được biểu diễn như Hình 1-11.

Ứng với các nhóm mạng neuron khác nhau thường áp dụng một số luật học



Hình 1-11. Phân loại mạng neuron nhân tạo [21]

nhất định. Nếu tồn tại hàng chục loại mạng neuron khác nhau thì các luật học dùng

trong mạng neuron có thể liệt kê gấp nhiều lần. Nếu coi cấu trúc mô hình mạng là phần thể xác thì các luật học là phần trí tuệ thông minh của mạng neuron và các công trình nghiên cứu luật học chiếm số lượng lớn nhất trong mấy chục năm qua.

Đối với mạng neuron phản hồi thường sử dụng luật Hebb và các luật cải tiến của nó để chỉnh trọng số mà không cần tín hiệu chỉ đạo từ bên ngoài. Đối với mạng neuron truyền thẳng thường sử dụng luật lan truyền ngược (back-propagation) để chỉnh trọng số với tín hiệu chỉ đạo từ bên ngoài.

1.2.5 Đặc điểm của mạng neuron nhân tạo

Mạng neuron nhân tạo không tiếp cận đến sự phức tạp của bộ não. Nhưng đã có hai sự tương quan cơ bản giữa mạng neuron nhân tạo và sinh học. Thứ nhất, cấu trúc khối tạo thành chúng đều là các thiết bị tính toán đơn giản (mạng neuron nhân tạo đơn giản hơn nhiều) được liên kết chặt chẽ với nhau (mạng neuron đôi khi được xem như là các mô hình liên kết (connectionist models), là các mô hình phân bố song song (parallel-distributed models)). Thứ hai, các liên kết giữa các neuron quyết định chức năng của mạng.

Một số những đặc trưng ưu việt mà mạng neuron có thể thu được từ việc mô phỏng bộ não con người như sau: [17]

- Tính chất phi tuyến

Một neuron có thể tính toán một cách phi tuyến hay tuyến tính. Nếu một mạng neuron được cấu thành bởi sự kết nối các neuron phi tuyến thì nó sẽ có tính phi tuyến. Và tính phi tuyến này sẽ được phân tán trên toàn mạng. Đây là một đặc tính rất quan trọng vì các cơ chế vật lý sinh ra các tín hiệu đầu vào thường là phi tuyến (ví dụ như tín hiệu tiếng nói).

- Tính chất tương ứng đầu vào-đầu ra

Mặc dù khái niệm "học" hay "tích lũy" (training) chưa được bàn đến nhưng để hiểu được mối quan hệ đầu vào-đầu ra của mạng neuron, chúng ta sẽ đề cập sơ qua về khái niệm này. Một mô hình học phổ biến được gọi là học với một người dạy hay học có giám sát liên quan đến việc thay đổi các trọng số synapse của mạng neuron bằng việc áp dụng một tập hợp các mẫu. Như vậy, mạng neuron học từ các ví dụ bằng cách xây dựng nên một quan hệ đầu vào-đầu ra cho vấn đề cần giải quyết. Lấy ví dụ phân loại mẫu, trong đó với mỗi số liệu "đầu vào" sẽ cho một kết quả "đầu ra". Các mạng neuron sẽ có cơ chế tự học sau khi được huấn luyện bằng một tập mẫu gồm cả thông tin "đầu vào" và "đầu ra" biết trước.

- Tính chất thích nghi

Các mạng neuron có khả năng thích nghi khi môi trường thay đổi. Tuy nhiên, mạng có khả năng tiếp nhận sự thay đổi khi được huấn luyện thêm bằng tập mẫu mới và sau khi được huấn luyện như vậy thì có thể nhận dạng cả hai loại mẫu.

- Tính chấp nhận sai sót

Một mạng neuron được cài đặt dưới dạng phần cứng, vốn có khả năng chấp nhận lỗi, Ví dụ, nếu một neuron hay các liên kết kết nối của nó bị hỏng, việc nhận dạng lại một mẫu được lưu trữ sẽ suy giảm về chất lượng. Nhưng do bản chất phân tán của thông tin lưu trữ trong mạng neuron nên sự hỏng hóc này cũng sẽ được trải ra trên toàn mạng.

1.2.6 Ứng dụng của mạng neuron nhân tạo

Có thể kể đến một số ứng dụng phổ biến của mạng neuron trong đời sống hiện nay như sau:

- Trong lĩnh vực không gian vũ trụ (Aerospace): Phi công tự động, giả lập đường bay, các hệ thống điều khiển lái máy bay, bộ phát hiện lỗi.

- Chế tạo các bộ điều khiển tự động cho động cơ (Automotive): Các hệ thống dẫn đường tự động cho ô tô, các bộ phân tích hoạt động của xe.

- Trong lĩnh vực ngân hàng (Banking): Bộ đọc séc và các tài liệu, tính tiền của thẻ tín dụng, dự báo chứng khoán, v.v.

- Trong quốc phòng (Defense): Định vị- phát hiện vũ khí, dò mục tiêu, phát hiện đối tượng, nhận dạng nét mặt, các bộ cảm biến thế hệ mới, xử lí ảnh radar, v.v.

 Trong lĩnh vực điện tử (Electronics): Dự đoán mã tuần tự, sơ đồ chip IC, điều khiển tiến trình, phân tích nguyên nhân hỏng chip, nhận dạng tiếng nói, mô hình phi tuyến. - Trong y khoa (Medical): Chẩn đoán, nhận dạng, ra quyết định các liệu pháp điều trị, v.v.

- Trong lĩnh vực giải trí (Entertaiment): Hoạt hình, các hiệu ứng đặc biệt, dự báo thị trường, thiết kế các Game, v.v.

- Trong lĩnh vực tài chính (Financial): Định giá bất động sản, cho vay, kiểm tra tài sản cầm cố, đánh giá mức độ hợp tác, phân tích đường tín dụng, chương trình thương mại qua giấy tờ, phân tích tài chính liên doanh, dự báo tỉ lệ tiền tệ.

 Trong lĩnh vực bảo hiểm (Insurance): Đánh giá việc áp dụng chính sách, tối ưu hoá sản phẩm.

- Trong lĩnh vực trắc địa - bản đồ: Các công tác dự báo, các bài toán về tối ru hóa, v.v.

1.2.7 Mang neuron Hopfield

Hình 1-12. Cấu trúc của một mạng Hopfield [15]

Năm 1982, Hopfield tập hợp một số nghiên cứu trước đó và trình bày phân tích toán học hoàn chỉnh dựa trên các mô hình Ising spin [55] để cho ra đời mạng Hopfield [15].

Các mạng Hopfield sử dụng cả hai quá trình truyền thẳng và phản hồi. Hình 1-12 minh họa một mạng Hopfield hồi quy đơn lớp. Mặc dù về cơ bản là một mạng đơn lớp nhưng nhờ cơ cấu phản hồi mà nó hoạt động hiệu quả như một mạng đa lớp. Trong mạng này, độ trễ trong quá trình phản hồi được đưa ra nhằm đóng vai trò ổn định mạng, điều này mang bản chất tự nhiên giống như độ trễ của các neuron sinh học ghi nhận khoảng cách của các khớp nối và tỉ lệ giới hạn của vòng thần kinh.

Trong đó: N_i : Neuron, D: Nút phân phối, F: Hàm kích hoạt, I_j : Đầu vào bias, w_{ij} : Trọng số.

1.2.8 Ứng dụng mạng neuron Hopfield trong các bài toán tối ưu hóa

Trong quá trình phát triển, mạng neuron Hopfiled đã được ứng dụng thành công trong rất nhiều lĩnh vực, đặc biệt trong bộ nhớ liên kết và trong các bài toán tối ưu. Từ sau công trình của Hopfiled và Tank (1985), mạng neuron Hopfiled đã được sử dụng nhiều vào việc giải bài toán tối ưu tổ hợp [83]. Trong các bài toán tối ưu này, mạng neuron Hopfield sẽ đạt tới trạng thái cân bằng khi hàm năng lượng của nó đạt tới giá trị cực tiểu. Từ bài toán cho trước, xây dựng một hàm mục tiêu F nào đó (đã được xử lý các ràng buộc) và đặt F = E (E là hàm năng lượng). Sau đó, tìm mối liên hệ giữa các biến của chúng. Do đó, mạng neuron Hopfiled rất phù hợp với các bài toán tối ưu tổ hợp, đặc biệt là các bài toán tối ưu thuộc lớp bài toán NP-đầy đủ như: bài toán tìm đường đi tối ưu cho tuyến đường xe chạy, bài toán người bán hàng, bài toán người đưa thư, bài toán lập thời khóa biểu,...

Có thể khái quát các bước sau đây trong việc sử dụng mạng neuron Hopfiled để giải các bài toán tối ưu hóa hay còn gọi là ánh xạ các bài toán tối ưu hóa lên mạng neuron [83]:

- Lập sơ đồ biểu diễn để các đầu ra của các neuron có thể giải mã thành các nghiệm có thể của bài toán tối ưu.
- Chọn hàm năng lượng sao cho giá trị cực tiểu của nó ứng với nghiệm "tốt nhất" của bài toán cần ánh xạ.
- Gán giá trị cho các tham số của hàm năng lượng điều này sẽ xác định các trọng số tương đối gán cho các thành phần khác nhau của hàm năng lượng.

- Rút ra phương trình động học của các neuron (tương ứng với việc xác định các trọng số liên kết và đầu vào ngoài).
- 5. Khởi tạo các giá trị đầu vào.
- Chú ý: Không có phương pháp ánh xạ trực tiếp các bài toán tối ưu có ràng buộc lên mạng neuron. Do đó phải thêm vào hàm mục tiêu các điều kiện để tránh các ràng buộc bị phá võ. Khi đó, hàm năng lượng được biểu diễn như tổng của các hàm mục tiêu và hàm điều kiện của bài toán.

Chính vì những lí do trên, mạng neuron Hopfiled đã được lựa chọn để tối ưu hóa sự phụ thuộc không gian trong các bài toán: nhận dạng mục tiêu từ ảnh viễn thám [73], siêu phân giải bản đồ [50] và bài toán tăng độ phân giải không gian, nâng cao độ chính xác của mô hình grid DEM trong luận án này.

1.3 Luận giải về tăng độ phân giải không gian grid DEM

Độ phân giải không gian (spatial resolution) của một tấm ảnh là khoảng cách tối thiểu giữa hai đối tượng mà chúng được phân chia và tách biệt với nhau trên ảnh [84].

Độ phân giải không gian bị giới hạn khoảng cách lấy mẫu trên mặt đất vì không thể phân tách được một đối tượng trong phạm vi một pixel.

Do đó, độ phân giải không gian của DEM dạng grid được thể hiện bằng trong một diện tích vuông có một giá trị độ cao. Ví dụ như grid DEM có độ phân giải là 30m/pixel có nghĩa là một điểm ảnh chứa một giá trị độ cao đại diện cho độ cao của tất cả các điểm nằm trong một ô vuông với kích thước 30m×30m trên mặt đất.

Độ phân giải không gian của mô hình số độ cao (DEM) thể hiện khả năng thể hiện địa hình với độ chi tiết cao hay thấp. Độ phân giải không gian có vai trò quan trọng tới độ chính xác của phân tích không gian và một số dữ liệu có được từ DEM như độ dốc, độ dốc theo hướng, các mô hình dòng chảy, v.v... Độ phân giải không gian của DEM dạng grid ảnh hưởng đến cả nội dung thông tin, tính chính xác của dữ liệu của các sản phẩm này. Một loạt nghiên cứu của các tác giả khác nhau đã chỉ ra tác động của độ phân giải không gian của DEM đối với các dữ liệu không gian được phân tích từ kết quả DEM [39, 81], nhất là về độ dốc và hướng dốc [40, 41], phân định ranh giới lưu vực và độ chính xác của các mô hình SWAT [75], các mô hình thoát nước và mô hình ba chiều của cảnh quan [51], và kết quả khảo sát địa chất [68]. Tất cả các nghiên cứu trên đã chỉ ra rằng, các DEM với độ phân giải không gian cao hơn có thể cho các dữ liệu có độ chính xác và tính chi tiết cao hơn.

Tăng độ phân giải hay siêu phân giải (super-resolution) là một thuật ngữ sử dụng cho các phương pháp nhằm tăng cường độ nét của hình ảnh hoặc video. Trong các phương pháp này, thông thường các ảnh chụp liên tiếp với độ phân giải thấp được sử dụng để tạo ra các ảnh có độ phân giải cao hơn. Phương pháp siêu phân giải hoạt động giải quyết sự hơi khác nhau của một số các hình ảnh có độ phân giải thấp của cùng một đối tượng để tạo ra một ảnh có độ phân giải cao hơn. Và như vậy, tổng lượng thông tin về một đối tượng do nhiều tấm ảnh mang lại sẽ cao hơn nhiều so với thông tin trong một tấm ảnh đơn lẻ.

Khái niệm siêu phân giải bản đồ (super-resolution mapping/sub-pixel mapping) lần đầu tiên được giới thiệu bởi Atkinson [30] trong đó mỗi điểm ảnh trên ảnh viễn thám gốc được chia thành nhiều tiểu điểm ảnh (sub-pixel) và số lượng của các tiểu điểm ảnh của các lớp được xác định bằng giá trị phần trăm của lớp phủ được xác định từ kết quả phân loại mềm. Vị trí của các tiểu điểm ảnh được xác định dựa trên nguyên lý về sự phụ thuộc không gian, đó là về nguyên tắc, xu hướng của các điểm ảnh lân cận thường có tính chất giống nhau hơn so với điểm ảnh ở xa. Do đó, các tiểu điểm ảnh của cùng một lớp phủ sẽ được sắp xếp cạnh nhau để tạo ra một bản đồ lớp phủ có mức liên kết không gian là lớn nhất hay sự phụ thuộc không gian giữa các tiểu điểm ảnh là lớn nhất. Một số kỹ thuật tăng độ phân giải không gian đã được phát triển như hoán đổi vị trí các tiểu điểm ảnh [53], trường ngẫu nhiên Markov [54, 65], mang neuron Hopfield (HNN) [73].

1.4 Một số nghiên cứu tiêu biểu về tăng độ phân giải không gian và tăng độ chính xác DEM

Các DEM có độ phân giải không gian cao và độ chính xác cao có thể thu được bằng cách sử dụng công nghệ Lidar hoặc đo đạc mặt đất hoặc bằng phương pháp đo ảnh. Tuy nhiên, đối với một số ứng dụng như xác định lưu vực sông, phân vùng xác định nguy cơ trượt lở, xác định mức độ sói mòn, mô hình thủy văn trên diện rộng, v.v. thì việc đo đạc xác định DEM ở độ chính xác cao bằng phương pháp LiDAR hoặc phương pháp đo vẽ mặt đất, đo vẽ bằng ảnh hàng không cự ly gần là quá tốn kém và chưa khả thi.

Trong trường hợp dữ liêu mô hình số đô cao DEM không cần phải có mức đô chính xác đến dưới 1 m như dữ liêu có được từ LiDAR và dữ liêu đầu vào chủ yếu sử dung các nguồn DEM có sẵn ở đô phân giải thấp và trung bình (từ 20m-90m), một số nghiên cứu sử dụng các phương pháp giảm kích thước pixel (downscale) để tăng độ phân giải bằng một số phương pháp tái chia mẫu (resampling). Các phương pháp tái chia mẫu được sử dụng nhiều nhất cho việc tăng độ phân giải là phương pháp nội suy song tuyến (Bilinear) và Bi-cubic [56]. Các phương pháp khác cũng có thể được sử dụng như tái chia mẫu B-spline và phương pháp lọc được sử dụng trong một phát minh của Atkins và các cộng sự [80] hay phương pháp Kriging [49]. Các phương pháp Kriging khác nhau cũng được sử dụng để tăng độ chính xác của mô hình số độ cao [78], nhóm nghiên cứu đã so sánh việc sử dụng kết hợp hai kỹ thuật địa thống kê: Ordinary Cokriging (OCK) và Kriging with an External (KED) để tăng độ chính xác của mô hình số độ cao. Bằng cách sử dụng một tập hợp các điểm kiểm tra được phân bố rải rác trên mặt đất, kết quả của nghiên cứu đã cho thấy kỹ thuật kết hợp trên đều tạo ra DEM chính xác hơn bất kỳ nguồn dữ liêu nào được sử dung riêng lẻ, trong đó kỹ thuật KED tao ra kết quả chính xác nhất. Cũng theo như các nghiên cứu này, việc chia nhỏ các pixel và xác đinh lai bằng các phương pháp như nôi suy song tuyến Bilinear, Bi-cubic, Kriging có thể phần nào tăng độ chính xác của dữ liệu raster, có nghĩa là có thể nâng cao độ chính xác của DEM dạng grid.

Tuy đã có một số nghiên cứu trên thế giới về việc sử dụng phương pháp tái chia mẫu để nâng cao độ chính xác của DEM dạng grid, việc đánh giá một cách đầy đủ về các thuật toán phổ biến này, bao gồm các thuật toán Bilinear, Bi-cubic, Kriging trong việc nâng cao độ chính xác của một số loại DEM dạng grid cũng như thử nghiệm thuật toán mới cho việc tăng độ chính xác của DEM dạng grid là cần thiết. Mặc khác, thông qua việc đánh giá độ chính xác, có thể bằng các chỉ số khác nhau có thể đánh giá được ưu, nhược điểm của các thuật toán này.

Ý tưởng về việc sử dụng thuật toán tăng độ phân giải không gian để tăng độ chính xác của DEM cũng đã được Mokarrama và Hojati [6] tiến hành nghiên cứu. Các tác giả này đã đề xuất chia pixel thành các pixel nhỏ (sub-pixel) và độ cao của tiểu điểm ảnh (các pixel nhỏ này) được xác định bằng hai phương pháp tính giá trị hấp dẫn là tiệm cận và góc phần tư cho DEM ở phía bắc của Darab, tỉnh Fars, Iran. Trong mô hình này, các điểm ảnh khác ở xa điểm ảnh trung tâm có thể nhân bất kỳ giá tri nào. Thiết kế của thuật toán được thử nghiêm với ảnh sử dụng DEM ASTER với độ phân giải không gian là 30m và DEM SRTM 90m. Trong mô hình này, các hệ số biến đổi là S = 2, S = 3, và S = 4) với hai phương pháp tiếp xúc lân cận (T = 1) và góc phần tư (T = 2) được lập trình tính toán trên phần mềm MATLAB. Thuật toán đánh giá độ chính xác bằng cách tính toán các sai số RMSE đã chứng minh được qua 487 mẫu thực nghiệm rằng mô hình thuật này đã tăng được độ phân giải không gian và cả độ chính xác của DEM kết quả, tuy nhiên mức độ tăng độ chính xác thể hiện ở sai số trung phương giảm rất ít, nguyên nhân là do phương pháp sử dụng để tính giá trị tương quan về độ cao giữa các điểm lân cận thông qua giá trị hấp dẫn (attraction) không phải là phương pháp tối ưu. Mặt khác, mục đích chính của các tác giả là sử dụng DEM cho mục đích phân loại các dạng địa hình nên việc đánh giá độ chính xác của DEM cũng không được quan tâm.

1.5 Kết luận chương 1

Trong chương này đã giới thiệu tổng quan về DEM và về mạng neuron. Trong đó, phần tổng quan về DEM gồm: các khái niệm và định nghĩa về DEM, các cấu trúc của DEM, các phương pháp thành lập, độ chính xác bề mặt DEM, các ứng dụng của DEM, các nghiên cứu trong và ngoài nước về DEM hiện nay. Đặc biệt, luận án đã giới thiệu một số các nghiên cứu tiêu biểu về cải thiện, đánh giá độ chính xác DEM. Trong phần tổng quan về mạng neuron, luận án đã giới thiệu tổng quan về cấu tạo, nguyên lý hoạt động của neuron sinh học và cấu trúc, đặc điểm, nguyên lý hoạt động, phân loại, ứng dụng của mạng neuron nhân tạo. Bên cạnh đó, luận án cũng đã giới thiệu một cách tổng quan về nguyên lý ứng dụng mạng neuron nhân tạo Hopfiled trong thuật toán tăng độ phân giải không gian và độ chính xác của mô hình grid DEM.

Trên cơ sở những vấn đề đã được nghiên cứu nhưng chưa được giải quyết triệt để, trong luận án này đề xuất các nội dung nghiên cứu như sau:

- Khảo sát về khả năng tăng độ chính xác của grid DEM bằng các thuật toán tái chia mẫu phổ biến như: Bilinear, Bi-cubic và Kriging.
- Đề xuất, thử nghiệm và đánh giá độ chính xác việc sử dụng thuật toán tăng độ phân giải và độ chính xác của DEM bằng mạng neuron Hopfield bằng các hàm mục tiêu và điều kiện.

Chương 2 KHẢO SÁT VỀ KHẢ NĂNG TĂNG ĐỘ CHÍNH XÁC CỦA GRID DEM BẰNG CÁC THUẬT TOÁN TÁI CHIA MẫU

Trong xử lý dữ liệu raster, việc tái chia mẫu (resampling) được sử dụng khá nhiều trong những trường hợp cần dữ liệu có kích thước pixel nhỏ hơn hoặc lớn hơn kích thước pixel gốc. Việc tái chia mẫu được kỳ vọng sẽ cải thiện về độ chính xác/mức độ chi tiết của dữ liệu. Các thuật toán tái chia mẫu phổ biến nhất được sử dụng là Bilinear và Bi-cubic. Đối với dữ liệu mô hình số độ cao DEM dạng grid, việc tái chia mẫu được sử dụng khá nhiều và đã có các đánh giá cho thấy tái chia mẫu bằng các phương pháp như Bilinear, Bi-cubic và Kringing có thể nâng cao độ chính xác của mô hình số độ cao. Đặc biệt là phương pháp Kriging đã cho phép tạo ra DEM với độ chính xác cao hơn DEM gốc khi so sánh giá trị sai số trung phương RMSE (Root Mean Square Error) [56, 49].

Để đánh giá một cách đầy đủ, toàn diện hơn về khả năng nâng cao độ chính xác mô hình số độ cao DEM dạng grid bằng các thuật toán tái chia mẫu, trong chương này đề xuất các phương pháp đánh giá độ chính xác các grid DEM theo một cách tiếp cận mới, sau đó sẽ áp dụng các phương pháp này để tiến hành đánh giá độ chính xác của các phương pháp tái chia mẫu được sử dụng hiện nay để đánh giá khả năng nâng cao độ chính xác bằng tái chia mẫu một cách toàn diện.

Việc đánh giá độ chính xác của dữ liệu bao gồm các phương pháp đánh giá trực quan và đánh giá định lượng. Các phương pháp trực quan được thực hiện bằng cách so sánh trực tiếp và so sánh biểu đồ phân tán của dữ liệu độ cao grid. Đánh giá định lượng thì ngoài đại lượng được sử dụng nhiều để đánh giá độ chính xác là sai số trung phương, trong chương này đề xuất sử dụng các tham số m, b của đường phân bố khi sử dụng hồi quy tuyến tính và hệ số tương quan R giữa các tập dữ liệu DEM chuẩn và DEM được tái chia mẫu.

2.1 Các phương pháp đánh giá độ chính xác của grid DEM

Người ta nhận thấy rằng độ chính xác của phép nội suy không gian của DEM chủ yếu phụ thuộc vào mật độ điểm và sự phân bố của dữ liệu đầu vào, kích thước

lưới (độ phân giải), độ phức tạp của địa hình và thuật toán nội suy được sử dụng [32].

Các biến trong các thông số nội suy có thể cải thiện đáng kể hoặc làm giảm đi độ chính xác. Vì vậy, cần phải kiểm tra thêm sự tác động của độ phức tạp địa hình và các phương pháp lấy mẫu dữ liệu khác nhau. Việc sử dụng kết hợp các mô hình variogram, độ chính xác của các bề mặt và sự tương quan không gian sẽ làm cho việc đánh giá độ chính xác của dữ liệu grid DEM được chính xác hơn [32].

Việc đánh giá độ chính xác của các dữ liệu grid DEM được thực hiện theo cả các cách đánh giá trực quan và các phương pháp đánh giá định lượng.

2.1.1 Phương pháp đánh giá trực quan

2.1.1.1 Sử dụng phương pháp so sánh trực tiếp

Phương pháp này so sánh hai ảnh của hai bộ dữ liệu DEM trực tiếp bằng mắt thường để thấy sự giống nhau hoặc sự sai khác, chênh lệch nếu có.

2.1.1.2 Sử dụng phương pháp mặt cắt

So sánh hai bề mặt DEM dựa vào mặt cắt: Trong phương pháp này, dựa trên giá trị các điểm độ cao của các bộ dữ liệu DEM, chúng ta tính toán và vẽ các mặt cắt dọc, các mặt cắt ngang tương ứng của các dữ liệu DEM là kết quả của tái chia mẫu và dữ liệu DEM mẫu ở cùng một độ phân giải. Sau đó, tiến hành so sánh giữa các mặt cắt tương ứng đó. Nếu các mặt cắt của DEM kết quả của các phương pháp tái chia mẫu càng gần hoặc sát với mặt cắt của DEM mẫu thì bề mặt DEM đó càng gần với bề mặt DEM mẫu (DEM tham chiếu), có nghĩa là dữ liệu DEM đó có độ chính xác càng cao (có độ sai lệch ít so với DEM mẫu).

Khi các mặt cắt của DEM kết quả sau khi tái chia mẫu trùng với mặt cắt của DEM mẫu thì bề mặt của DEM theo các phương pháp tái chia mẫu hoàn toàn trùng khóp với bề mặt của DEM tham khảo (DEM mẫu).

2.1.1.3 So sánh bằng biểu đồ phân tán

Từ các dữ liệu điểm độ cao của các bộ dữ liệu DEM, chúng ta đi xây dựng các biểu đồ phân tán (Scatter diagram) của các bộ dữ liệu này. Sau đó so sánh hai bề mặt DEM bằng biểu đồ phân tán.

Biểu đồ phân tán là một đồ thị biểu thị mối tương quan giữa bề mặt DEM kết quả của các phương pháp tái chia mẫu và bề mặt DEM tham khảo. Trong đó, trục tung Y thể hiện giá trị độ cao của DEM sau khi tái chia mẫu, trục hoành X là giá trị độ cao của DEM tham khảo (DEM mẫu).

Trong các biểu đồ phân tán này, nếu các điểm trên biểu đồ phân tán càng nằm sát đường hồi quy thì hai bề mặt DEM sẽ càng gần giống nhau, còn nếu các điểm nằm xa đường hồi quy thì hai bề mặt DEM không khớp nhau.

2.1.2 Phương pháp đánh giá định lượng

Để định lượng và phân tích độ lệch giữa hai bộ dữ liệu độ cao (bộ dữ liệu DEM mới được xây dựng theo các phương pháp tái chia mẫu và bộ dữ liệu DEM mẫu), các cách thông thường là tính sai số trung phương dựa trên các giá trị độ cao của các bộ dữ liệu DEM và phương pháp sử dụng các tham số thống kê để đánh giá độ chính xác của dữ liệu DEM (đánh giá độ chính xác không gian của các bộ dữ liệu).

2.1.2.1 Sử dụng giá trị sai số trung phương

Sai số trung phương là đại lượng được sử dụng nhiều nhất để đánh giá độ chính xác trong DEM. Sai số trung phương được trình bày trong công thức 1.1.

2.1.2.2 Sử dụng các giá trị thống kê R (Hệ số tương quan) và phương trình hồi quy được đại diện bằng 2 tham số m và b

Hệ số tương quan là thước đo về giá trị bằng số của một số loại tương quan giữa hai biến, nghĩa là mối quan hệ thống kê giữa hai biến. Các biến có thể là hai cột của một tập hợp dữ liệu quan sát nhất định, thường được gọi là mẫu hoặc hai thành phần của ngẫu nhiên đa biến với một phân phối đã biết. Các giá trị của hệ số tương quan nằm trong khoảng -1.0 đến 1.0. Giá trị chính xác là 1 có nghĩa là có một mối quan hệ tích cực hoàn hảo giữa hai biến. Đối với sự gia tăng tích cực trong một biến, cũng có sự gia tăng tích cực trong biến thứ hai. Giá trị -1.0 có nghĩa là có một

mối quan hệ phủ định hoàn hảo giữa hai biến. Điều này cho thấy các biến di chuyển theo hướng ngược lại đối với sự gia tăng tích cực trong một biến, có sự giảm trong biến thứ hai. Nếu tương quan giữa hai biến là 0, nghĩa là không có mối quan hệ tuyến tính giữa chúng [46].

Có một số loại hệ số tương quan, nhưng loại phổ biến nhất là tương quan Pearson (*R*). Hệ số này dùng để tính toán giá trị và hướng của mối quan hệ tuyến tính giữa hai biến. Nó không thể sử dụng để đánh giá các mối quan hệ phi tuyến giữa hai biến và không thể phân biệt giữa các biến phụ thuộc và biến độc lập.

Trong toán học, các mô hình hồi quy là một kỹ thuật thống kê dùng để ước lượng phương trình phù hợp nhất với một tập hợp các kết quả quan sát của biến phụ thuộc và biến độc lập. Nó cho phép đạt được kết quả ước lượng tốt nhất về mối quan hệ chân thực giữa các biến số. Từ phương trình ước lượng được này, người ta có thể dự báo về biến phụ thuộc (chưa biết) dựa vào giá trị cho trước của biến độc lập (đã biết).

Hình 2-1 biểu thị tập hợp các kết quả quan sát của một mô hình hồi quy dưới dạng đồ thị và chúng ta phải tìm phương trình của đường thẳng trên đồ thị với điều kiện nó phù hợp nhất với số lượng mẫu mà chúng ta thu thập được, vì một đường như vậy sẽ đem lại kết quả dự báo tốt nhất cho biến phụ thuộc. Đường thẳng phù hợp nhất với số liệu phải được lựa chọn sao cho giá trị của tổng bình phương các độ lệch (khoảng cách) theo phương thẳng đứng giữa các điểm và đường thẳng là nhỏ nhất. Phương pháp số bình phương nhỏ nhất thông thường này được ứng dụng trong hầu hết các phân tích hồi quy. Tính phù hợp của đường hồi quy với các kết quả quan sát mẫu được phản ánh bằng hệ số tương quan R. Ở Hình 2-1, đường màu đỏ là đường mà mô hình ước lượng được.

Trong nội dung luận án này, để đánh giá kết quả của các phương pháp khác nhau, các mô hình hồi quy tuyến tính đã được gắn vào mối quan hệ giữa dữ liệu tham chiếu và dữ liệu được tái chia mẫu. Sự tương tự nhau của hai loại DEM cũng có thể được đánh giá định lượng bằng cách sử dụng các hệ số hồi quy tuyến tính (m, b) và hệ số tương quan R.

Hệ số tương quan R được tính theo công thức (2.1) [32]:

$$R = \frac{[(Z_{di} - Z_{ri})^2]}{n.m^2}$$
(2.1)

Trong đó:

R là hệ số tương quan; Z_{di} là giá trị độ cao thứ i trên bề mặt DEM kết quả của phương pháp tái chia mẫu; Z_{ri} là giá trị độ cao thứ i trên bề mặt DEM tham khảo; n là số lượng điểm độ cao kiểm tra.



Hình 2-1. Mô hình đường hồi quy tổng quát [62]

Khi so sánh hai DEM (dữ liệu DEM kết quả sau khi tái chia mẫu và dữ liệu DEM tham chiếu – DEM mẫu), nếu gọi độ cao của một pixel (*i*, *j*) trong tập dữ liệu tham chiếu là y và độ cao của pixel tương ứng ở vị trí (*i*, *j*) trong tập dữ liệu chuẩn sử dụng làm mẫu là x, nếu như hai tập dữ liệu độ cao này giống nhau hoàn toàn thì giá trị $y_{(i, j)} = x_{(i, j)}$ với mọi (*i*, *j*). Nếu coi giá trị độ cao tham chiếu y là một hàm của x, vẽ đồ thị phân tán với một trục tọa độ là x và trục kia là giá trị y tương trừng trung trục tộa độ thị phân tán sẽ nằm trên đường thẳng y = x, gọi là đường trùng khớp tuyệt đối, hay đại lượng y và x có sự tương quan thuận tuyệt đối (Hình 2-2) [14].

Tuy nhiên, khi so sánh giữa các DEM với mọi giá trị x và y thì việc trùng khớp tuyệt đối không thể xảy ra, khi đó người ta sử dụng đường hồi quy y = mx + b



Hình 2-2. Đường trùng khóp tuyệt đối theo tương quan thuận [14]

để đặc trưng cho xu thế biến thiên của *y* theo *x*. Nếu các tham số hồi quy m = 1 và b = 0 thì giữa DEM chuẩn và DEM tái chia mẫu có sự tương quan thuận tuyệt đối. Tuy nhiên do giá trị độ cao trên DEM cần đánh giá sẽ không thể trùng khóp với giá trị độ cao trên DEM chuẩn nên $y \neq x$ và sẽ có giá trị y = mx + b với $m \neq 1$ và $b \neq 0$. Dựa vào các giá trị m và b có thể xác định được thành phần sai số hệ thống trong các DEM. Giả sử b > 0 có thể thấy DEM tái chia mẫu sẽ có xu hướng thấp hơn mặt DEM chuẩn một giá trị b, người lại b < 0 thì DEM tái chia mẫu sẽ có xu hướng cao hơn DEM chuẩn một giá trị b. Tương tự khi m > 1 thì DEM tái chia mẫu có xu thế thấp hơn DEM chuẩn và m < 1 thì DEM tái chia mẫu có xu thế cao hơn DEM tái chia mẫu được đánh giá có sai số hệ thống nhỏ khi m càng gần giá trị bằng 1 và b càng cần giá trị bằng 0. Như vậy, hai thành phần m và b thể hiện sự xuất hiện của sai số hệ thống trong dữ liệu cần đánh giá độ chính xác.

Hệ số tương quan *R* được sử dụng để đo sự liên kết giữa hai bộ dữ liệu, do đó sẽ đánh giá được sự phân phối các điểm dữ liệu trong các biểu đồ phân tán xung quanh đường hồi quy. Hình 2-3 cho thấy tính phân tán của các điểm trên đồ thị phân tán tương ứng với mỗi giá trị R^2 . Giá trị của R^2 sẽ đặc trưng cho thành phần sai số ngẫu nhiên khi so sánh DEM. Khi giá trị của R^2 càng gần đến 1 thì càng có nhiều điểm dữ liệu nằm gần đường hồi quy, nghĩa là sai số ngẫu nhiên có giá trị nhỏ. Nếu $R^2 = 1$ thì thành phần sai số ngẫu nhiên = 0. Một sự trùng khớp tuyệt đối giữa hai bộ dữ liệu DEM là tương ứng với việc cả hai thành phần sai số ngẫu nhiên và sai số hệ thống đều nhỏ, có nghĩa là khi tất cả các điểm dữ liệu được đặt sát với đường khớp tuyệt đối (y = x) và hệ số xác định R^2 càng gần 1. Điều đó có nghĩa là hai bộ dữ liệu sẽ giống hệt nhau nếu các giá trị m = 1, b = 0 và $R^2 = 1$ đồng thời xuất hiện.



Hình 2-3. Biểu đồ phân tán với các giá trị tương quan khác nhau tương ứng với thành phần sai số ngẫu nhiên lớn hay nhỏ [12]

2.2 Một số thuật toán tái chia mẫu (Resampling) phổ biến nhằm tăng độ phân giải không gian cho DEM dạng grid

Tái chia mẫu (resampling) là thuật ngữ đề cập đến việc tăng dày số lượng mẫu của một tập mẫu với mục tiêu tạo ra tập mẫu mới có độ chính xác và độ chi tiết cao hơn. Nếu coi một mô hình số độ cao là một tập mẫu gồm các điểm ảnh (pixel) mang giá trị độ cao đại diện cho ô vuông mà nó bao trọn trên mặt đất, thì việc tái chia mẫu là chia nhỏ các điểm ảnh này thành các điểm ảnh có kích thước nhỏ hơn với mong muốn các điểm ảnh mới sẽ biểu diễn tốt bề mặt địa hình một cách chi tiết hơn và sát thực tế hơn các điểm ảnh cũ có độ phân giải thấp. Có nhiều phương pháp tái chia mẫu khác nhau. Trong các phương pháp tái chia mẫu, độ cao của các điểm ảnh nhỏ được xác định bằng kỹ thuật nội suy. Phương pháp nội suy có thể là phương pháp trung bình có trọng số từ các pixel xung quanh. Trọng số phụ thuộc vào khoảng cách giữa vị trí của pixel mới và các pixel lân cận. Phương pháp tái chia mẫu đơn giản nhất là phương pháp Dựa vào điểm lân cận gần nhất (Nearest Neighbor); một số các phương pháp khác cho độ chính xác cao hơn bằng việc xét thêm các điểm dữ liệu xung quanh để tạo ra kết quả chính xác hơn.

Để đánh giá hiệu quả của phương pháp tăng độ phân giải không gian của DEM khi sử dụng mô hình HNN, chúng tôi cũng đã sử dụng ba phương pháp tái chia mẫu phổ biến để tăng độ phân giải không gian của DEM là phương pháp song tuyến (Bilinear), Bi-cubic và phương pháp Kriging. Sau đó tiến hành thực nghiệm tất cả các phương pháp trên trên cùng hai bộ dữ liệu là bộ dữ liệu DEM đã làm suy giảm độ phân giải và bộ dữ liệu DEM mẫu ở độ phân giải không gian 20m, 60m và 90m.

2.2.1 Phương pháp tái chia mẫu Bilinear (song tuyến)

Trong toán học, phép nội suy song tuyến là một phần mở rộng của phép nội suy tuyến tính để nội suy các hàm có hai biến (ví dụ: x và y) trên lưới mặt phẳng 2D. Nội suy song tuyến được thực hiện bằng cách sử dụng phép nội suy tuyến tính theo một hướng trước, sau đó là theo hướng còn lại. Mặc dù mỗi bước nội suy theo các hướng là tuyến tính nhưng khi nội suy theo hai hướng thì phép toán nội suy lại không phải là tuyến tính mà là bậc hai.

Giả sử cần xác định một giá trị của một hàm *f* chưa biết tại điểm (*x*, *y*), ví dụ giá trị độ cao của điểm (*x*, *y*), từ bốn điểm $Q_{11} = (x_1, y_1), Q_{12} = (x_1, y_2), Q_{21} = (x_2, y_1)$ và $Q_{22} = (x_2, y_2)$ (Hình 2-5) [33].

Để tính giá trị của điểm nội suy (x, y):

- Đầu tiên phải tính toán nội suy theo hướng x theo công thức (2.2):

$$f(x, y_1) = \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}),$$

$$f(x, y_2) = \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22})$$
(2.2)
Sau đó tính toán nội suy theo hướng y ta được giá trị của điểm (x, y) theo công thức (2.3):

$$f(x,y) = \frac{1}{(x_2 - x_1)(y_2 - y_1)} \begin{bmatrix} x_2 - x & x - x_1 \end{bmatrix} \begin{bmatrix} f(Q_{11}) & f(Q_{11}) \\ f(Q_{11}) & f(Q_{11}) \end{bmatrix} \begin{bmatrix} y_2 - y \\ y - y_1 \end{bmatrix}$$
(2.3)

Hình 2-5. Tái chia mẫu song tuyến để ước tính giá trị f của điểm P (x, y) từ điểm Q11 = (x1, y1), Q12 = (x1, y2), Q21 = (x2, y1) và Q22 = (x2, y2). Bốn điểm màu đỏ là các điểm dữ liệu để nội suy và điểm màu xanh lá cây P là điểm cần nội suy [33]



Hình 2-4. Ví dụ minh họa về nội suy song tuyến [7]

* Ví dụ minh họa về nội suy song tuyến (Hình 2-4):

- Tính toán nội suy theo hướng x:

$$f(20,14.5) = \frac{15-14.5}{15-14} \cdot 91 + \frac{14.5-14}{15-14} \cdot 210 = 150.5$$
$$f(21,14.5) = \frac{15-14.5}{15-14} \cdot 162 + \frac{14.5-14}{15-14} \cdot 950 = 128.5$$

- Tính toán nội suy theo hướng y ta được giá trị của điểm cần nội suy là:

$$f(20.2,14.5) = \frac{21-20.2}{21-20} \cdot 150.5 + \frac{20.2-20}{21-20} \cdot 128.5 = 146.1$$

2.2.2 Phương pháp nội suy dựa vào điểm lân cận gần nhất (Nearest Neighbor)

Phương pháp nội suy dựa vào điểm lân cận gần nhất (Nearest Neighbor) là một phương pháp nội suy rất đơn giản. Thuật toán nội suy điểm lân cận gần nhất sẽ chọn giá trị điểm nội suy là giá trị của điểm gần nhất, hoàn toàn không xem xét giá trị của các điểm lân cận khác để tính toán nội suy. Vì các điểm ảnh liền kề có xu hướng có cùng mức cường độ nên mỗi pixel đầu ra của thuật toán nội suy này sẽ được gán giá trị của một pixel gần nhất trong ảnh đầu vào. Hàng xóm gần nhất của nó có thể là bất kỳ một trong các pixel trên, dưới, trái và phải, ví dụ điểm P(x, y) trên trục hoành và giá trị của pixel P(x, y+1) lân cận của nó trên trục tung [7].

Hình 2-6 minh họa về các điểm được nội suy theo thuật toán điểm lân cận gần nhất (Nearest Neighbor) [7]. Trong đó, các ô vuông màu trắng là các điểm sẽ được nội suy.



Hình 2-6. Ví dụ về phương pháp nội suy Nearest Neighbor [7]

Ưu điểm

Kỹ thuật nội suy dựa vào điểm lân cận gần nhất (Nearest Neighbor) rất đơn giản trong tính toán.

Nhược điểm

Có các biến dạng nội suy như làm mờ ảnh và có các quầng sáng ở cạnh.

2.2.3 Phương pháp tái chia mẫu Bi-cubic

Trong xử lý hình ảnh, phương pháp nội suy Bi-cubic thường được chọn thay cho phương pháp nội suy song tuyến (Bilinear) hoặc phương pháp nội suy dựa vào điểm lân cận gần nhất (Nearest neighbor) khi tốc độ tính toán không phải là vấn đề. Trong khi ở phương pháp nội suy song tuyến chỉ xét đến 4 pixel (2x2), thì trong tính toán nội suy Bi-cubic tính đến 16 pixel (4x4). Do đó, hình ảnh sau nôi suy bằng phương pháp Bi-cubic thường mượt mà hơn và ít biến dạng hơn [7].

Giả sử giá trị hàm f có các dẫn xuất là f_{x, f_y} , và f_{xy} được biết đến ở bốn góc (0,0), (1,0), (0,1) (1,1) của hình vuông đơn vị. Bề mặt nội suy sau đó có thể được viết:

$$P(x,y) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij} x^{i} x^{j}$$
(2.4)

Trong bài toán nội suy Bi-cubic, sau khi xác định được 16 hệ số a_{ij} , chúng ta thay các giá trị đó vào hàm P(x, y) để tạo ra bốn phương trình:

1.
$$f(0,0) = P(0,0) = a_{00}$$

2. $f(1,0) = P(1,0) = a_{00} + a_{10} + a_{20} + a_{30}$
3. $f(0,1) = P(0,1) = a_{00} + a_{01} + a_{02} + a_{03}$
4. $f(1,1) = P(1,1) = \sum_{i=0}^{3} \sum_{j=0}^{3} a_{ij}$

Ưu điểm

Nội suy Bi-cubic thường được chọn thay vì nội suy song tuyến (Bilinear) hoặc phương pháp nôi suy dựa vào điểm lân cận gần nhất (Nearest neighbor) khi tốc độ tính toán không phải là một vấn đề, vì nó cho kết quả có các biến dạng nội suy ít hơn.

Nhược điểm

Phương pháp nội suy Bi-cubic thường tính toán phức tạp hơn nên cần nhiều thời gian để tạo đầu ra hơn so với hai phương pháp nội suy song tuyến (Bilinear) hoặc phương pháp nôi suy dựa vào điểm lân cận gần nhất (Nearest neighbor) đã mô tả ở trên. Mặc khác, tùy thuộc vào quy luật phân bố của các điểm trong khu vực mà nội suy song tuyến hay nội suy Bi-cubic sẽ cho kết quả tốt hơn. Trong trường hợp các điểm có mức độ tương quan thấp, nghĩa là các điểm ở xa có mức độ ảnh hưởng không đáng kể so với các điểm ở gần thì việc sử dụng Bi-cubic sẽ gây ra các biến dạng cho kết quả nội suy.

2.2.4 Phương pháp nội suy Kriging

Kriging là một kỹ thuật nội suy địa lý khi xét cả khoảng cách và mức độ biến đổi giữa các điểm dữ liệu đã biết để ước tính giá trị các điểm ở các khu vực không xác định. Giá trị nội suy được chính là một hàm tuyến tính có trọng số của các giá trị điểm đã biết được sử dụng để nội suy (thường là các điểm xung quanh điểm cần nội suy).

Bản chất của phương pháp nội suy Kriging là dự đoán giá trị của hàm tại một điểm nhất định bằng cách tính trung bình trọng số của các điểm đã biết nằm trong vùng lân cận của điểm nội suy.

Phương pháp nội suy Kriging gần giống với phương pháp tỷ lệ nghịch khoảng cách. Nhưng trong phương pháp tỷ lệ nghịch khoảng cách, trọng số chỉ phụ thuộc vào khoảng cách tới vị trí của điểm nội suy theo một hàm đơn giản, còn trong phương pháp nội suy Kriging, các trọng số được tính toán không chỉ dựa vào khoảng cách giữa các điểm đo và vị trí của điểm nội suy mà còn dựa vào mối quan hệ không gian giữa các điểm đo xung quanh điểm nội suy (các tương quan không gian phải được định lượng).

Phép nội suy Kriging thường gồm hai bước chính là:

- Bước 1: Tìm ra sự tương quan không gian của các điểm dữ liệu và tính toán định lượng sự tương quan này thông qua các giá trị phương sai (variance) và hiệp phương sai (covariance), hoặc bán phương sai (semivariance).

- Bước 2: Tính toán nội suy dữ liệu.

Kriging sử dụng hàm trọng số khác nhau tùy thuộc vào cả khoảng cách và hướng địa lý của điểm mẫu đến nút được tính toán. Vấn đề là người dùng không thể, ngay từ cái nhìn đầu tiên, có thể biết chính xác cách thức một bộ dữ liệu thay đổi từ bất kỳ một vị trí nào liên quan đến khoảng cách và hướng. Tuy nhiên, có



Hình 2-7. Ví dụ về dữ liệu không có phương sai theo chiều ngang nhưng thay đổi rất nhiều dọc theo trục dọc của tập dữ liệu [72]

nhiều kỹ thuật có sẵn để giúp xác định điều này, phổ biến nhất là phân tích phương sai.

Phương pháp Kriging sử dụng một thông số để thể hiện mức độ tương quan giữa các điểm trên một bề mặt được gọi là semi-variogram. Semi-variogram chỉ đơn giản là một nửa phương sai của sự khác biệt giữa tất cả các điểm có thể cách nhau một khoảng cách không đổi. Khác với giá trị hiệp phương sai covariance, giá trị của semi-variogram càng lớn thì độ tương quan không gian của các điểm càng thấp.

Về mặt lý thuyết, semi-variogram ở khoảng cách d = 0 sẽ bằng 0, vì không có sự khác biệt giữa các điểm được so sánh với chính chúng. Tuy nhiên, khi các điểm được so sánh với các điểm ngày càng xa, thì giá trị semi-variogram sẽ tăng lên. Ở một khoảng cách nào đó, được gọi là khoảng bão hòa (range), giá trị semivariance sẽ không thay đổi và các điểm sẽ coi như không có sự tương quan với nhau. Khoảng bão hòa này sẽ xác định vùng lân cận tối đa mà các điểm dữ liệu nằm trong vùng này mới được chọn để nội suy cho điểm nút lưới, dựa trên mối tương quan thống kê giữa các điểm.

Phương pháp này liên quan chặt chẽ về mặt toán học với phương pháp phân tích hồi quy. Phương pháp Kriging nhằm mục đích rút ra được một cách nội suy tuyến tính tốt nhất không sai số hệ thống, dựa trên các giả định về hiệp phương sai, sử dụng lý thuyết của Gauss-Markov để chứng minh tính độc lập của giá trị nội suy và sai số, sử dụng các công thức rất giống nhau.

Kriging dựa trên giá trị semi-variogram như sau:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{1}^{N(h)} \left[Z(x_i) - Z(x_j) \right]^2$$
(2.5)

Trong đó: (*h*) là giá trị variogram, N là số điểm dữ liệu mẫu, h là khoảng cách giữa hai điểm dữ liệu, $Z(x_i)$ và $Z(x_j)$ lần lượt là giá trị dữ liệu của điểm x_i và x_j .

Giá trị semi-variogam đầu tiên được ước tính từ các điểm dữ liệu mẫu và sau đó sử dụng nó để ước tính giá trị dữ liệu của điểm dự đoán dựa trên công thức (2.6):

$$Z(x_0) = \sum_{1}^{n} w_i(x_0) Z(x_i)$$
(2.6)

Trong đó: w_i (x_0) là giá trị trọng số được tính toán dựa trên hàm semi-variogram.

Hiệu quả của phương pháp nội suy Kriging phụ thuộc vào độ chính xác của các tham số semivariogram và mô hình của sự thay đổi của địa hình (tức là, giá trị trung bình thay đổi theo khoảng cách như thế nào) [25].

2.3 Thực nghiệm tăng độ phân giải không gian của mô hình số độ cao DEM dạng grid bằng các thuật toán tái chia mẫu phổ biến

2.3.1 Dữ liệu và thực nghiệm

Hai loại dữ liệu đã được sử dụng để đánh giá độ chính xác của các DEM sau khi tăng độ phân giải bằng các phương pháp tái chia mẫu phổ biến (đã trình bày ở trên) là: các bộ dữ liệu DEM giảm độ phân giải (Degraded DEMs) và các bộ dữ liệu DEM thực (Sampled DEMs).

Quá trình thực nghiệm với dữ liệu giảm độ phân giải được trình bày như Hình 2-8. Các bước thực nghiệm bao gồm:

- Từ một nguồn dữ liệu ban đầu được sử dụng làm DEM chuẩn (giả sử ở độ phân giải 20m), sử dụng phương pháp giảm độ phân giải (image degrading) thành DEM có kích thước pixel là 60m bằng cách xác định độ cao của pixel trên DEM giảm độ phân giải bằng trung bình độ cao của các pixel của DEM chuẩn nằm trong đường bao của pixel này. Chẳng hạn độ cao của một pixel bất kỳ trên DEM giảm độ



DEM đầu vào (60 m)

Hình 2-8. Quá trình thực nghiệm sử dụng dữ liệu DEM giảm độ phân giải

phân giải 60m sẽ bằng giá trị độ cao trung bình tính từ 9 pixel 20m trên DEM chuẩn nằm trong đường bao của pixel này.

- DEM giảm độ phân giải (60m) sẽ được sử dụng làm đầu vào cho các thuật toán Resampling bao gồm song tuyến Bilinear, Bi-cubic và Kriging để tạo ra DEM tái chia mẫu có kích thước pixel là 20m.

- DEM tái chia mẫu sẽ được so sánh với DEM chuẩn ban đầu bằng các phương pháp sử dụng Sai số trung phương, sử dụng hệ số hồi quy tuyến tính và tương quan.

Có thể nói rằng DEM được giảm độ phân giải trong trường hợp này không chứa sai số. Không có sai số trong trường hợp này có nghĩa là các giá trị độ cao của các pixel trong DEM không chứa sai số nội suy và sai số đo khi ta giả thiết bề mặt DEM chuẩn ban đầu là bề mặt trái đất thực. Sự khác biệt giữa các DEM giảm độ phân giải (60m) và DEM chuẩn (20m) là do sự khác nhau về độ phân giải. Hay có thể nói là sai số ảnh hưởng của độ phân giải của DEM dạng grid đến kết quả xác định độ cao.

Chỉ riêng những dữ liệu này có thể đủ để đánh giá khả năng ứng dụng của thuật toán, nhưng nó vẫn có thể dẫn đến sự hoài nghi vì nhiều người cho rằng thực nghiệm không được tiến hành trên bề mặt trái đất thực (DEM thực). Các DEM thực hầu hết được xác định từ tập các điểm độ cao đo, dữ liệu đám mây điểm hoặc từ dữ liệu đường bình độ thay vì được lấy trung bình độ cao của các pixel con nằm trong một pixel gốc. Trên thực tế, độ cao của một pixel của DEM biểu thị độ cao của diện tích bề mặt được bao phủ bởi pixel này. Do đó, nó phải là độ cao trung bình của tất cả các điểm trong phạm vi bề mặt đó.



Hình 2-9. Các bước thực nghiệm với dữ liệu grid DEM thực

Các thuật toán nội suy được sử dụng để nội suy các độ cao của các pixel này từ dữ liệu các điểm độ cao hoặc đường bình độ nên độ cao của pixel trong các grid DEM thực là độ cao trung bình của tất cả các điểm trong đường biên của điểm ảnh (footprint) của pixel này kèm theo sai số đo và sai số nội suy. Để thực hiện việc đánh giá một cách toàn diện hơn các thuật toán, ngoài các dữ liệu có được từ việc giảm độ phân giải, chúng tôi tiến hành thu thập hai tập dữ liệu được tạo ra từ một tập các điểm độ cao đo và đường bình độ. Dữ liệu điểm độ cao và đường bình độ sẽ được chuyển thành dữ liệu DEM dạng grid thông qua các thuật toán nội suy ở 2 độ phân giải khác nhau trong đó một tập các điểm độ cao với mật độ đảm bảo được sử dụng để tạo ra DEM 5m, một tập khác với mật độ thấp hơn được sử dụng để tạo ra DEM có độ phân giải 20m. Để đánh giá độ chính xác cho kết quả của thuật toán, có thể sử dụng thuật toán để tăng kích thước pixel của DEM 20m lên độ phân giải 5m, rồi so sánh với DEM 5m chuẩn. Quá trình đánh giá độ chính xác đối với 2 tập dữ liệu này được trình bày như (Hình 2-9), gồm các bước cụ thể như sau:

 Dữ liệu ban đầu được thu thập bao gồm tập các điểm độ cao hoặc đường bình độ.

- Từ dữ liệu điểm độ cao/đường bình độ nội suy để xác định DEM chuẩn, ví dụ ở độ phân giải 5m. Mặt khác, dữ liệu điểm độ cao/đường bình độ được lược bỏ, ví dụ từ dữ (liệu 1000 điểm độ cao sẽ lược bỏ ngẫu nhiên lấy 500 điểm độ cao hoặc dữ liệu đường bình độ với khoảng cao đều 5m sẽ lược bỏ thành đường bình độ với khoảng cao đều 10m.

 Từ dữ liệu điểm độ cao/đường bình độ được lược bỏ, sử dụng nội suy để tạo ra DEM có độ phân giải thấp hơn, ví dụ DEM 20m.

- Dữ liệu DEM độ phân giải thấp sẽ được tái chia mẫu theo 3 phương pháp song tuyến, Bi-cubic và Kriging để tạo ra DEM kích thước pixel nhỏ (ví dụ 5m).

- Dữ liệu DEM được tái chia mẫu được so sánh với DEM có độ phân giải cao được tạo ra từ dữ liệu độ cao và đường bình độ được lấy làm chuẩn bằng các phương pháp so sánh sử dụng sai số trung phương, các tham số hồi quy (m, b) và hệ số tương quan R.

Trong luận án sử dụng 4 bộ dữ liệu DEM cho thực nghiệm. Độ phân giải không gian cho cả bốn bộ dữ liệu DEM thực nghiệm trong nghiên cứu này đã được chọn trong khoảng từ 5m đến 90m và theo đó giá trị hệ số thu phóng là 3 hoặc 4. Có hai lý do cho việc lựa chọn độ phân giải không gian này: Thứ nhất là vì hầu hết các nguồn dữ liệu DEM dạng grid sẵn có hiện có trên thế giới hiện nay như SRTM,



Hình 2-10. Tăng độ phân giải không gian của DEM khu vực Yên Thành, Nghệ An từ độ phân giải 60m đến 20m (bộ dữ liệu D1)

Trong đó: (a)-DEM chuẩn ở độ phân giải 20m; (b)-DEM giảm độ phân giải ở độ phân giải 60m; (c)-DEM 20m được tạo thành sau khi tái chia mẫu sử dụng mô hình song tuyến; (d)-DEM ở độ phân giải 20m sau khi tái chia mẫu sử dụng mô hình Bi-cubic; (e)-DEM ở độ phân giải 20m sau khi sử dụng phép nội suy Kriging.

ASTER đều nằm trong phạm vi độ phân giải này. Lý do thứ hai, quan trọng hơn là việc tăng độ chính xác của các dữ liệu này về mặt độ phân giải không gian có ý nghĩa rất lớn cho nhiều ứng dụng. Các dữ liệu grid DEM có độ phân giải cao hơn cũng có thể được lấy từ máy quét LiDAR hoặc quét Laser 3D (3D Laser Scanners) và chúng đảm bảo độ chính xác cho hầu hết các ứng dụng. Tuy nhiên, các dữ liệu này có giá thành cao, khả năng thu thập trên diện rộng thấp. Do đó, việc tăng thêm

độ chính xác hoặc độ phân giải cho các loại dữ liệu này thực sự không cần thiết và có ý nghĩa.

Ngoài 4 bộ dữ liệu như mô tả ở trên, một nhóm dữ liệu đo trực tiếp ngoài thực địa đã được sử dụng để đánh giá mô hình. Dữ liệu này gồm 236 điểm độ cao được xác định bằng máy toàn đạc điện tử ở cùng khu vực của bộ dữ liệu để thành lập DEM Lạng Sơn 20m và 5m với độ chính xác tương đương.

2.3.1.1 DEM giảm độ phân giải

Bộ dữ liệu DEM giảm độ phân giải thứ nhất (được gọi là bộ dữ liệu D1) có diện tích khoảng 3,5 km x 3,5 km và được thu thập tại huyện Yên Thành, tỉnh Nghệ An, Bắc Trung Bộ. Khu vực này nằm ở vị trí (18°58'57.03" N, 105°22'44.87" E), cách thành phố Vinh khoảng 45 km. Dữ liệu DEM này được sản xuất từ các bản đồ địa hình tỷ lệ 1: 10.000. Độ phân giải không gian của DEM gốc ban đầu là 20m (Hình 2-10(a)) và DEM này đã được làm giảm độ phân giải xuống 60m bằng cách lấy trung bình giá trị độ cao thông qua đường biên của điểm ảnh (footprint) của các pixel trong DEM gốc độ phân giải 20m như ở Hình 2-10(b).



Hình 2-11. Tăng độ phân giải không gian của DEM SRTM khu vực Yên Thành, Nghệ An) từ độ phân giải 90m đến 30m (bộ dữ liệu D2)

Trong đó: (a)-DEM tham chiếu ở độ phân giải 30m; (b)-DEM giảm độ phân giải ở độ phân giải 90m (lưu ý: điều này tạo thành đầu vào duy nhất cho các thuật toán); (c)- DEM độ phân giải 30m sau khi tái chia mẫu song tuyến; (d)-DEM độ phân giải 30m sau khi tái chia mẫu Bi-cubic; (e)-DEM độ phân giải 30m sau khi nội suy Kriging.

Bộ dữ liệu DEM giảm độ phân giải thứ hai (được gọi là bộ dữ liệu D2) là DEM SRTM được cung cấp bởi USGS Earth Explorer (http://earthexplorer.usgs.gov/) (Hình 2-11(a)). Bộ dữ liệu này cũng bao phủ ở cùng một khu vực với bộ dữ liệu DEM thứ nhất nhưng với độ phân giải không gian là 30m. Dữ liệu này được làm giảm độ phân giải xuống 90m để tạo ra một bộ dữ liệu thử nghiệm thứ hai cho các thuật toán tái chia mẫu (Hình 2-11(b)).

2.3.1.2 Các bộ dữ liệu thực nghiệm bằng DEM thực (Sampled DEM)

Bộ dữ liệu DEM thực thứ nhất (được đặt tên là bộ dữ liệu S1) được thu thập bằng phương pháp đo đạc trực tiếp trên thực địa ở tỉnh Lạng Sơn của Việt Nam. Khu vực thực nghiệm có diện tích khoảng 200m x 200m thuộc phường Mai Pha, thành phố Lạng Sơn, cách Hà Nội khoảng 150km.

Một tập hợp gồm 533 điểm độ cao được thu thập theo phương pháp đo đạc trực tiếp ngoài thực địa, sau đó sử dụng phép nội suy Kriging để tạo ra một bộ dữ liệu DEM có độ phân giải không gian 5m dùng làm dữ liệu DEM tham chiếu như trong Hình 2-12(a). Độ chính xác của DEM tham chiếu này được đánh giá dựa trên Tiêu chuẩn ASPRS 1990 (Tiêu chuẩn dữ liệu không gian địa lý kỹ thuật số) với một bộ gồm 234 điểm kiểm tra. Kết quả đánh giá trong Bảng 2-1 cho thấy chất lượng của DEM tham chiếu tốt hơn một chút so với tiêu chuẩn của DEM độ phân giải 5m trong tiêu chuẩn ASPRS 1990.

Bảng 2-1. Đánh giá độ chính xác của DEM dựa trên tiêu chuẩn ASPRS cho dữ liệu không
gian địa lý kỹ thuật số

Bộ dữ liệu và tiêu chuẩn	Độ SSTP khi không có	Khoảng cao đều đường bình độ tương ứng với giá trị của SSTP			
	lớp phủ thực vật (m)	không có thực phủ (NVA at 95%) (cm)	trong vung co thực phủ (VVA at 95%) (cm)		
DEM Mai Pha, Lạng Sơn	0.483	1.449-mét	1.449	1.449-mét	
Tiêu chuẩn ASPRS class VIII (66.7-cm)	0.667	2-mét	200	2-mét	

Cũng từ tập dữ liệu điểm độ cao nói trên, chúng tôi tiến hành tạo ra bộ DEM với độ phân giải 20m. Bộ dữ liệu DEM có độ phân giải 20m này được sử dụng làm dữ liệu đầu vào cho thuật toán tăng độ phân giải DEM trong nghiên cứu này. Và sau

khi tăng độ phân giải DEM từ 20m lên 5m, chúng tôi đã tiến hành so sánh DEM kết quả này với DEM tham chiếu ở độ phân giải 5m.



Hình 2-12. Dữ liệu DEM khu vực Mai Pha, Lạng Sơn sau khi tăng độ phân giải không gian từ 20m lên 5m (bộ dữ liệu S1)

Trong đó: (a)-Dữ liệu DEM tham chiếu ở độ phân giải 5m; (b)- Dữ liệu DEM giảm độ phân giải xuống 20m (lưu ý: điều này tạo thành đầu vào duy nhất cho các thuật toán); (c)-DEM ở độ phân giải 5m được nội suy theo phương pháp song tuyến; (d)-DEM ở độ phân giải 5m được nội suy theo phương pháp Bi-cubic; (e)-DEM ở độ phân giải 5m được nội suy theo phương pháp Kriging.

Bộ dữ liệu DEM thực thứ hai (được đặt tên là bộ dữ liệu S2) được tạo từ dữ liệu đường bình độ ở khu vực huyện Đắc Hà, tỉnh Kontum, Việt Nam. Vị trí của khu vực thu thập bộ dữ liệu DEM này là (14.671794 ° N và 107,967292 ° E). Diện tích của vùng thực nghiệm là khoảng 6,6 km x 6,6 km.



Hình 2-13. Dữ liệu DEM khu vực Đắc Hà, Kon Tum sau khi tăng độ phân giải không gian từ 90m lên 30m (Bộ dữ liệu S2)

Trong đó: (a)-Dữ liệu đường bình độ với khoảng cao đều 5m, (b)-Dữ liệu đường bình độ với khoảng cao đều 10m, (c)-Dữ liệu DEM tham chiếu ở độ phân giải 30m (được tạo từ dữ liệu đường bình độ với khoảng cao đều 5m),(d)-Dữ liệu DEM đầu vào ở độ phân giải 90m (được tạo từ dữ liệu đường bình độ với khoảng cao đều 10m) (lưu ý: điều này tạo thành đầu vào duy nhất cho các thuật toán), (e)-DEM ở độ phân giải 30m sau khi tái chia mẫu theo phương pháp song tuyến, (f)-DEM ở độ phân giải 30m sau khi tái chia mẫu theo phương pháp Bi-cubic, (g)-DEM ở độ phân giải 30m được nội suy theo phương pháp Kriging.

Từ dữ liệu đường bình độ gốc với khoảng cao đều 5m (Hình 2-13(a)), đã tạo ra một DEM ở độ phân giải 30m và được sử dụng làm dữ liệu DEM tham chiếu (Hình 2-13(c)). Sau đó, đường bình độ gốc được lược bỏ thành bộ dữ liệu đường

bình độ với khoảng cao đều là 10m (Hình 2-13(b)) để tạo ra bộ dữ liệu DEM đầu vào ở độ phân giải không gian 90m (Hình 2-13(d)).

Việc đánh giá độ chính xác được thực hiện bằng cách so sánh DEM kết quả ở độ phân giải 30m (DEM sau khi đã được tăng độ phân giải từ 90m lên 30m với hệ số thu phóng bằng 3) với dữ liệu DEM chuẩn (DEM tham chiếu).

2.3.2 Phân tích về độ chính xác

2.3.2.1 Đánh giá trực quan

2.3.2.1.1 So sánh trực tiếp

So sánh trực quan cho thấy các DEM kết quả được tạo ra sau khi tái chia mẫu bằng các phương pháp song tuyến, Bi-cubic và Kriging tương đối khớp với DEM chuẩn hơn so với các dữ liệu đầu vào là DEM giảm độ phân giải và DEM có độ phân giải thấp tạo ra từ tập điểm độ cao và đường bình độ.

Sự cải thiện về mức độ giống nhau khi so sánh trực quan giữa các DEM sau khi tái chia mẫu và DEM tham chiếu rất rõ ràng khi so sánh giữa các DEM 20m và 30m trong các bộ dữ liệu giảm độ phân giải ở khu vực Nghệ An (Hình 2-10 và Hình 2-11) và các DEM sau khi tái chia mẫu thành độ phân giải 5m và 30m với các DEM tham chiếu (Hình 2-12 và Hình 2-13). Trong khi đó, hình ảnh của các DEM gốc và DEM sau khi tái chia mẫu theo phương pháp Bi-cubic, ảnh đã bị mờ đi do nhiễu và hình dạng của các thuộc tính địa hình trong trường hợp này cũng đã bị méo. Hình ảnh của phép nội suy Kriging khi giảm độ phân giải DEM trong Hình 2-10(e), Hình 2-11(e), Hình 2-12(e) và Hình 2-13(g) nhìn có vẻ không có nhiễu và rất giống với DEM chuẩn trong Hình 2-10(a), Hình 2-11(a), Hình 2-12(a) và Hình 2-14(c). Sự cải thiện rõ ràng nhất của việc tái cấu trúc hình dạng địa hình từ dữ liệu DEM độ phân giải thô có thể được nhìn thấy trong các khu vực được đánh dấu trong Hình 2-10, Hình 2-11, Hình 2-12 và Hình 2-13.

2.3.2.1.2 Sử dụng phương pháp mặt cắt

Để đánh giá tác động của thuật toán trên từng khu vực địa hình khác nhau, chúng tôi tiến hành xây dựng các mặt cắt trên bề mặt DEM và so sánh các DEM

theo các mặt cắt này. Các mặt cắt được xây dựng theo cả chiều dọc và chiều ngang của khu vực thực nghiệm. Các vị trí mặt cắt cho cả 4 khu vực thực nghiệm được trình bày như Hình 2-14.



Hình 2-14. Sơ đồ vị trí các mặt cắt của 4 khu vực thực nghiệm

Trong đó: (a)-Khu vực thực nghiệm Nghệ An với DEM 20m, (b)-Khu vực thực nghiệm Nghệ An với DEM 30m, (c)-Khu vực thực nghiệm Lạng sơn với DEM 5m, (d)-Khu vực thực nghiệm Đắc Hà, Kon Tum với DEM 30m.

Việc so sánh các bề mặt của các DEM kết quả được thể hiện qua các mặt cắt ở Hình 2-15, Hình 2-16, Hình 2-17 và Hình 2-18 cho thấy các phương pháp tái chia mẫu đều cải thiện độ chính xác của bề mặt DEM. Các bề mặt DEM tái chia mẫu nằm sát hơn so với bề mặt DEM chuẩn so với DEM độ phân giải thấp chưa được tái chia mẫu. Trong Hình 2-15, Hình 2-16, Hình 2-17 và Hình 2-18, có thể thấy trong các phương pháp tái chia mẫu cơ bản thì phương pháp Kriging và phương pháp song tuyến Bilinear là các phương pháp tốt hơn so với phương pháp Bi-cubic. Các mặt cắt đều cho thấy bề mặt là kết quả của phương pháp nội suy Kriging gần với



các mặt cắt của DEM tham chiếu hơn so với các mặt cắt theo phương pháp tái chia mẫu song tuyến Bilinear và Bi-cubic đối với cả 4 bộ dữ liệu thực nghiệm.

Hình 2-15. Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM giảm độ phân giải 20m khu vực Nghệ An

Điều này có thể nhìn thấy rõ ràng qua các mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM 5m Lạng Sơn (Hình 2-17) ở những nơi như đỉnh đồi hoặc đáy thung lũng. Trong các hình ảnh này, có thể quan sát thấy các bề mặt DEM được tạo ra theo phương pháp tái chia mẫu Bi-cubic thì gần với bề mặt DEM gốc hơn, trong khi bề mặt được tạo ra bởi phương pháp song tuyến và nội suy Kriging thì lại gần với bề mặt DEM tham chiếu 5m hơn.



Hình 2-16. Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM giảm độ phân giải 30m ở Nghệ An

Phép nội suy Kriging thực hiện chính xác hơn nhiều so với các phương pháp tái chia mẫu song tuyến và Bi-cubic tại các vùng địa hình có sự thay đổi độ cao lớn như đỉnh của các rặng núi, đồi hoặc đáy thung lũng, đặc biệt là các thung lũng hình chữ V, các sườn và đỉnh núi nhọn.



Hình 2-17. Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM 5m được lấy làm DEM tham chiếu khu vực Lạng Sơn, Việt Nam

So sánh giữa các bề mặt DEM trên mặt cắt được đánh dấu bởi các đường được đánh dấu bao gồm: DEM chuẩn (DEM gốc), DEM chưa được tăng độ phân giải (DEM NN), DEM được tái chia mẫu bằng phương pháp song tuyến (DEM Bilinear), Bi-cubic (DEM Bi-cubic) và DEM được tái chia mẫu bằng phương pháp Kriging (DEM Kriging) cho thấy tại hầu hết các điểm thì kết quả của hai phương pháp song tuyến và Bi-cubic luôn khá gần nhau và khác so với phương pháp Kriging.



Hình 2-18. Một số mặt cắt dọc và mặt cắt ngang của bộ dữ liệu DEM 30m được lấy làm DEM tham chiếu khu vực Đắc Hà, Việt Nam

Trong khi phương pháp song tuyến chỉ sử dụng các pixel lân cận điểm cần nội suy còn phương pháp Bi-cubic thì sử dụng thêm các pixel ở xa điểm cần nội suy. Điều đó cho thấy khi nội suy để xác định độ cao đối với grid DEM, các pixel ở xa ngoài pixel lân cận sẽ không ảnh hưởng nhiều đến kết quả nội suy.

2.3.2.1.3 So sánh bằng biểu đồ phân tán (Scatter)

Kết quả so sánh trực quan bằng biểu đồ phân tán được thể hiện trong các Hình 2-19, Hình 2-20, Hình 2-21 và Hình 2-22. Trong các biểu đồ phân tán này, hai bộ dữ liệu DEM được coi là gần nhau hơn nếu tất cả các điểm dữ liệu đều nằm gần nhau và nằm gần đường hồi quy. Điều đó có nghĩa là phần sai số ngẫu nhiên có giá trị thấp và các điểm đo có độ chính xác (precision) tốt. Ngược lại, nếu các điểm dữ liệu tương ứng trên hai DEM khác nhau và giá trị sai khác này xuất hiện một cách ngẫu nhiên, các điểm trên biểu đồ phân tán sẽ nằm rời rạc và cách xa đường thẳng hồi quy.

Như vậy, dựa trên biểu đồ phân tán có thể xác định được một cách trực quan mức độ khớp nhau của hai tập dữ liệu độ cao.

So sánh trực quan bằng biểu đồ phân tán trong các Hình 2-19, Hình 2-20, Hình 2-21, Hình 2-22 cho thấy: Độ phân tán giữa DEM tái chia mẫu của dữ liệu độ cao chưa được tăng độ phân giải (N-DEM) là lớn nhất, độ phân tán này giảm dần theo thứ tự là DEM sau tái chia mẫu Bi-cubic (C-DEM), DEM sau tái chia mẫu song tuyến Bilinear (B-DEM), DEM kết quả của phương pháp Kriging (K-DEM). Điều này có nghĩa là phương pháp Kriging có thể tạo ra DEM có giá trị độ cao sát với giá trị độ cao trên DEM tham chiếu hơn các phương pháp tái chia mẫu còn lại.



Hình 2-19. Các biểu đồ tán xạ của các DEM tham chiếu có độ phân giải không gian cao so sánh với các DEM được tăng độ phân giải, đối với bộ dữ liệu DEM giảm độ phân giải 20m tại khu vực Nghệ An

Trong đó: (a)-DEM tham chiếu và DEM giảm độ phân giải ban đầu (N-DEM),(b)-DEM tham chiếu và DEM sau tái chia mẫu theo mô hình song tuyến (B-DEM), (c)-DEM tham chiếu và DEM sau khi tái chia mẫu theo mô hình Bi-cubic (C-DEM), (d)-DEM tham chiếu và DEM được nội suy theo phương pháp Kriging (K-DEM).



Hình 2-20. Các biểu đồ tán xạ của các DEM tham chiếu độ phân giải không gian cao so sánh với các DEM được tăng độ phân giải – đối với bộ dữ liệu DEM giảm độ phân giải 30m Nghệ An

Trong đó: (a)-DEM tham chiếu và DEM giảm độ phân giải ban đầu (N-DEM), (b)-DEM tham chiếu và DEM sau khi tái chia mẫu theo mô hình song tuyến (B-DEM), (c)-DEM tham chiếu và DEM sau khi tái chia mẫu theo mô hình Bi-cubic (C-DEM), (d)-DEM tham chiếu và DEM được nội suy theo mô hình Kriging (K-DEM).



Hình 2-21. Các biểu đồ tán xạ của các DEM tham chiếu độ phân giải không gian cao so sánh với các DEM được tăng độ phân giải - đối với bộ dữ liệu DEM mẫu độ phân giải 5m khu vực Lạng Sơn

Trong đó: (a)-DEM tham chiếu và DEM giảm độ phân giải ban đầu (N-DEM), (b)-DEM tham chiếu và DEM sau khi tái chia mẫu theo phương pháp song tuyến (B-DEM), (c)-DEM tham chiếu và DEM sau khi tái chia mẫu theo phương pháp Bi-cubic (C-DEM), (d)-DEM tham chiếu và DEM được nội suy theo phương pháp Kriging (K-DEM).



Hình 2-22. Các biểu đồ tán xạ của các DEM tham chiếu độ phân giải không gian cao so sánh với các DEM được tăng độ phân giải - đối với bộ dữ liệu DEM mẫu độ phân giải 30m tại Đắc Hà

Trong đó: (a)-DEM tham chiếu và DEM giảm độ phân giải ban đầu (N-DEM), (b)-DEM tham chiếu và DEM sau khi tái chia mẫu theo phương pháp song tuyến (B-DEM), (c)-DEM tham chiếu và DEM sau khi tái chia mẫu theo phương pháp Bi-cubic (C-DEM), (d)-DEM tham chiếu và DEM được nội suy theo phương pháp Kriging (K-DEM).

Mặt khác, các tham số của đường hồi quy sẽ phản ánh xu thế của các điểm độ cao trên mặt DEM cần đánh giá so với bề mặt DEM chuẩn, phản ánh thành phần sai số hệ thống có trong DEM. Nếu hệ số độ dốc m gần với giá trị 1 hơn và hệ số chặn b gần với giá trị 0 hơn thì thành phần sai số hệ thống tồn tại trong các điểm ảnh sẽ nhỏ hơn.

Các biểu đồ tán xạ của các kết quả tái chia mẫu trong Hình 2-19, Hình 2-20, Hình 2-21 và Hình 2-22 cho thấy sự khớp nhau hơn giữa dữ liệu DEM tham chiếu và các dữ liệu DEM sau khi tái chia mẫu (B-DEM, C-DEM, K-DEM) khi so sánh với dữ liệu DEM gốc ban đầu (N-DEM). So sánh giữa dữ liệu DEM sau khi tái chia mẫu song tuyến Bilinear (Hình 2-19(b), Hình 2-20(b), Hình 2-21(b) và Hình 2-22(b)) dữ liệu DEM sau tái chia mẫu theo Bi-cubic (Hình 2-19(c), Hình 2-20(c), Hình 2-21(c) và Hình 2-22(c)), dữ liệu DEM sau nội suy Kriging (Hình 2-19(d), Hình 2-20(d), Hình 2-21(d) và Hình 2-22(d)). Sự cải thiện độ phân giải này có thể thấy rõ nhất với bộ dữ liệu DEM lấy mẫu độ phân giải 5m khu vực Lạng Sơn và bộ dữ liệu DEM giảm độ phân giải ở độ phân giải 20m khu vực Nghệ An. Các điểm đữ liệu trong các biểu đồ phân tán trong Hình 2-20, Hình 2-21 cho thấy nó rất gần, (có nhiều điểm trùng khít) với đường khớp tuyệt đối và hệ số phù hợp nhất trong các biểu đồ phân tán này gần với giá trị 1 và 0.

Khi so sánh bốn bộ dữ liệu DEM này, chúng ta thấy rằng: các điểm dữ liệu trong các biểu đồ phân tán trong Hình 2-19(b), Hình 2-20(b), Hình 2-21(b) và Hình 2-22(b) (dữ liệu DEM sau khi tái chia mẫu theo phương pháp song tuyến), Hình 2-19(c), Hình 2-20(c), Hình 2-21(c) và Hình 2-22(c) (dữ liệu DEM sau khi tái chia mẫu theo phương pháp Bi-cubic), Hình 2-19(d), Hình 2-20(d), Hình 2-21(d) và Hình 2-22(d) (dữ liệu DEM sau khi được nội suy theo phương pháp Kriging) ít bị phân tán ra khỏi đường khóp tuyệt đối nhất so với dữ liệu DEM gốc.

2.3.2.2 Đánh giá định lượng

2.3.2.2.1 Sử dụng giá trị sai số trung phương

Việc đánh giá định lượng được dựa trên tiêu chuẩn về Sai số trung phương (RMSE), thể hiện trong các Bảng 2-2, Bảng 2-3, Bảng 2-4 và Bảng 2-5. Trên các bảng kết quả này là sai số trung phương của dữ liệu đầu vào, dữ liệu DEM tái chia mẫu theo phương pháp song tuyến Bilinear, phương pháp Bi-cubic và tái chia mẫu sử dụng nội suy Kriging. Các sai số trung phương được tính cho tổng thể toàn bộ bề mặt và cho từng mặt cắt theo các tuyến được đánh dấu như trên Hình 2-14. Kết quả đánh giá định lượng cho thấy sự cải thiện về độ chính xác của DEM cũng phù hợp với kết quả quan sát trực quan.

							-
Bộ dữ liệu	SSTP của DEM gốc (m)	SSTP khi tái lấy mẫu song tuyến (Bilinear) (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi tái lấy mẫu Bi- cubic (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi nội suy Kriging (m)	ĐCX được cải thiện so với DEM gốc (%)
RMSE	6.9326	3.3026	52.4%	3.3716	51.4%	2.8874	58.4%
CP 1*	4.5389	3.0986	31.7%	3.1431	30.8%	3.2242	29.0%
CP 2	4.4169	2.8131	36.3%	2.8973	34.4%	2.8798	34.8%
CP 3	4.3370	2.7674	36.2%	2.8041	35.3%	2.6625	38.6%
CP 4	4.4689	2.9057	35.0%	2.9731	33.5%	2.7696	38.0%
CP 5	4.0911	2.9148	28.8%	2.9445	28.0%	2.9336	28.3%
CP 6	3.8029	2.5245	33.6%	2.5619	32.6%	2.6330	30.8%
CP 7	4.6677	3.1959	31.5%	3.2344	30.7%	3.1115	33.3%
CP 8	4.8884	2.9958	38.7%	3.0833	36.9%	2.9249	40.2%
CP 9	5.1846	2.9851	42.4%	3.0731	40.7%	2.7065	47.8%
CP 10	5.2172	3.3379	36.0%	3.4256	34.3%	3.2270	38.1%
CP 11	4.3794	2.5489	41.8%	2.6209	40.2%	2.4393	44.3%
RP 1	6.9375	3.7005	46.7%	3.6816	46.9%	3.6813	46.9%
RP 2	6.4972	2.9903	54.0%	3.0293	53.4%	2.8799	55.7%
RP 3	4.5824	2.8843	37.1%	2.9332	36.0%	2.8899	36.9%
RP 4	7.0182	3.4087	51.4%	3.4013	51.5%	3.1660	54.9%
RP 5	6.5620	3.5779	45.5%	3.5906	45.3%	3.4508	47.4%
RP 6	6.9686	3.3586	51.8%	3.4280	50.8%	3.3106	52.5%
RP 7	6.8329	3.1977	53.2%	3.2778	52.0%	3.1780	53.5%
RP 8	7.7733	3.7850	51.3%	3.7997	51.1%	3.7522	51.7%
RP 9	5.7281	2.7969	51.2%	2.9109	49.2%	2.7287	52.4%
RP 10	5.0358	2.3813	52.7%	2.4803	50.7%	2.2053	56.2%
RP 11	2.3477	1.3837	41.1%	1.4051	40.1%	1.3916	40.7%

Bảng 2-2. Sai số trung phương (RMSE) của các phương pháp tái chia mẫu sử dụng mô hình song tuyến (Bilinear), Bi-cubic và Kriging cho DEM 20m khu vực Nghệ An

Bảng 2-3. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng mô hình song tuyến (Bilinear), tái lấy mẫu Bi-cubic và theo thuật toán nội suy Kriging cho DEM SRTM 30m khu vực Nghệ An

Bộ dữ liệu	SSTP của DEM gốc (m)	SSTP khi tái lẫy mẫu song tuyến (Bilinear) (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi tái lẫy mẫu Bi- cubic (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi nội suy Kriging (m)	ĐCX được cải thiện so với DEM gốc (%)
Overall	11.1379	8.8105	20.9%	8.8736	20.3%	8.5719	23.0%
RMSE							
CP 1	8.5013	6.8408	19.5%	6.9101	18.7%	6.9668	18.1%
CP 2	9.7106	8.4326	13.2%	8.4863	12.6%	8.5101	12.4%
CP 3	11.6961	10.7635	8.0%	10.8141	7.5%	11.0702	5.4%
CP4	10.0198	8.9907	10.3%	9.0225	10.0%	9.2379	7.8%
CP 5	9.2745	7.0420	24.1%	7.2130	22.2%	7.3050	21.2%
CP 6	11.5945	9.8018	15.5%	9.8618	14.9%	9.7752	15.7%
CP 7	9.7925	8.3543	14.7%	8.4407	13.8%	8.4730	13.5%
RP 1	10.4429	9.8024	6.1%	9.8357	5.8%	9.6199	7.9%
RP 2	9.9168	8.0953	18.4%	8.0897	18.4%	8.0478	18.8%
RP 3	10.5144	9.6251	8.5%	9.6645	8.1%	9.5933	8.8%
RP 4	9.9849	7.7341	22.5%	7.8310	21.6%	7.6501	23.4%
RP 5	9.8911	8.4770	14.3%	8.5192	13.9%	8.1091	18.0%
RP 6	8.8079	7.7367	12.2%	7.7801	11.7%	7.3437	16.6%
RP 7	6.6352	6.4032	3.5%	6.4005	3.5%	6.2829	5.3%

Bộ dữ liệu	SSTP của DEM gốc (m)	SSTP của DEM tái lấy mẫu theo mô hình song tuyến (Bilinear) (m)	ĐCX được cải thiện so với DEM gốc KDEM (%)	SSTP của DEM tái lấy mẫu Bi- cubic (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi nội suy Kriging (m)	ĐCX được cải thiện so với DEM gốc (%)
SSTP tổng thể	2.8493	1.5139	78.25	1.6000	88.39	1.2092	42.35
CP 1	1.4960	1.2419	16.98	1.2912	13.69	1.0288	31.23
CP 2	1.6962	1.1635	31.40	1.1821	30.31	1.4066	17.08
CP 3	2.0641	1.4043	31.97	1.4791	28.34	1.2680	38.57
CP4	2.2345	1.3591	39.18	1.4586	34.72	1.1416	48.91
CP 5	2.2705	1.3006	42.72	1.3728	39.53	1.4232	37.32
CP 6	2.3084	1.7034	26.21	1.7805	22.87	1.1232	51.34
CP 7	2.0349	1.6198	24.40	1.6569	18.57	1.1355	44.20
RP 1	1.9569	1.4024	28.33	1.4348	26.68	0.9771	50.07
RP 2	2.2873	1.6555	27.62	1.7196	24.82	1.5087	34.04
RP 3	2.3612	1.6712	29.22	1.7451	26.09	1.2566	46.78
RP 4	1.9510	1.4361	26.39	1.5174	22.23	1.6807	13.85
RP 5	1.7489	1.4228	18.64	1.4657	16.19	1.2011	31.32
RP 6	1.6217	1.4081	18.55	1.4297	17.31	1.4138	18.22
RP 7	1.3897	1.1567	28.67	1.2101	25.38	0.7408	54.32

Bảng 2-4. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng mô hình song tuyến (Bilinear), Bi-cubic và thuật toán Kriging cho DEM 5m khu vực Lạng Sơn

vực Đác Hà									
Bộ dữ liệu	SSTP của DEM gốc (m)	SSTP của DEM tái chia mẫu theo mô hình song tuyến (Bilinear) (m)	ĐCX được cải thiện so với DEM gốc KDEM (%)	SSTP khi nội suy tái chia mẫu theo mô hình Bi- cubic (m)	ĐCX được cải thiện so với DEM gốc (%)	SSTP khi nội suy Kriging (m)	ĐCX được cải thiện so với DEM gốc (%)		
SSTP tổng thể	5.0680	2.3284	54.1%	2.4218	52.2%	2.1095	58.4%		
CP 1	4.0702	2.3434	42.4%	2.4436	40.0%	2.2330	45.1%		
CP 2	4.0203	2.0594	48.8%	2.2048	45.2%	1.9960	50.4%		
CP 3	3.8541	2.0370	47.1%	2.0956	45.6%	2.1132	45.2%		
CP 4	3.5399	2.2395	36.7%	2.2698	35.9%	2.1836	38.3%		
CP 5	3.4595	1.8231	47.3%	1.9178	44.6%	1.7027	50.8%		
CP 6	2.3885	1.3172	44.9%	1.3603	43.0%	1.3345	44.1%		
CP 7	2.6743	1.1377	57.5%	1.2514	53.2%	1.0253	61.7%		
CP 8	2.2896	1.0186	55.5%	1.1031	51.8%	0.9693	57.7%		
CP 9	2.0938	1.0068	51.9%	1.0624	49.3%	0.9942	52.5%		
RP 1	4.6476	1.8714	59.7%	1.9762	57.5%	2.0200	56.5%		
RP 2	4.6907	2.2973	51.0%	2.3739	49.4%	2.2684	51.6%		
RP 3	4.5113	1.7205	61.9%	1.8051	60.0%	1.4827	67.1%		
RP 4	3.6187	1.4932	58.7%	1.5675	56.7%	1.3472	62.8%		
RP 5	3.9713	1.2816	67.7%	1.4015	64.7%	1.1203	71.8%		
RP 6	3.2681	1.0181	68.8%	1.0540	67.7%	0.9603	70.6%		
RP 7	2.8494	1.5355	46.1%	1.5676	45.0%	1.3829	51.5%		
RP 8	3.0366	1.0666	64.9%	1.0934	64.0%	1.1003	63.8%		
RP 9	4.2630	1.9581	54.1%	2.0194	52.6%	1.8029	57.7%		
RP 10	5.2869	2.3473	55.6%	2.4070	54.5%	2.5440	51.9%		

Bảng 2-5. Sai số trung phương (RMSE) cho các phương pháp tái chia mẫu sử dụng phương pháp tái chia mẫu song tuyến (Bilinear), Bi-cubic và thuật toán Kriging cho DEM 30m khu vực Đắc Hà

Qua các kết quả đánh giá định lượng trên cho thấy: phương pháp tăng độ phân giải không gian của mô hình DEM dạng grid theo các phương pháp tái chia mẫu và theo phương pháp nội suy Kriging cho độ chính xác cao hơn DEM ban đầu khi chạy thử nghiệm trên cả bốn bộ dữ liệu DEM.

Sai số trung phương cho các phương pháp tái chia mẫu theo phương pháp song tuyến, Bi-cubic và phương pháp nội suy Kriging lần lượt là 3.3026 m, 3.3716 m và 2.8874 m. So sánh với sai số trung phương của dữ liệu DEM gốc độ phân giải 60m ban đầu, sai số trung phương của các DEM giảm độ phân giải ở độ phân giải 20m giảm đáng kể cho cả ba phương pháp. Còn đối với dữ liệu DEM giảm độ phân giải 30m khu vực Nghệ An, mặc dù độ chính xác của thuật toán tái chia mẫu cho bộ dữ liệu thử nghiệm này cũng tăng nhưng không lớn bằng của bộ dữ liệu DEM giảm độ phân giải 20m. Tuy nhiên, mức độ tăng độ chính xác cũng khá thuyết phục, với sai số trung phương giảm khoảng 20% so với DEM 90m ban đầu. Giống như kết quả đạt được từ các bộ dữ liệu thực nghiệm giảm độ phân giải, độ chính xác cho các dữ liệu thực nghiệm trên DEM thực (sampled datasets) cũng tăng đáng kế. Sai số trung phương của dữ liệu DEM độ phân giải 5m khu vực Lạng Sơn đã giảm mạnh khi sử dụng các thuật tái chia mẫu DEM, với các giá trị 1,5139 m cho phương pháp song tuyến, 1,6 m cho phương pháp Bi-cubic và 1,2092 m cho phép nội suy Kriging. Kết quả thử nghiệm trên dữ liệu DEM 30m khu vực Đắc Hà không ấn tượng bằng trên dữ liệu DEM 5m Lạng Sơn. Tuy nhiên, sự cải thiện độ chính xác của DEM vẫn có ý nghĩa vì sai số trung phương đã giảm khoảng 50% so với dữ liệu DEM gốc ban đầu.

Các thống kê này chứng minh rằng các phương pháp tái chia mẫu có thể làm tăng độ chính xác của DEM dạng grid khi nó được sử dụng để tăng độ phân giải không gian của DEM thành một DEM mới có độ phân giải không gian cao hơn.

Việc tăng độ chính xác về mặt sai số trung phương cho các mặt cắt đã chứng minh sự ảnh hưởng của các thuộc tính địa hình đối với thuật toán này. Đối với các bộ dữ liệu DEM 20m và DEM 30m ở khu vực Nghệ An, việc tăng độ chính xác giữa các DEM gốc và các DEM sau khi được tái chia mẫu là tương đối ổn định. Đối với bộ dữ liệu DEM 30m, mức tăng độ chính xác của hầu hết các mặt cắt là từ 20% và đối với dữ liệu DEM lấy mẫu 5m khu vực Lạng Sơn đã có sự biến động hơn, với giá trị nhỏ nhất là 18% và giá trị lớn nhất là 60%. Điều này là do hầu hết các mặt cắt ngang 2, 4, 9) - là các mặt cắt tại các khu vực có địa hình đặc trưng như đáy thung lũng

hoặc ngọn đồi. Ngược lại, các mặt cắt có độ chính xác tăng ít hơn xảy ra chủ yếu ở các vùng sườn núi, nơi bề mặt của DEM gốc (thô) tương đối gần với DEM tham chiếu. Sự tăng độ chính xác có biến thiên nhỏ hơn đối với các dữ liệu DEM giảm độ phân giải 20m, 30m và bộ dữ liệu DEM thực 30m khu vực Đắc Hà là vì hầu hết các mặt cắt này nằm dọc theo nhiều loại địa hình khác nhau, không phải trên các dạng địa hình đặc trưng. Phạm vi tăng độ chính xác của bộ dữ liệu DEM 20m là 20% và từ 50% đến 70%.

2.3.2.2.2 <u>Sử dụng giá trị thống kê *R* (Hệ số tương quan) và phương trình hồi quy (đại diện bằng hai tham số *m* và *b*)</u>

Sự giống nhau giữa hai loại DEM cũng có thể được đánh giá định lượng bằng cách sử dụng các hệ số hồi quy tuyến tính (m, b) và hệ số tương quan R. Khi so sánh hai DEM, nếu gọi độ cao của một pixel trong tập dữ liệu tham chiếu là x và độ cao của pixel tương ứng trong tập dữ liệu so sánh là y, thì đường khớp tuyệt đối sẽ là y = x sao cho m = 1 và b = 0. Vì giá trị của m có thể lớn hơn hoặc nhỏ hơn 1 và giá trị của b có thể lớn hơn hoặc nhỏ hơn 0, nên việc so sánh giữa các giá trị khác nhau của m và b để xác định độ gần của chúng tương ứng với 1 và 0, đôi khi không dễ đánh giá.

Để dễ dàng hơn cho việc đánh giá này, các tham số phụ như | 1-m | và | b | đã được tính toán như trong Bảng 2-6. Theo đó, khi giá trị của các tham số phụ | 1-m | và | b | đều đồng thời càng nhỏ thì hai bộ dữ liệu càng giống nhau. Tham số thứ ba để đánh giá sự khớp nhau của hai bộ dữ liệu là hệ số tương quan R. Hệ số tương quan này dùng để đo sự liên kết giữa hai bộ dữ liệu, do đó sẽ đánh giá được sự phân phối các điểm dữ liệu trong các biểu đồ phân tán xung quanh đường khớp tuyệt đối. Khi giá trị của R^2 càng gần đến 1 thì càng có nhiều điểm dữ liệu được đặt gần với đường khớp tuyệt đối. Một sự trùng khớp tuyệt đối giữa hai bộ dữ liệu DEM là khi tất cả các điểm dữ liệu được đặt trên đường khớp tuyệt đối (y = x) và hệ số xác định $R^2 = 1$. Điều đó có nghĩa là hai bộ dữ liệu sẽ giống hệt nhau nếu các giá trị m = 1, b = 0 và $R^2 = 1$ một cách đồng thời.

Để đánh giá kết quả của các phương pháp khác nhau, các mô hình hồi quy tuyến tính đã được sử dụng để mô tả vào mối quan hệ giữa dữ liệu chuẩn và dữ liệu được tái chia mẫu như trong Bảng 2-6. Các giá trị hệ số cho thấy DEM được nội suy theo phương pháp Kriging giống với các DEM tham chiếu hơn so với các DEM gốc ban đầu hay các DEM tái chia mẫu theo các phương pháp song tuyến tính, Bicubic. Đối với tất cả bốn bô dữ liêu, giá tri của các tham số m, b và R^2 của DEM nôi suy theo phương pháp Kriging gần với các giá trị 1, 0 và 1 hơn so với các giá trị này của DEM gốc ban đầu hay của các DEM tái chia mẫu theo các phương pháp song tuyến tính và Bi-cubic. Đối với bô dữ liêu DEM được tái chia mẫu ở đô phân giải 5m khu vực Lang Sơn, các giá tri |1-m| = 0.0550, |b| = 16.3717 và $R^2 = 0.9884$ cho DEM nôi suy theo phương pháp Kriging cho thấy sư tương đồng lớn hơn với DEM tham chiếu so với DEM thô ban đầu (| 1-m | = 0,0 310, | b | = 9.3306 và R^2 = 0.9425), DEM tái chia mẫu theo phương pháp song tuyến (|1-m| = 0.0399, |b| =12,3782 và $R^2 = 0.9793$), DEM tái chia mẫu theo phương pháp Bi-cubic (| 1-m | = 0,0342, |b| = 10,6432 và $R^2 = 0,9763$). Các hệ số hồi quy tuyến tính cho bộ dữ liệu DEM tham chiếu (DEM mẫu) khu vực Đắc Hà cũng chỉ ra sự gần giống hơn giữa phương pháp nội suy Kriging với DEM tham chiếu, với (| 1-m | = 0,0078, | b | = 7,3917 và $R^2 = 0.9967$) so với (| 1-m | = 0.0128, | b | = 12.1453 và $R^2 = 0.9960$) của phương pháp tái chia mẫu song tuyến và (|1-m| = 0,0115, |b| = 10,9118 và $R^2 =$ 0,9959) của phương pháp tái chia mẫu Bi-cubic. Trong ba phương pháp tái chia mẫu, kết quả không chính xác nhất là DEM theo phương pháp Bi-cubic, nhưng so với dữ liêu DEM gốc ở đô phân giải ban đầu, những kết quả này là chính xác hơn nhiều. Mặt khác có thể thấy, việc các giá tri chặn b đều mạng dấu dương và có giá trị khá lớn cho thấy các bề mặt DEM tái chia mẫu đều có xu thế cao hơn so với DEM chuẩn mà nguyên nhân của nó là do tính khái quát hóa của việc sử dụng các grid có kích thước lớn để mô phỏng địa hình.

Bảng 2-6. Các hệ số hồi quy tuyến tính cho các bộ dữ liệu DEM sau khi tái chia mẫu độ phân giải 20m khu vực Nghệ An và DEM độ phân giải 30m khu vực Nghệ An và các bộ dữ liệu DEM lấy mẫu độ phân giải 5m khu vực Lang Sơn và DEM độ phân giải 30m khu vực Đắc Hà

Bộ dữ liệu		Các hệ số hồi quy tuyến tính						
		т	1 - m	b	b	R^2		
	DEM giảm độ phân giải ở độ phân giải 60m	0.9822	0.0178	2.1147	2.1147	0.9770		
Bộ dữ liệu DEM 20m khu vực Nghệ An	DEM tái lấy mẫu theo pp song tuyến ở độ phân giải 20m	0.9765	0.0235	2.5368	2.5368	0.9951		
	DEM tái lấy mẫu theo pp Bi-cubic ở độ phân giải 20m	0.9781	0.0219	2.3680	2.3680	0.9948		
	DEM 20m khi nội suy Kriging	0.9832	0.0168	1.8217	1.8217	0.9962		
	DEM giảm độ phân giải ở độ phân giải 90m	0.9659	0.0341	1.2658	1.2658	0.9412		
Bộ dữ liệu DEM 30m khu vực Nghệ An	DEM tái lấy mẫu theo pp song tuyến ở độ phân giải 30m	0.9500	0.0500	3.2057	3.2057	0.9646		
	DEM tái lấy mẫu theo pp song tuyến ở độ phân giải 20m 30m	0.9529	0.0471	2.8723	2.8723	0.9639		
	DEM 30m khi nội suy Kriging	0.9608	0.0392	1.9291	1.9291	0.9964		
	DEM gốc ban đầu ở độ phân giải 20m	0.9690	0.0310	9.3306	9.3306	0.9425		
Bộ dữ liệu DEM khu vực Lạng	DEM tái lấy mẫu theo pp song tuyến ở độ phân giải 5m	0.9601	0.0399	12.3782	12.3782	0.9793		
Sơn	DEM tái lấy mẫu theo pp Bi- cubic ở độ phân giải 5m	0.9658	0.0342	10.6432	10.6432	0.9763		
	DEM 5m khi nội suy Kriging	0.9450	0.0550	16.3717	16.3717	0.9884		
	DEM gốc ban đầu ở độ phân giải 90m	0.9934	0.0066	6.1988	6.1988	0.9810		
Bộ dữ liệu DEM khu vực Đắc Hà	DEM tái lấy mẫu theo pp song tuyến ở độ phân giải 30m	0.9872	0.0128	12.1453	12.1453	0.9960		
	DEM tái lấy mẫu theo pp Bi- cubic ở độ phân giải DEM 30m	0.9885	0.0115	10.9118	10.9118	0.9959		
	DEM 30m khi nội suy Kriging	0.9922	0.0078	7.3917	7.3917	0.9967		

Các hệ số hồi quy tuyến tính cho bộ dữ liệu DEM giảm độ phân giải ở độ phân giải 20m khu vực Nghệ An cho thấy rằng DEM sau khi tái chia mẫu rất khớp
với DEM tham chiếu. So sánh cũng cho thấy xu thế chung DEM gốc ban đầu với các tham số |1-m| = 0,0178 và |b| = 2.1147 thường có biến thiên tỷ lệ thuận với DEM chuẩn hơn là các DEM sau khi được tái chia mẫu theo phương pháp song tuyến (|1-m| = 0,0235 và |b| = 2.5368) và tái chia mẫu Bi-cubic (|1-m| = 0,0219 và |b| = 2,3680). Tuy nhiên, nhiều điểm dữ liệu của DEM theo phương pháp tái chia mẫu song tuyến ($R^2 = 0,9951$) và Bi-cubic ($R^2 = 0,9948$) được phân phối gần với đường khớp tuyệt đối (y = x) hơn so với DEM 20m ban đầu ($R^2 = 0,9770$). Đối với bộ dữ liệu DEM giảm độ phân giải ở độ phân giải 30m, sự tăng độ chính xác của các phương pháp tái chia mẫu đã thấy rõ ràng hơn khi so sánh các tham số hồi quy tuyến tính của ba phương pháp.

Như đã phân tích trong phần đầu của chương này, các giá tri m và b phản ánh phần ảnh hưởng của sai số hệ thống nằm trong DEM trong khi giá trị R^2 phản ánh phần sai số ngẫu nhiên. Có thể thấy đối với cả 3 phương pháp tái chia mẫu đều làm giảm thành phần sai số ngẫu nhiên trong các DEM khi các giá trị R^2 đều tăng khi so sánh với DEM ban đầu trong đó nội suy Kriging có mức độ cải thiện hơn rất nhiều khi so với các phương pháp khác. Khi so sánh các tham số độ dốc *m* và tham số chặn b của các đường khớp tuyệt đối (đường hồi quy tuyệt đối) của cả bốn bộ dữ liệu cho thấy, rõ ràng tất cả các tham số độ dốc *m* của các DEM tái chia mẫu nhỏ hơn 1 và tham số chặn b lớn hơn 0. Hiện tượng này cho thấy các sai số nằm trong giá trị độ cao của các DEM tái chia mẫu có tính hệ thống. Trong đó hầu hết các giá tri đô cao của các pixel nằm tai các vi trí cục bô thấp (thường là đáy của các thung lũng) được tao bởi các phương pháp này có thể cao hơn các pixel tượng ứng trong DEM tham chiếu. Ngược lại, đối với các vi trí có đô cao lớn (như đỉnh đồi hoặc rặng núi), độ cao của các pixel trong dữ liệu DEM tái chia mẫu có thể thấp hơn so với các pixel tương ứng trong DEM tham chiếu. Hiện tượng này xảy ra là do hiệu ứng làm mịn (được gọi là sự sai lệch có điều kiện, trong đó các điểm có độ cao lớn được nội suy xuống thấp hơn và ở các điểm độ cao thấp thì có xu thế cho kết quả nội suy cao hơn thực tế).

2.4 Kết luận chương 2

Trong chương này, các thuật toán tái chia mẫu để tăng độ phân giải không gian và độ chính xác của các DEM dạng grid như phương pháp Bilinear (song tuyến), Bi-cubic và Kriging đã được tiến hành và đánh giá kết quả bằng cách sử dụng các dữ liệu DEM có độ phân giải không gian ở các độ phân giải khác nhau và các khu vực có địa hình khác nhau. Các thực nghiệm về thuật toán tái chia mẫu đã được thực hiện trên hai loại bộ dữ liệu độ cao: DEM giảm độ phân giải ở độ phân giải không gian 20m và 30m ở tỉnh Nghệ An, Việt Nam và DEM tham khảo (DEM mẫu) ở độ phân giải không gian 5m khu vực tỉnh Lạng Sơn (từ dữ liệu khảo sát mặt đất) và độ phân giải không gian 30m ở Đắc Hà, tỉnh Kontum, Việt Nam (được tạo ra từ các đường đồng mức).

Trong chương này, tác giả cũng đã đề xuất và thử nghiệm các phương pháp mới để đánh giá độ chính xác của các DEM kết quả một cách toàn diện hơn như: các phương pháp đánh giá trực quan (phương pháp so sánh trực tiếp bằng mắt thường, phương pháp sử dụng các mặt cắt, phương pháp so sánh bằng biểu đồ phân tán), các phương pháp đánh giá định lượng (phương pháp sử dụng giá trị sai số trung phương, phương pháp sử dụng các giá trị thống kê và phương trình hồi quy).

Các kết quả thử nghiệm đã cho thấy sự tăng đáng kể về độ chính xác đối với các DEM được tái chia mẫu, đặc biệt là từ phương pháp nội suy theo phương pháp Kriging so với DEM ban đầu. Đánh giá trực quan cho thấy sự tương đồng lớn hơn giữa các dữ liệu DEM các DEM được tạo ra bằng các phương pháp tái chia mẫu song tuyến và Bi-cubic, nội suy theo phương pháp Kriging với dữ liệu DEM chuẩn so với DEM có độ phân giải thấp ban đầu. Trong 3 phương pháp đánh giá độ chính xác định lượng dựa trên tiêu chuẩn sai số trung phương cho thấy đã có sự tăng độ chính xác của dữ liệu DEM khi nội suy theo thuật toán Kriging so với các phương pháp tái chia mẫu song tuyến và Bi-cubic. Sai số trung phương của phương pháp nội suy Kriging đã giảm khoảng 58%, 23%, 50% và 58% tương ứng cho các DEM giảm độ phân giải 5m của tỉnh Lạng Sơn và dữ liệu DEM độ phân giải 30m khu vực Đắc Hà, Kon Tum, Việt Nam.

Việc đánh giá độ chính xác bằng cách sử dụng các hệ số hồi quy tuyến tính và hệ số tương quan m, b và R^2 để so sánh giữa DEM độ phân giải không gian cao với DEM gốc ban đầu, DEM sau khi tái chia mẫu theo phương pháp song tuyến và Bi-cubic, DEM nội suy theo phương pháp Kriging cũng cho thấy cải thiện về độ chính xác khi sử dụng tái chia mẫu nếu xét trên phương diện sai số ngẫu nhiên. Tuy nhiên, qua phân tích cũng cho thấy DEM được tạo ra từ tái chia mẫu có xu thế chứa một số sai số mang tính hệ thống khiến bề mặt DEM tạo ra cao hơn thực tế tại các điểm trũng, tụ thủy và có xu thế thấp hơn tại các điểm cao, các đường phân thủy.

Chương 3 NGHIÊN CỨU NÂNG CAO ĐỘ CHÍNH XÁC CỦA MÔ HÌNH SỐ ĐỘ CAO DẠNG GRID BẰNG PHƯƠNG PHÁP SỬ DỤNG MẠNG NEURON HOPFIED

3.1 Cơ sở khoa học của việc ứng dụng mạng neuron Hopfield để tăng độ phân giải không gian và độ chính xác của mô hình số độ cao dạng grid

Độ phân giải không gian là một yếu tố rất quan trọng ảnh hưởng đến độ chính xác của DEM. Khi độ phân giải càng cao (nghĩa là một đơn vị diện tích trên mặt đất thể hiện được nhiều giá trị cao độ) thì mô hình số độ cao DEM càng mô phỏng địa hình mặt đất chính xác hơn và ngược lại. Để nâng độ phân giải không gian của DEM, cần tiến hành đo đạc, thu thập dữ liệu với mức độ chi tiết cao hơn. Hiện nay có các phương pháp thu thập dữ liệu với mật độ điểm rất cao như sử dụng máy bay không người lái UAV đo ở khoảng cách gần, sử dụng LiDAR đo với mật độ điểm và độ chính rất cao. Tuy nhiên việc thu thập dữ liệu với mức độ chi tiết như vậy sẽ mất khá nhiều thời gian và kinh phí. Các thuật toán tái chia mẫu/chia lại điểm ảnh (resampling) cũng được nhiều nghiên cứu cho thấy có thể tăng độ mức độ chi tiết cũng như độ chính xác lên một mức độ nhất định. Tuy nhiên, các thuật toán này còn một số nhược điểm và gây ra một số sai số có tính chất hệ thống trong các DEM được tái chia mẫu.

Trong một số nghiên cứu trước đây, mô hình mạng neuron Hopfield (HNN) đã được thay đổi và ứng dụng trong một số bài toán như bài toán làm mịn và tăng cường độ phân giải không gian của các ảnh viễn thám đa phổ [52]. Vì các ảnh viễn thám và các DEM dạng grid đều có cấu trúc dữ liệu raster nên có thể kỳ vọng rằng các phương pháp sử dụng mô hình mạng neuron được phát triển cho các ảnh viễn thám có thể được cải tiến để tăng độ chính xác cũng như mức độ chi tiết của các DEM dạng grid bằng cách chia nhỏ các điểm ảnh trên DEM độ phân giải thấp thành các điểm ảnh nhỏ và độ cao của các điểm ảnh nhỏ này được xác định bằng phương pháp khuếch đại sự phụ thuộc không gian. Trong chương này đề xuất một phương pháp mới để tăng độ phân giải thông qua chia nhỏ các điểm ảnh đồng thời áp dụng nguyên lý tối đa hóa sự phụ thuộc không gian để làm cho độ cao của các pixel lân cận nhau sẽ có xu hướng có giá trị gần nhau. Thuật toán này được đề xuất dựa trên mô hình thuật toán mạng neuron Hopfield đã áp dụng để làm trơn ảnh viễn thám thông qua tái chia mẫu kết hợp với nguyên lý tối đa hóa sự tương quan không gian và các điều kiện ràng buộc. Do tính chất và đặc điểm của dữ liệu, các hàm mục tiêu đại diện cho sự tương quan không gian và hàm điều kiện ràng buộc cũng cần thay đổi cho phù hợp với bài toán tăng độ phân giải không gian, tăng độ chính xác của grid DEM. Chính vì thế, việc sử dụng mạng neuron Hopfield để tăng độ phân giải không gian của mô hình số độ cao DEM dạng grid sẽ tập trung vào việc xây dựng các hàm này trong hàm tối ưu, thử nghiệm và đánh giá, so sánh kết quả của thuật toán với các thuật toán phổ biến để sử dụng tái chia mẫu đang được thực hiện hiện nay.

Mô hình mạng neuron Hopfield để nâng cao độ chính xác mô hình số độ cao DEM sẽ là biến thể của mô hình mạng nơ bởi Tatem và nnk [73]. Mô hình tăng độ phân giải không gian của một DEM dạng grid là một phiên bản phát triển từ mô hình mạng neuron Hopfield (HNN) được thiết kế cho thuật toán tăng độ phân giải [73], [50]. Trong mô hình HNN tăng độ phân giải không gian của các lớp phủ mặt đất (land cover), một điểm ảnh (pixel) trên ảnh viễn thám gốc được chia thành $m \times m$ các tiểu điểm ảnh hay các điểm ảnh con (sub-pixel) và mỗi tiểu điểm ảnh được đại diên bởi một neuron trong HNN. Mô hình này hoạt động dựa trên một hàm điều kiên và hai hàm mục tiêu. Hàm điều kiên ràng buộc ở đây là tổng số các tiểu điểm ảnh của mỗi một lớp phủ bề mặt phải bằng số lượng các tiểu điểm ảnh của các lớp được xác định bằng giá trị phần trăm của lớp phủ từ kết quả phân loại mềm. Các hàm mục tiêu đóng vai trò tối đa hoá sự phụ thuộc không gian của các tiểu điểm ảnh trong phạm vi một điểm ảnh gốc. Như vậy, các tiểu điểm ảnh của cùng một lớp phủ sẽ được sắp xếp cạnh nhau để tạo ra một bản đồ lớp phủ có mức liên kết không gian là lớn nhất. Sau khi chia nhỏ các điểm ảnh trên DEM gốc, độ phân giải không gian của DEM sẽ được tăng lên *m* lần. Tỷ lệ *m* được định nghĩa là tỷ lệ giữa độ phân giải ban đầu và độ phân giải cuối cùng theo các hàng và các cột, có thể gọi giá trị m là độ phóng đại. Quá trình tăng độ phân giải còn được mô tả bằng một thuật ngữ tiếng Anh là downscaling trong đó mỗi điểm ảnh trên DEM gốc được chia thành $m \times m$ điểm ảnh nhỏ hơn trên DEM mới.

3.2 Mạng neuron Hopfield ứng dụng cho siêu phân giải bản đồ (superresolution mapping/sub-pixel mapping)

3.2.1 Xây dựng mô hình

Mạng neuron Hopfield là mạng được kết nối hoàn toàn có hồi quy và chúng phần lớn được sử dụng cho việc liên kết tự động và tối ưu hoá [64]. Một mạng đơn giản của 5 neuron có thể được nhìn thấy như Hình 3-1 [64].



Hình 3-1. Mạng neuron Hopfield 5 nút ([64]

Mạng nơ ron trên sẽ đạt được trạng thái ổn định khi hàm năng lượng E, có dạng như công thức (3.1) đạt giá trị cực tiểu.

$$E = -\frac{1}{2} \sum_{i} \sum_{j} w_{ij} v_{i} v_{j} - \sum_{i=1} v_{i} I_{i}$$
(3.1)

Ở đây: w_{ij} là trọng số kết nối từ neuron j tới i

- v_i và v_j là các đầu ra của neuron i và j
- I_i là đường chéo bên ngoài hướng về phía neuron i.

Để tối ưu hoá những kết quả, hàm năng lượng được tính toán dựa vào những hàm mục tiêu và những điều kiện ràng buộc, mà có thể được định nghĩa :

Năng lượng = Mục tiêu + Điều kiện
$$(3.2)$$

Cho nhiệm vụ giải quyết siêu phân giải bản đồ lớp phủ, mạng neuron Hopfield được sử dụng như là một công cụ tối ưu hoá được giới thiệu bởi Tatem *et al.* [73]. Giá trị phần trăm của mỗi lớp từ một phép phân loại mềm được sử dụng như những sự ràng buộc để tối thiểu hàm năng lượng của mạng.

3.2.2 Thiết lập các hàm mục tiêu và điều kiện

Tatem *et al.* [73] sử dụng mạng neuron Hopfield cho sự nhận dạng mục tiêu từ ảnh viễn thám. Bản mô tả của một mạng neuron Hopfield được thiết kế cho siêu phân giải bản đồ được trình bày trong nghiên cứu của Nguyen Quang Minh, 2011 [24].

Trong mô hình này, mỗi tiểu pixel (i,j) được đại diện bởi một neuron trong mạng neuron Hopfield. Mỗi pixel (x, y) trong ảnh của bộ cảm vệ tinh được đại diện bởi một ma trận của (x, y) các tiểu pixel có một trung tâm tại (xz + int(z/2), yz + int(z/2)), ở đây *int* là một giá trị nguyên (z là hệ số phóng).

Hàm năng lượng *E* được định nghĩa như sau:

$$E = -\sum_{i} \sum_{j} (k_1 G \mathbf{1}_{ij} + k_2 G \mathbf{2}_{ij} + k_3 P_{ij})$$
(3.3)

Ở đây k_1 , k_2 và k_3 là các hằng số trọng số, $G1_{ij}$ và $G2_{ij}$ là các giá trị đầu ra của neuron (i,j) được đóng góp bởi hai hàm mục tiêu, và P_{ij} là giá trị đầu ra được đóng góp bởi điều kiện ràng buộc. Hàm mục tiêu đầu tiên $G1_{ij}$ cho phép tăng giá trị đầu ra của neuron trung tâm v_{ij} tới giá trị 1 nếu đầu ra trung bình của 8 neuron phụ cận lớn hơn giá trị 0.5.

$$\frac{dG1_{ij}}{dv_{ij}} = \frac{1}{2} \left(1 + \tanh\left(\frac{1}{8} \sum_{k=i-1(k\neq j)}^{i+1} \sum_{l=j-1(l\neq j)}^{j+1} v_{kl} - 0.5\right) \lambda\right) \times (v_{ij} - 1)$$
(3.4)

Ở đây λ là hằng số, chúng điều khiển độ dốc (độ hội tụ) của hàm *tanh*. Hàm *tanh* điều khiển ảnh hưởng của các neuron lân cận. Nếu đầu ra trung bình của các neuron lân cận nhỏ hơn 0.5 thì trên hàm năng lượng sẽ không có ảnh hưởng. Nếu

trung bình của hàm phụ cận lớn hơn 1 và đầu ra của neuron v_{ij} nhỏ hơn 1, thì biểu đồ đầu ra là số âm và nó phụ thuộc vào sự tăng thêm neuron đầu ra.

Hàm mục tiêu thứ hai nhằm vào sự giảm bớt đầu ra của neuron trung tâm tới $v_{ij} = 0$ nếu đầu ra trung bình của 8 neuron phụ cận nhỏ hơn 0.5.

$$\frac{dG2_{ij}}{dv_{ij}} = \frac{1}{2} \left(1 - \tanh\left(\frac{1}{8} \sum_{k=i-1(k\neq j)}^{i+1} \sum_{l=j-1(l\neq j)}^{j+1} v_{kl}\right) \lambda \right) \times v_{ij}$$
(3.5)

Lúc này, hàm *tanh* ước lượng tới giá trị 0 nếu trung bình đầu ra của các neuron phụ cận lớn hơn 0.5. Nếu trung bình đầu ra của các neuron phụ cận nhỏ hơn 0.5 thì hàm *tanh* ước lượng tới 1 và neuron trung tâm đầu ra v_{ij} nhỏ hơn 1, thì biểu đồ đầu ra là số dương và nó đòi hỏi vào sự giảm bớt neuron đầu ra. Giá trị của những đầu ra neuron trở nên ổn định nếu chỉ một trong số những điều kiện này được thoả mãn và $G1_{ij} + G2_{ij} = 0$.

Hai hàm mục tiêu của mạng neuron Hopfield đóng vai trò là sự cực đại hoá sự phụ thuộc không gian trong siêu phân giải bản đồ. Bằng việc điều chỉnh những giá trị của hai hàm mục tiêu, hàm năng lượng của mạng neuron Hopfield là cực tiểu hoặc các tiểu pixel liền kề sẽ có giá trị lớp phủ giống nhau.

Tuy nhiên, nếu chỉ sử dụng các hàm mục tiêu thì kết quả thu được là tất cả các giá trị đầu ra neuron nhận giá trị 1 hoặc 0, điều đó đồng nghĩa với việc tất cả các tiểu phần tử sẽ thuộc về một lớp nào đó. Bởi vậy, cần có một điều kiện ràng buộc sao cho số lượng các tiểu phần tử thuộc về một lớp phủ trong mỗi điểm ảnh sẽ bằng số phần trăm tương ứng với lớp đó được xác định bằng phân loại mềm. Điều kiện ràng buộc P_{ij} nhằm giữ lại những tỷ lệ lớp pixel của những ảnh phân loại mềm. Điều kiện ràng buộc có thể được miêu tả như sau:

$$\frac{dP_{ij}}{dv_{ij}} = \frac{1}{2z^2} \left(\sum_{k=xz}^{xz+z} \sum_{l=yz}^{yz+z} (1 + \tanh(v_{kl} - 0.55)\lambda) \right) - a_{xy}$$
(3.6)

Nếu tỷ lệ vùng của sự ước lượng cho pixel gốc (x,y) là thấp hơn hoặc lớn hơn vùng mục tiêu, thì những giá trị đầu ra của các neuron sẽ được tăng lên hoặc giảm bớt để giải quyết vấn đề. Chỉ khi sự ước lượng tỷ lệ vùng là đồng nhất với tỷ lệ

vùng mục tiêu cho mỗi pixel, thì biểu đồ sẽ đạt giá trị "0", tương ứng với $P_{ij} = 0$ trong hàm năng lượng. Các hằng số trọng số k_1 , k_2 và k_3 được sử dụng để xếp loại sự quan trọng của các phần tử trong phương trình (3.3). Tatem và cộng sự [73] đã đề xuất rằng những giá trị của $k_1 = k_2 = 150$ và $k_3 = 130$ là thích hợp cho sự xác định mục tiêu lớp phủ.

Trong mô hình HNN để tăng độ phân giải lớp phủ, đầu ra v_{ij} của một neuron (một tiểu điểm ảnh) (*i*, *j*) là:

$$v_{ij} = g(u_{ij}) = \frac{1}{2} (1 + tanh\lambda u_{ij})$$
 (3.7)

Trong đó: $g(u_{ij})$ là hàm kích hoạt của mỗi neuron, u_{ij} là giá trị đầu vào của mỗi neuron và λ là độ hội tụ của hàm *tanh*.

Giá trị đầu vào u_{ij} được xác định tại thời điểm t như sau:

$$u_{ij}(t) = u_{ij}(t - dt) + \frac{du_{ij}}{dt}dt$$
(3.8)

Trong đó dt là bước thời gian, $u_{ij}(t - dt)$ là giá trị đầu ra tại thời điểm (t - dt)và du_{ij}/dt được định nghĩa như sau:

$$\frac{du_{ij}}{dt} = \frac{dE_{ij}}{dv} \tag{3.9}$$

Trong công thức (3.5), E là hàm năng lượng, được định nghĩa là

$$E = Mục tiêu + Điều kiện và$$

$$\frac{dE_{ij}}{dv} = \left(\sum_{e}^{K} \frac{dGoal_{e}}{dv} + \frac{dConstraint}{dv}\right)$$
(3.10)

Trong đó, *K* là số hàm mục tiêu. Tùy thuộc vào từng ứng dụng cụ thể, hàm mục tiêu và hàm điều kiện ràng buộc có thể được sửa đổi để tối ưu hóa. Ví dụ, trong [73], việc sử dụng mạng neuron Hopfield cho phân giải bản đồ lớp phủ đã sử dụng hàm điều kiện ràng buộc là các giá trị phần trăm lớp phủ từ kết quả phân loại mềm và hàm mục tiêu là hàm có mục đích làm cho các tiểu điểm ảnh cạnh nhau sẽ có cùng nhãn lớp phủ.

Mạng HNN trong các trường hợp trên sẽ chạy cho đến khi tổng năng lượng E của mạng đạt đến một giá trị cực tiểu được xác định là:

$$E = \sum_{i} \sum_{j} \left(\sum_{f} \left(k_{f} v_{ij}^{Goal_{f}} \right) + \sum_{g} \left(k_{g} v_{ij}^{Constraint_{g}} \right) \right) = min$$
(3.11)

Hoặc:

$$E(t) - E(t - dt) = 0$$
(3.12)

3.3 Xây dựng thuật toán nâng cao độ chính xác của mô hình số độ cao dạng grid

3.3.1 Xây dựng mô hình mạng neuron Hopfield nhằm tăng độ phân giải không gian của mô hình số độ cao DEM dạng grid

Để sử dụng mô hình mạng neuron Hopfield tăng độ phân giải của DEM dạng grid, chúng ta sẽ chia một pixel trong DEM gốc ở độ phân giải thấp có kích thước pixel lớn thành $m \times m$ pixel con, mỗi pixel con được đại diện bởi một neuron trong HNN và giá trị độ cao chính là trạng thái đầu ra (output) của các neuron trong mạng neuron Hopfield. Giá trị đầu ra cũng chính là giá trị độ cao của mỗi neuron (tiểu điểm ảnh) sẽ được xác định thông qua hàm mục tiêu đảm bảo giá trị semi-variogram giữa các neuron lân cận tiến về giá trị nhỏ nhất. Ngoài ra, các giá trị độ cao của mỗi pixel con được ràng buộc bởi hàm điều kiện là giá trị trung bình độ cao của các pixel con nằm trong phạm vi của một pixel trong DEM gốc phải bằng giá trị độ cao của pixel trên DEM gốc.

Ý tưởng của phương pháp mới dựa trên giả thiết rằng độ cao của mỗi pixel con phải gần bằng hoặc có xu thế gần giống với độ cao của các pixel con liền kề (giả thiết về sự phụ thuộc không gian). Việc xác định sự phụ thuộc không gian trong trường hợp này được xác định thông qua giá trị semi-variance được định nghĩa là:

.....

$$\gamma(h) = \frac{1}{2N(h)} \sum_{1}^{N(h)} \left[v_{ij} - v_{ij+h} \right]^2$$
(3.13)

Trong đó $\gamma(h)$ là giá trị của hệ số semi-variogram ở bước nhảy khoảng cách h, h là khoảng cách giữa một cặp điểm pixel con có giá trị độ cao lần lượt là: v_{ij} và v_{ij+h} , và N(h) là số cặp điểm cách nhau một khoảng cách h.

Hàm semi-variogram nói trên được sử dụng nhiều trong công tác trắc địa, đặc biệt là khi sử dụng để nội suy theo thuật toán Kriging. Hàm semi-variogram thể hiện quy luật về sự khác biệt về độ cao giữa các cặp điểm theo khoảng cách giữa các cặp điểm này.

Sự phụ thuộc không gian ở đây được định nghĩa là sự giống nhau về mặt giá trị giữa các cặp điểm có khoảng cách gần nhau, có nghĩa là giá trị semi-variogram $\gamma(h)$ sẽ nhỏ khi khoảng cách h giữa hai điểm (i, j) và (i, j + h) nhỏ. Đối với mô hình DEM được tăng độ phân giải, nếu giữa các pixel con có sự phụ thuộc không gian, thì hệ số semi-variance sẽ nhỏ ở bước nhảy h nhỏ. Điều này có nghĩa là khi hệ số semi-variogram đạt cực tiểu thì chức năng tối đa hóa sự phụ thuộc không gian trong mô hình HNN mới này sẽ tăng hoặc giảm giá trị đầu ra của pixel con nằm ở vị trí trung tâm cho tới khi bằng giá trị độ cao trung bình của tám pixel con xung quanh.

Để tìm được giá trị cực tiểu của hàm $\gamma(h)$ cần sử dụng giá trị đạo hàm của hàm này. Giá trị cực tiểu của semi-variogram được xác định theo đạo hàm như sau:

$$\frac{\partial \gamma(h)}{\partial \nu} = 0 \tag{3.14}$$

$$\operatorname{Va} \frac{\partial \gamma(h)}{\partial v} = \frac{\partial \left(\frac{1}{2N(h)} \sum_{1}^{N(h)} (2v_{ij} - 2v_{ij+h})\right)}{\partial v} = v_{ij} - \frac{\sum_{1}^{N(h)} v_{ij+h}}{N(h)} = 0$$

Vậy, từ đó suy ra:

$$v_{ij}^{expected} = \frac{\sum_{1}^{N(h)} v_{ij+h}}{N(h)}$$
(3.15)

Đối với mạng neuron Hopfield, trong trạng thái khi chưa cân bằng thì giá trị đầu ra của mỗi neuron (chính là giá trị độ cao của các tiểu điểm ảnh) sẽ không bằng được giá trị đầu ra (độ cao) mong muốn. Do vậy, cần phải thay đổi giá trị độ cao để đạt được độ cao mong muốn. Giá trị độ cao được thay đổi như sau:

$$du_{ij}^{sd} = v_{ij}^{expected} - v_{ij} \tag{3.1}$$

Điều này có nghĩa là giá trị độ cao của pixel con nằm ở giữa có độ cao là v_{ij} sẽ bằng giá trị độ cao trung bình của các pixel con xung quanh với bước nhảy h ($v_{ij + h}$). Trong mô hình tăng độ phân giải của grid DEM, các pixel con có bước nhảy nhỏ nhất nghĩa là người ta sẽ sử dụng 8 điểm xung quanh pixel con v_{ij} .





Trong Hình 3-2 biểu thị một ví dụ về mô hình mới được đề xuất để tăng kích thước điểm ảnh của một DEM dạng grid với kích thước 2 x 2 pixel. Một pixel trong DEM gốc được chia thành 4 × 4 pixel con trong DEM mới (hệ số thu phóng m = 4). Vì vậy, từ một DEM gốc kích thước 2 × 2 được tái chia mẫu thành một DEM gồm 8 × 8 pixel con. Mỗi pixel con được đại diện bởi một neuron trong mô hình HNN và có giá trị ban đầu là giá trị độ cao của pixel trong DEM gốc (hoặc có thể được gán ngẫu nhiên). Độ cao giả lập của pixel con sau khi thực hiện tối đa hóa sự phụ thuộc không gian được tính bằng cách sử dụng một cửa sổ 3 × 3 và giá trị độ cao của pixel con nằm giữa bằng giá trị độ cao trung bình của 8 pixel con xung quanh.

Nếu hàm để tối đa hóa sự không gian phụ thuộc không gian là hàm duy nhất được sử dụng trong mô hình thì độ cao của tất cả các pixel con trong DEM mới (sau khi tăng độ phân giải) cuối cùng sẽ giống nhau và như vậy độ cao của mô hình DEM gốc sẽ không được giữ lại. Để giải quyết vấn đề này, cần sử dụng một *hàm điều kiện để ràng buộc*. Nguyên tắc của hàm này là độ cao trung bình của tất cả các pixel con nằm trong một pixel gốc phải bằng giá trị độ cao của pixel đó trong DEM gốc. Ví dụ, giá trị độ cao trung bình của tất cả các pixel con trong pixel (1,1) của DEM gốc trong Hình 3-2 phải bằng độ cao của pixel (1,1).

$$du_{ij}^{ep} = Elevation_{x,y} - \frac{\sum_{(x-1)\times m}^{x\times m} \sum_{(y-1)\times m}^{y\times m} v_{pq}}{m \times m}$$
(3.27)

Trong đó, Elevation_{xy} là giá trị độ cao của pixel (x, y) trong DEM gốc, v_{pq} là pixel con (p, q) nằm trong pixel (x, y) trong DEM mới và m là hệ số thu phóng. Nếu giá trị độ cao của tất cả các pixel con trong một pixel nhỏ hơn giá trị *Elevation_{xy}* thì một giá trị được thêm vào giá trị độ cao v_{pq} của tất cả các pixel con thuộc pixel (x, y). Ngược lại thì một giá trị được bớt đi từ giá trị đầu ra v_{pq} của neuron (p, q).

Sau đó, một giá trị đầu vào của mỗi neuron (pixel con) được tính dựa trên công thức (3.8) với giá trị du_{ij}/dt là:

$$\frac{du_{ij}}{dt} = \frac{dE_{ij}}{dv} = du_{ij}^{sd} + du_{ij}^{ep}$$
(3.38)

Từ đây có thể tính được giá trị của hàm năng lượng E của toàn bộ mạng neuron Hopfield tại thời điểm *t* là:

$$E = \sum_{i} \sum_{j} \left(du_{ij}^{sd} + du_{ij}^{ep} \right)$$
(3.49)

Giá trị đầu ra v_{ij} của mỗi neuron được tính bằng cách sử dụng hàm kích hoạt $g(u_{ij})$. Tuy nhiên, trong mô hình mới này, chức năng kích hoạt $g(u_{ij})$ không giống như trong công thức (3.21) vì nó không được sử dụng để đẩy giá trị đầu ra của neuron lên 0 hoặc 1 như trong trường hợp tăng độ phân giải bản đồ lớp phủ. Thay vào đó, một hàm kích hoạt tuyến tính được trình bày trong nghiên cứu của Tank và Hopfield [3] đã được sử dụng trong cách tiếp cận mới này như sau:

$$v_{ij} = g(u_{ij}) = a \times u_{ij} + b \tag{3.20}$$

mô hình này, các tham số a = 1 và b = 0.

Quá trình hiệu chỉnh như trên sẽ thực hiện lặp nhiều lần. Sau mỗi lần hiệu chỉnh thì hàm năng lượng *E* được tính lại và so sánh với lần hiệu chỉnh trước. Mạng HNN sẽ chạy cho đến khi hàm năng lượng *E* đạt cực tiểu:

$$E = \sum_{i} \sum_{j} \left(du_{ij}^{sd} + du_{ij}^{ep} \right) = min$$
(3.21)

Hoặc, E(t)-E(t-dt) = 0, trong đó (t - dt) và t là hai lần lặp liên tiếp của mạng Hopfield.

3.3.2 Sơ đồ khối của thuật toán

Trên Hình 3-3 là sơ đồ khối của mô hình HNN đã sửa đổi cho thuật toán tăng độ phân giải không gian của DEM, lấy ví dụ minh họa về tăng độ phân giải DEM từ độ phân giải 20m ban đầu lên độ phân giải không gian 5m.

Đầu tiên, DEM gốc độ phân giải thấp (20m) được tái chia mẫu thành DEM độ phân giải 5m theo phương pháp tái chia mẫu đơn giản như phương pháp Dựa vào điểm lân cận gần nhất (Nearest Neighbor) để cung cấp dữ liệu khởi tạo cho mạng HNN. Mỗi pixel con trong DEM 5m này tương ứng với một neuron trong mô hình HNN và giá trị đầu vào khởi tạo cho các neuron này là các giá trị độ cao của các pixel con. Các giá trị đầu ra của các neuron trong mạng HNN (các giá trị độ cao. Hàm mục tiêu ở đây nhằm giảm thiểu giá trị của semi-variance) và hàm điều kiện là sự trùng khớp về độ cao so với DEM độ phân giải gốc ban đầu. Tổng các giá trị hàm năng lượng của mạng không đạt giá trị cực tiểu, thì sau đó giá trị cực tiểu này sẽ là giá trị của hàm năng lượng mới và mạng HNN sẽ chạy cho lần lặp tiếp theo. Quá trình lặp sẽ tiếp tục cho đến khi giá trị hàm năng lượng của hai lần lặp liên tiếp không thay đổi và giá trị đầu ra cuối cùng của mỗi neuron trong mạng HNN là độ cao của pixel con tương ứng với neuron này.





- 3.3.3 Thiết kế chương trình tăng độ phân giải không gian và nâng cao độ chính xác của grid DEM sử dụng mạng neuron Hopfiled
 - <u>Phát biểu bài toán</u>: Cho một dữ liệu DEM dạng grid có độ phân giải không gian thấp, kích thước grid lớn, sử dụng thuật toán mạng neuron Hopfiled để làm tăng độ phân giải không gian và độ chính xác của grid DEM đã có.
 - <u>Ánh xa lên bài toán lên mang neuron Hopfiled:</u>
 Để ánh xạ bài toán tối ưu lên mạng neuron Hopfiled, chúng ta cần đưa ra các trọng số và các ràng buộc của bài toán [83].
 - Từ một pixel kích thước lớn trong grid DEM gốc ở độ phân giải thấp, chia thành *m x m* pixel con. Mỗi pixel được đại diện bởi một neuron trong mạng neuron Hopfield ứng với đầu ra là v_{ij}. Như vậy, số neuron trong mạng chính bằng số pixel con và bằng *m*².

 Mỗi phương án của bài toán ứng với một giá trị semi-variance được định nghĩa là:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{1}^{N(h)} \left[v_{ij} - v_{ij+h} \right]^2$$
(3.22)

Trong đó: $\gamma(h)$ là giá trị của hệ số semi-variogram ở bước nhảy khoảng cách h, h là khoảng cách giữa một cặp điểm pixel con có giá trị độ cao lần lượt là: v_{ij} và v_{ij+h} , và N(h) là số cặp điểm cách nhau một khoảng cách h.

Ta cần phải giải bài toán tìm giá trị cực tiểu của hàm γ(h) với các ràng buộc sau:

+ Giá trị độ cao của pixel con nằm ở giữa có độ cao là v_{ij} sẽ bằng giá trị độ cao trung bình của các pixel con xung quanh với bước nhảy h (v_{ij+h}):

$$v_{ij}^{expected} = \frac{\sum_{1}^{N(h)} v_{ij+h}}{N(h)}$$
(3.23)

+ Giá trị độ cao được thay đổi như sau:

$$du_{ij}^{sd} = v_{ij}^{expected} - v_{ij} \tag{3.24}$$

 Giá trị của hàm năng lượng là tổng tất cả các điều chỉnh của tất cả các neuron:

$$E = \sum_{i} \sum_{j} \left(du_{ij}^{sd} + du_{ij}^{ep} \right) = min$$
(3.25)

- Từ đây, lấy đạo hàm riêng $\frac{dE_{ij}}{dv}$, ta xác định được phương trình động học của các neuron:

$$\frac{dE_{ij}}{dv} = \frac{du_{ij}}{dt} = du_{ij}^{sd} + du_{ij}^{ep}$$
(3.26)

- Chọn hàm kích hoạt của các neuron là hàm y = f(x).
- Như vậy ta đã ánh xạ được bài toán tăng độ phân giải của grid DEM lên mạng neuron Hopfield. Hay nói cách khác là mạng Hopfield đã được thiết kế để tối ưu hóa hàm *E* theo (3.25) với điều kiện trọng số và đầu ra của mạng xác định cụ thể.

 Mã giả của chương trình tăng độ phân giải không gian và nâng cao độ chính xác của grid DEM sử dụng mạng neuron Hopfield như sau:

{# Định nghĩa các hàm:

Hàm mở file

Hàm tính giá trị của hàm Spatial dependence maximization (Hàm tối ưu hóa sự phụ thuộc không gian)

Hàm điều kiện ràng buộc về độ cao, Elevation Function;

Hàm xác định kích thước của DEM gốc và DEM được tăng độ phân giải

Hàm tính độ cao trung bình

Hàm tính giá trị hiệu chỉnh của hàm độ cao cho từng pixel con

Hàm ghi kết quả ra file.

Nhập các dữ liệu đầu vào:

Tạo đối tượng str1 chỉ đến file độ cao

Lấy các tham số vị trí của DEM gốc

Lấy thông tin phép chiếu của DEM gốc

Gán chọn Hệ số thu phóng (Zoom factor).

Xác lập thông tin vị trí cho DEM được chia nhỏ bằng HNN.

 # Khởi tạo giá trị độ cao ban đầu của DEM được chia nhỏ = giá trị độ cao trên DEM gốc.

Khởi tạo giá trị ban đầu của hàm năng lượng.

Khởi tạo giá trị ban đầu của ngưỡng mạng neuron Hopfiled (Threshold).

Khởi tạo giá trị của biến tính toán sự thay đổi sau mỗi chu trình của Hàm năng lượng (Enery_dif).

Khởi tạo biến chạy của chu trình (iteration).

While abs(Energy_dif) > Threshold:

Tính toán hàm mục tiêu là tối ưu hóa sự phụ thuộc không gian: độ cao của các pixel con phải bằng độ cao trung bình của các pixel con xung quanh nó.

Tính toán hàm điều kiện ràng buộc: trung bình độ cao của các pixel con trong một pixel to phải bằng độ cao của pixel to đó.

Hàm năng lượng mới = Tổng (abs(Hàm mục tiêu) + abs(Hàm điều kiện ràng buộc))

Hiệu chỉnh giá trị độ cao mới cho từng pixel con

Tính giá trị của Enery_dif (sự thay đổi của hàm năng lượng mới và hàm năng lượng cũ)

Iteration = iteration +1 (Tăng giá trị biến chạy của chu trình).

Khi hàm năng lượng đạt giá trị min (kết quả của hàm năng lượng E trong hai lần lặp liên tiếp giống nhau hoặc không thay đổi nhiều) thì:

In ra ("Vòng lặp số:", iteration; Giá trị hàm năng lượng mới nhất, Energy_new))

Tạo file kết quả của chương trình Downscaling - file DEM dạng raster.}

<u>Lập chương trình:</u>

Chương trình tăng độ phân giải không gian và nâng cao độ chính xác của mô hình số độ cao dạng grid sử dụng mạng neuron Hopfield được lập trình bằng ngôn ngữ Python tích hợp trên nền QGIS.





Đoạn code của chương trình được giới thiệu trong phần Phụ lục 1.

3.4 Thực nghiệm tăng độ phân giải không gian và độ chính xác của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng neuron Hopfield (mô hình HNN đã biến đổi)

3.4.1 Dữ liệu thực nghiệm

Để có thể dễ dàng so sánh hiệu quả của thuật toán với các phương pháp tái chia mẫu được sử dụng nhiều như song tuyến Bilinear, Bi-cubic và Kriging, trong luận án này sẽ sử dụng cùng các bộ dữ liệu đã được sử dụng để đánh giá các thuật toán này ở Chương 2. Trong 4 bộ dữ liệu thực nghiệm đó, có 2 bộ dữ liệu DEM được xây dựng trên nguyên tắc giảm độ giải và 2 bộ dữ liệu DEM được xây dựng từ dữ liệu thực tế. Hai bộ dữ liệu được xây dựng bằng nguyên tắc giảm độ phân giải từ dữ liệu chuẩn là Bộ dữ liệu D1 và D2 ở khu vực Yên Thành, Nghệ An. Dữ liệu DEM chuẩn của bộ dữ liệu D1 có độ phân giải 20m, sau đó được giảm độ phân giải xuống 60m để sử dụng làm đầu vào cho các mô hình tăng độ phân giải bằng các thuật toán song tuyến, Bi-cubic, Kriging và HNN (Hình 3-5).



Hình 3-5. Tăng độ phân giải không gian của DEM từ độ phân giải 60m lên 20m

Trong đó: (a) DEM tham chiếu ở độ phân giải 20m, (b) DEM gốc ban đầu đã giảm độ phân giải ở độ phân giải 60m được sử dụng làm dữ liệu đầu vào cho các thuật toán), (c) DEM sau khi tăng độ phân giải lên 20m bằng thuật toán HNN, (d) DEM ở 20m sau khi tái chia mẫu sử dụng mô hình song tuyến, (e) DEM ở độ phân giải 20m sau khi tái chia mẫu sử dụng mô hình Bi-cubic, (f) DEM ở độ phân giải 20m sau khi tái chia mẫu bằng phương pháp Kriging.



Hình 3-6. Tăng độ phân giải không gian của DEM SRTM khu vực Yên Thành, Nghệ An từ độ phân giải 90m lên 30m (tập dữ liệu D2)

Trong đó: (a)-DEM tham chiếu ở độ phân giải 30m, (b)-DEM gốc ban đầu đã giảm độ phân giải ở độ phân giải 90m được sử dụng làm dữ liệu đầu vào cho các thuật toán, (c)-DEM sau khi tăng độ phân giải lên 30m bằng cách sử dụng mô hình HNN, (d)-DEM độ phân giải 30m sau khi tái chia mẫu sử dụng mô hình song tuyến, (e-DEM độ phân giải 30m khi tái chia mẫu sử dụng mô hình Bi-cubic, (f)-DEM độ phân giải 30m có được từ tái chia mẫu sử dụng Kriging.



Hình 3-7. Tăng độ phân giải không gian của DEM cho tập dữ liệu S1

Trong đó: (a)-Dữ liệu DEM tham chiếu ở độ phân giải 5m khu vực Mai Pha-Lạng Sơn, (b)-Dữ liệu DEM gốc ban đầu ở độ phân giải 20m được sử dụng làm dữ liệu đầu vào cho các thuật toán, (c)-DEM sau khi tăng độ phân giải lên 5m bằng cách sử dụng mô hình HNN, (d)-DEM độ phân giải 5m sau khi tái chia mẫu sử dụng mô hình song tuyến, (e)-DEM độ phân giải 5m sau khi tái chia mẫu sử dụng mô hình Bi-cubic, (f)- DEM độ phân giải 5m khi tái chia mẫu sử dụng mô hình Kriging.





Trong đó: (a)-Dữ liệu DEM tham chiếu ở độ phân giải 30m khu vực Đắc Hà-Kon Tum (được tạo từ dữ liệu đường bình độ với khoảng cao đều 5m), (b)-Dữ liệu DEM độ phân giải 90m (được tạo từ dữ liệu đường bình độ với khoảng cao đều 10m-dữ liệu đầu vào cho các thuật toán, (c)-DEM sau khi tăng độ phân giải lên 30m bằng thuật toán sử dụng mô hình HNN, (d)-DEM độ phân giải 30m sau khi tái chia mẫu theo mô hình song tuyến, (e)-DEM độ phân giải 30m sau khi tái chia mẫu theo mô hình song tuyến, sau khi nội suy Kriging.

Bộ dữ liệu D2 được xây dựng trên nguyên tắc giảm độ phân giải trong đó DEM chuẩn được cung cấp bởi USGS Earth Explorer có nguồn gốc là DEM thuộc dự án Mission Radar Topography Mission (SRTM) (Hình 3-6(a)). Bộ dữ liệu này bao phủ cùng một khu vực với bộ dữ liệu D1, nhưng với độ phân giải không gian là 30m. DEM này cũng đã được giảm độ phân giải xuống 90m để tạo ra một bộ dữ liệu thử nghiệm thứ hai cho mô hình HNN (Hình 3-6(b)).

Hai bộ dữ liệu thực nghiệm được xây dựng trên dữ liệu thực tế là bộ dữ liệu S1 và bộ dữ liệu S2. Trong đó S1 được thành lập từ dữ liệu điểm độ cao đo đạc mặt đất, sau đó được nội suy thành 2 bộ dữ liệu DEM gồm DEM 5m sử dụng nhiều điểm dữ liệu hơn và DEM 20m. DEM 5m được đánh giá độ chính xác theo tiêu chuẩn ASPRS của Mỹ, DEM đạt yêu cầu về độ chính xác. Thông tin đánh giá độ chính xác của DEM 5m này được trình bày trong Bảng 2-1 (Chương 2). Dữ liệu DEM 20m được thành lập từ tập hợp các điểm độ cao đã lược bỏ bớt điểm, được sử dụng làm đầu vào cho các thuật toán tái chia mẫu song tuyến, Bi-cubic, Kriging và HNN (Hình 3-7).

Bộ dữ liệu thực nghiệm S2 được thu thập và thành lập từ dữ liệu đường bình độ ở huyện Đắc Hà ở tỉnh Kontum, Việt Nam. Diện tích của khu vực thực nghiệm là khoảng 6,6km x 6,6km. DEM chuẩn để so sánh được xây dựng dựa trên đường đồng mức với khoảng cao đều 5m (Hình 2-13(a)). Sau đó, từ dữ liệu đường đồng mức được lọc bỏ với khoảng cao đều là 10m (Hình 2-13(b)) tạo ra DEM đầu vào có độ phân giải 90m (Hình 3-8). Từ DEM có độ phân giải 90m, sử dụng các thuật toán song tuyến, Bi-cubic, Kriging và HNN để tăng độ phân giải của DEM lên 30m với kết quả được trình bày trên Hình 3-8.

3.4.2 Kết quả thực nghiệm, phân tích độ chính xác

Việc đánh giá định lượng được thực hiện chủ yếu dựa trên tiêu chuẩn sai số trung phương cho toàn bộ các DEM và các mặt cắt của bốn bộ dữ liệu như được trình bày trong Bảng 3-1, trong đó có sự cải thiện về sai số trung phương của toàn bộ, các sai số trung phương lớn nhất và nhỏ nhất của các mặt cắt đã được thể hiện rõ. Cùng với các sai số trung phương, các hệ số hồi quy tuyến tính như độ dốc m, tham số chặn b và hệ số tương quan R đã được sử dụng để đánh giá sự phù hợp giữa các DEM sau khi được tăng độ phân giải và các DEM tham chiếu cho cả bốn bộ dữ liệu (Bảng 3-2).

Bộ dữ liệu		Phương pháp tái chia mẫu song tuyến	Phương pháp tái chia mẫu Bi-cubic	Phương pháp nội suy Kriging	Phương pháp dùng mô hình HNN	Độ chính xác được cải thiện so với p.p. tái chia mẫu song tuyến
	Tổng thể	3.3026	3.3716	2.8874	1.9853	39.9%
DEM 20m Nghệ An	Min MCD	2.5245	2.5619	2.4393	1.9124	24.25%
	Max MCD	3.3379	3.4256	3.2270	2.0171	39.57%
(D1)	Min MCN	1.3837	1.4051	1.3916	1.5229	10.06%
	Max MCN	3.7005	3.7997	3.7522	2.3575	36.29%
	Tổng thể	8.8105	8.8736	8.5719	8.3510	5.21%
DEM 30m Nghệ An	Min MCD	8.5013	6.8408	6.9101	6.9668	18.05%
	Max MCD	11.6961	10.7635	10.8141	11.0702	5.35%
(D2)	Min MCN	6.6352	6.4032	6.4005	6.2829	5.31%
	Max MCN	10.5144	9.8024	9.8357	9.6199	8.51%
	Tổng thể	1.5139	1.6000	1.2092	0.8493	43.90%
DEM 5m Lạng Sơn	Min MCD	1.1635	1.1821	1.0288	0.5102	56.15%
	Max MCD	1.6198	1.7805	1.4232	0.9587	40.81%
(S1)	Min MCN	1.1567	1.2101	0.7408	0.5897	49.02%
	Max MCN	1.6712	1.7451	1.6807	1.1155	33.25%
	Tổng thể	2.3284	2.4218	2.1095	2.0946	10.0%
DEM 30m	Min MCD	2.0938	1.0068	1.0624	0.9942	52.52%
Đắc Hà (S2)	Max MCD	4.0702	2.3434	2.4436	2.2330	45.14%
	Min MCN	2.8494	1.0181	1.0505	0.9603	66.29%
	Max MCN	4.6807	2.3473	2.4070	2.5440	45.65%

Bảng 3-1. Sai số trung phương của các phương pháp tái chia mẫu song tuyến, Bicubic, nội suy Kriging và phương pháp dùng mô hình mạng neuron Hopfiled HNN

Dê d <u>a</u> liên		Hệ số hồi quy tuyến tính			
	Độ đữ liệu	m	b	R ²	
DEM 20m Nghệ An (D1)	DEM 20m tăng độ phân giải sử dụng mô hình HNN	0.9973	0.2949	0.9981	
	DEM 20m sau khi tái chia mẫu theo mô hình song tuyến	0.9765	2.5368	0.9951	
	DEM 20m sau khi tái chia mẫu theo mô hình Bi-cubic	0.9781	2.3680	0.9948	
	DEM 20m sau khi nội suy Kriging	0.9832	1.8217	0.9962	
DEM 30m Nghệ An (D2)	DEM 30m tăng độ phân giải sử dụng mô hình HNN	0.9904	-1.6013	0.9686	
	DEM 30m sau khi tái chia mẫu theo mô hình song tuyến	0.9500	3.2057	0.9646	
	DEM 30m sau khi tái chia mẫu theo mô hình Bi-cubic	0.9529	2.8723	0.9639	
	DEM 30m sau khi nội suy Kriging	0.9608	1.9291	0.9694	
DEM 5m Lạng Sơn (S1)	DEM 5m tăng độ phân giải sử dụng mô hình HNN	1.0195	-5.908	0.9937	
	DEM 5m sau khi tái chia mẫu theo mô hình song tuyến	0.9601	12.3782	0.9793	
	DEM 5m sau khi tái chia mẫu theo mô hình Bi-cubic	0.9658	10.6432	0.9763	
	DEM 5m sau khi nội suy Kriging	0.945	16.3717	0.9884	
DEM 30m Đắc Hà (S2)	DEM 30m tăng độ phân giải sử dụng mô hình HNN	1.0043	-4.1179	0.9968	
	DEM 30m sau khi tái chia mẫu theo mô hình song tuyến	0.9872	12.1453	0.9960	
	DEM 30m sau khi tái chia mẫu theo mô hình Bi-cubic	0.9885	10.9118	0.9959	
	DEM 30m sau khi nội suy Kriging	0.9922	7.3917	0.9967	

Bảng 3-2. Các hệ số hồi quy tuyến tính cho cả bốn bộ dữ liệu D1, D2, S1, S2

Trong Bảng 3-2, SSTP tổng thể là sai số trung phương nói chung, Min MCD là sự cải thiện độ chính xác nhỏ nhất theo mặt cắt dọc của DEM, Max MCD là sự cải thiện độ chính xác lớn nhất theo mặt cắt dọc của DEM, Min MCN là sự cải thiện độ chính xác nhỏ nhất theo mặt cắt ngang của DEM, Max MCN là sự cải thiện độ chính xác lớn theo mặt cắt ngang của DEM.

3.4.3 Đánh giá trực quan

3.4.3.1 So sánh trực quan bằng mắt thường

So sánh trực quan cho thấy các DEM mặc dù các phương pháp song tuyến và Kriging đã cải thiện hình ảnh của DEM để khá giống với DEM chuẩn, cải thiện mức độ vỡ do độ phân giải thấp, kết quả sau khi tăng độ phân giải bằng mô hình HNN mới được đề xuất vẫn cải thiện hình ảnh DEM khá nhiều so với các phương pháp tái chia mẫu "kinh điển" này (tái chia mẫu sử dụng mô hình song tuyến, Bicubic, Kriging). Quan sát cho thấy sự cải thiện có thể nhìn thấy được ở kết quả của 4 bộ dữ liệu.

Sự cải thiện về độ tương tự giữa các hình ảnh trực quan của các DEM sau khi tăng độ phân giải và các DEM tham chiếu đã thấy rất rõ ràng khi so sánh giữa DEM 20m của bộ dữ liệu D1 (Hình 3-5), DEM 30m của bộ dữ liệu D2 (Hình 3-6), DEM 5m của bộ dữ liệu S1 (Hình 3-7) và 30m DEM của bộ dữ liệu S2 (Hình 3-8) với các DEM tham chiếu tương ứng. Hình ảnh của các DEM đầu vào và các DEM được tạo ra bằng các phương pháp "kinh điển", đặc biệt là các ảnh được tạo ra bằng phương pháp tái chia mẫu song tuyến, Bi-cubic, đã bị mờ đi do nhiễu và hình dạng của các đặc điểm địa hình trong các ảnh này trông bị méo.

Ngược lại, ảnh của các DEM sau khi tăng độ phân giải bằng thuật toán sử dụng mô hình HNN trong Hình 3-5(c), Hình 3-6(c), Hình 3-7(c) và Hình 3-8(c) trông không bị nhiễu và tương tự như các DEM tham chiếu trong Hình 3-5(a), Hình 3-6(a), Hình 3-7(a) và Hình 3-8(c). Sự cải thiện rõ ràng nhất trong việc tái cấu trúc hình dạng địa hình từ dữ liệu độ phân giải thô ban đầu được dùng làm dữ liệu đầu vào có thể được nhìn thấy trong các khu vực được đánh dấu trong Hình 3-5, Hình 3-6, Hình 3-7 và Hình 3-8.

3.4.3.2 Sử dụng các mặt cắt để đánh giá trực quan

So sánh giữa các bề mặt của các DEM kết quả, sử dụng các mặt cắt từ các vị trí được hiển thị trong Hình 3-9, cho thấy lợi thế rõ ràng của phương pháp tăng độ phân giải bằng thuật toán sử dụng mô hình HNN so với các phương pháp tái chia mẫu sử dụng các điểm nội suy.

Khi so sánh các mặt cắt của bốn bộ dữ liệu: D1 (Hình 3-10), D2 (Hình 3-11), S1 (Hình 3-12) và S2 (Hình 3-13), ta thấy các mặt cắt địa hình xác định từ các DEM sau khi tăng độ phân giải bằng thuật toán sử dụng mô hình HNN gần với các mặt cắt của các DEM chuẩn hơn so với của các DEM được tạo ra theo các phương pháp tái chia mẫu và điều này đúng với cả bốn bộ dữ liệu.

Sự cải thiện quan sát được rõ nhất trong bộ dữ liệu D1 trong Hình 3-10 (một số mặt cắt dọc và mặt cắt ngang ví dụ của bộ dữ liệu DEM Nghệ An 20m) ở những nơi như đỉnh đồi hoặc đáy thung lũng và bộ dữ liệu S1 trong Hình 3-12 (một số mặt cắt dọc và mặt cắt ngang ví dụ của bộ dữ liệu DEM Lạng Sơn 5m). Trong các ảnh này, có thể thấy rằng các bề mặt được tạo ra bằng các phương pháp tái chia mẫu gần với bề mặt DEM đầu vào (DEM NN) có độ phân giải thấp, trong khi bề mặt được tạo bởi DEM sau khi tăng độ phân giải bằng mô hình HNN gần với bề mặt DEM chuẩn (DEM gốc) hơn.



Hình 3-9. Vị trí của các mặt cắt để đánh giá độ chính xác của DEM Trong đó: (a)-Tập dữ liệu D1, (b)-Tập dữ liệu D2, (c)-Tập dữ liệu S1, (d)-tập dữ liệu S2.



Hình 3-10. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu DEM Nghệ An 20m (D1)



Hình 3-11. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu DEM Nghệ An 30m (D2)



Hình 3-12. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu DEM Lạng Sơn 5m (S1)



Hình 3-13. So sánh bề mặt tham chiếu (DEM gốc), DEM sau khi tăng độ phân giải bằng mô hình HNN (DEM HNN), DEM đầu vào (DEM NN), DEM sau tái chia mẫu song tuyến (DEM Bilinear), DEM sau tái chia mẫu Bi-cubic (DEM Bi-cubic), DEM sau nội suy Kriging (DEM Kriging) dựa trên các mặt cắt – đối với bộ dữ liệu DEM Đắc Hà 30m (S2)

Các kết quả trên cho thấy rằng: phương pháp tăng độ phân giải bằng mô hình HNN thực hiện chính xác hơn nhiều so với các phương pháp tái chia mẫu thông thường (Bilinear, Bi-cubic, Kriging) với các dạng địa hình đặc biệt, nơi có sự thay đổi lớn về độ cao như đỉnh của các dãy núi và đồi hoặc đáy thung lũng, đặc biệt là các thung lũng hình chữ V, các gờ và đồi có đỉnh nhọn. Phân tích có thể thấy là để tạo ra được bề mặt như vậy, nhờ tác động của hàm điều kiện ràng buộc độ cao, giúp giảm hoặc tăng độ cao tại các điểm đáy và đỉnh của các đối tượng địa hình, trong khi chức năng tối đa hóa sự phụ thuộc không gian (thông qua hàm mục tiêu sử dụng giá trị semi-variance tối thiểu) đảm bảo rằng độ cao của các pixel con liền kề thay đổi từ từ, như địa hình thực bên ngoài.

3.4.3.3 Đánh giá trực quan bằng biểu đồ phân tán

Các biểu đồ phân tán của các DEM kết quả sau khi tăng độ phân giải không gian cho thấy độ khớp hơn giữa dữ liệu DEM chuẩn và dữ liệu DEM sau khi tăng độ phân giải không gian bằng mô hình HNN so với các DEM kết quả của các phương pháp tái chia mẫu.

Sự giống nhau giữa hai tập dữ liệu có thể so sánh bằng biểu đồ phân tán như trong các hình: Hình 3-14 và Hình 3-15. Các so sánh cũng cho thấy các DEM kết quả của phương pháp tăng độ phân giải không gian bằng mô hình HNN trùng khớp với dữ liệu DEM gốc (DEM tham chiếu) hơn so với các DEM kết quả của các phương pháp tái chia mẫu song tuyến (Bilinear), Bi-cubic và Kriging.

Sự tăng độ chính xác có thể thấy rõ nhất đối với bộ dữ liệu DEM Nghệ An 20m (D1) ở và bộ dữ liệu DEM Lạng Sơn 5m (S1). Các biểu đồ phân tán ở Hình 3-14 và Hình 3-15 cho thấy hai bộ dữ liệu này là gần nhất với đường khớp tuyệt đối, các hệ số hồi quy (*m*, *b*) trong các biểu đồ phân tán này rất gần với các giá trị 1 và 0. So sánh hai bộ dữ liệu, các điểm dữ liệu trong các biểu đồ phân tán trong Hình 3-14(b) và Hình 3-15(b) (các DEM sau khi tái chia mẫu theo mô hình song tuyến), Hình 3-14(c) và Hình 3-15(c) (các DEM sau khi tái chia mẫu theo mô hình Bi-cubic), Hình 3-14(d) và Hình 3-15(b) (các DEM sau khi tái chia mẫu theo mô hình Bi-subic), Hình 3-14(d) và Hình 3-15(b) (các DEM sau khi tăng độ phân giải bằng thuật toán sử dụng mô hình HNN (Hình 3-14 (a), Hình 3-15(a)).



Hình 3-14. Các biểu đồ tán xạ của DEM tham chiếu độ phân giải không gian cao so với các DEM tăng độ phân giải của bộ dữ liệu DEM Nghệ An 20m (D1)



Hình 3-15. Các biểu đồ tán xạ của DEM tham chiếu độ phân giải không gian cao so với các DEM tăng độ phân giải của bộ dữ liệu DEM Lạng Sơn 5m (S1)
3.4.4 Đánh giá định lượng

3.4.4.1 Đánh giá định lượng dựa trên Sai số trung phương (RMSE)

Kết quả đánh giá định lượng hoàn toàn tương đồng với kết quả đánh giá trực quan. Việc đánh giá định lượng dựa trên sai số trung phương (Bång 3-1) đã cho thấy phương pháp tăng độ phân giải không gian bằng mô hình HNN cho độ chính xác cao hơn so với các phương pháp tái chia mẫu thông thường cho cả bốn bộ dữ liệu. Trong hai bộ dữ liệu giảm độ phân giải D1 và D2, mức tăng độ chính xác cao hơn cho bộ dữ liệu D1. Sai số trung phương của DEM sau khi tăng độ phân giải không gian bằng mô hình HNN là 1.9853m, trong khi sai số trung phương cho các phương pháp tái chia mẫu song tuyến, Bi-cubic và Kriging lần lượt là 3,3716m, 3,3716m và 2,8874m.

So sánh với sai số trung phương của dữ liệu DEM 20m thu được bằng phương pháp tái chia mẫu theo mô hình song tuyến từ dữ liệu 60m của Nghệ An (bộ dữ liệu D1), sai số trung phương của DEM 20m thu được sau khi tăng độ phân giải không gian bằng mô hình HNN đã giảm đáng kể là 39,9%. Đối với bộ dữ liệu D2 (DEM SRTM tăng độ phân giải từ 90m lên 30m), mức tăng độ chính xác sau khi sử dụng thuật toán tăng độ phân giải không gian bằng mô hình HNN không lớn như kết quả của các bộ dữ liệu khác nhưng vẫn khá thuyết phục với sai số trung phương giảm 5% so với DEM 30m thu được sau khi tăng độ phân giải từ DEM 90m theo phương pháp tái chia mẫu theo mô hình song tuyến.

Thực nghiệm trên bộ dữ liệu DEM thực cũng cho thấy mức độ tăng độ chính xác tương tự như đối với các DEM giảm độ phân giải. Sai số trung phương của dữ liệu DEM 5m Lạng Sơn (bộ dữ liệu S1) đã giảm mạnh đối với DEM tăng độ phân giải theo mô hình HNN xuống còn 0,8493m, từ 1,5139m đối với DEM 5m thu được sau khi tái chia mẫu theo mô hình song tuyến (43,9%), 1,6m đối với DEM thu được sau khi tái chia mẫu theo mô hình Bi-cubic và 1,2092m cho DEM thu được sau khi nội suy Kriging. Kết quả cho bộ dữ liệu kiểm tra S2 không ấn tượng như bộ dữ liệu S1. Tuy nhiên, việc tăng độ chính xác của DEM cũng rất có ý nghĩa với sai số trung phương giảm 5,2% so với phương pháp tái chia mẫu theo mô hình song tuyến.

Các thống kê này cho thấy, phương pháp sử dụng mô hình HNN đề xuất có thể tăng độ chính xác xác định độ cao của các pixel con của một grid DEM bằng cách tăng độ phân giải không gian. Hơn nữa, mô hình HNN sử dụng trong các thực nghiệm trên dường như hoạt động hiệu quả hơn với các DEM độ phân giải không gian trong khoảng 20m đến 5m.

Đánh giá cho từng khu vực có địa hình khác nhau trên DEM thông qua các mặt cắt cho thấy sự ảnh hưởng của các thuộc tính địa hình đến hiệu quả của thuật toán tăng độ phân giải không gian sử dụng mô hình HNN. Đối với bộ dữ liệu D1 độ phân giải 20m và bộ dữ liệu D2 độ phân giải 30m Nghệ An, mức tăng độ chính xác giữa các DEM được tái chia mẫu và giảm độ phân giải là tương đối đều trên các mặt cắt. Đối với bộ dữ liệu DEM thực 30m (tập dữ liệu S2), mức tăng độ chính xác của hầu hết các mặt cắt là từ 10% và đối với bộ dữ liệu DEM lấy mẫu 5m Lạng Sơn (bộ dữ liệu S1) thì mức tăng có nhiều thay đổi với giá trị tăng nhỏ nhất là 21% và giá trị tăng lớn nhất là 58 %. Điều này có thể giải thích là do hầu hết các mặt cắt có độ chính xác tăng hơn 50% (chẳng hạn như mặt cắt dọc 2, 4 và mặt cắt ngang 2, 4, 9 (Hình 3-9) nằm ở các khu vực địa hình đặc trưng như đáy thung lũng hoặc đỉnh đồi. Ngược lại, các mặt cắt có độ chính xác tăng ít hơn xảy ra chủ yếu ở các sườn núi, nơi bề mặt được biểu thị bởi DEM đầu vào tương đối giống với DEM chuẩn (độ phân giải cao). Mức tăng độ chính xác khá đồng đều giữa các mặt cắt đối với các DEM 20m và 30m giảm độ phân giải (D1 và D2) và DEM 30m của bộ dữ liệu S2 xảy ra do hầu hết các mặt cắt cắt qua nhiều dang địa hình khác nhau. Mức đô cải thiện về độ chính xác (so với mô hình song tuyến) cho bộ dữ liệu 20m là 16% và từ 24% đến 41%.

3.4.4.2 Đánh giá định lượng bằng cách sử dụng các hệ số hồi quy tuyến tính và hệ số tương quan R

Để đánh giá kết quả của các phương pháp khác nhau, các mô hình hồi quy tuyến tính được xác định từ dữ liệu DEM tham chiếu và DEM tăng độ phân giải bằng HNN và các thuật toán tái chia mẫu (Bảng 3-2). Các giá trị hệ số đã cho thấy DEM tăng độ phân giải không gian sử dụng mô hình HNN khóp với các DEM chuẩn hơn so với các DEM tái chia mẫu bằng các phương pháp thông thường. Đối với tất cả bốn bộ dữ liệu, giá trị của các tham số *m*, *b* và R^2 của các DEM tăng độ phân giải không gian sử dụng mô hình HNN gần hơn với các giá trị của 1, 0 và 1 hơn so với các giá trị tham số đó của các DEM tái chia mẫu song tuyến, Bi-cubic và Kriging.

Đối với bộ dữ liệu S1 Lạng Sơn, các DEM tạo thành sau khi tăng độ phân giải không gian sử dụng mô hình HNN có các giá trị |1-m| = 0.0195, |b| = 5.9080 và $R^2 = 0.9937$ cho thấy các DEM này giống với DEM tham chiếu hơn là các DEM sau khi tái chia mẫu song tuyến (|1-m| = 0.0399, |b| = 12.3782 và $R^2 = 0.9793$), DEM sau tái chia mẫu Bi-cubic (|1-m| = 0.0342, |b/ = 10.6432 và $R^2 = 0.9763$) và DEM sau nội suy Kriging (1-m| = 0.0550, |b| = 16.3717 và $R^2 = 0.9884$).

Thống kê hồi quy tuyến tính cho bộ dữ liệu lấy mẫu khu vực Đắc Hà (S2) cũng cho thấy, các DEM tạo thành sau khi tăng độ phân giải không gian sử dụng mô hình HNN khớp hơn với DEM chuẩn (|1-m| = 0.0043, |b| = 4.1179 và $R^2 = 0.9968$) so với các DEM tái chia mẫu song tuyến (|1-m| = 0.0128, |b| = 12.1453 và $R^2 = 0.9960$), các DEM tái chia mẫu Bi-cubic (|1-m| = 0.0115, |b| = 10.9118 và $R^2 = 0.9959$) và các DEM sau nội suy Kriging (|1-m| = 0.0078, |b| = 7.3917 và $R^2 = 0.9967$).

Các hệ số hồi quy tuyến tính cho DEM giảm độ phân giải 20m Nghệ An (bộ dữ liệu D1) cũng cho thấy sự rất khớp nhau giữa các DEM tạo thành sau khi tăng độ phân giải không gian sử dụng mô hình HNN (các hệ số (|1-m| = 0.0019, |b| = 0.2949 và $R^2 = 0.9973$) và DEM chuẩn.

Đối với bộ dữ liệu 30m (D2), sự tăng độ chính xác của DEM tăng độ phân giải bằng mô hình HNN đã được thấy rõ khi so sánh các tham số hồi quy tuyến tính của bốn phương pháp khác khau. Mặc dù hệ số xác định của DEM kết quả sau khi tăng độ phân giải bằng mô hình HNN ($R^2 = 0.9686$) chỉ lớn hơn không nhiều so với hệ số của phương pháp tái chia mẫu song tuyến ($R^2 = 0.9646$) và của phương pháp tái chia mẫu Bi-cubic ($R^2 = 0.9639$) và thậm chí còn nhỏ hơn của phương pháp Kriging ($R^2 = 0.9694$). Nhìn chung trên cơ sở so sánh các giá trị tương quan R^2 có thể thấy rằng thành phần sai số ngẫu nhiên nằm trong các giá trị độ cao của DEM có được từ phương pháp tăng độ phân giải sử dụng HNN đã được giảm đi rất nhiều so với kết quả của các phương pháp tái chia mẫu được sử dụng hiện nay cũng như phương pháp Kriging.

Đường khóp tuyệt đối của bộ dữ liệu cũng cho thấy có ít độ lệch (Bias) hơn giữa DEM sau khi tăng độ phân giải bằng mô hình HNN với DEM chuẩn (|1-m| = 0.0096, |b| = 1.6013) hơn là với các phương pháp tái chia mẫu song tuyến (|1-m| = 0.0500, |b| = 3.2057), phương pháp tái chia mẫu Bi-cubic (|1-m| = 0.0471, |b| = 2.8723) và phương pháp nội suy Kriging (|1-m| = 0.0392, |b| = 1.9291). Như đã phân tích ở trên, điều này cho thấy trong DEM được tăng độ phân giải bằng HNN chứa ít thành phần sai số hệ thống hơn so với các phương pháp truyền thống (song tuyến Bilinear, Bi-cubic và Kriging).

Như đã phân tích ở Chương 2, khi so sánh tham số độ dốc *m* và tham số chặn *b* của các đường khớp tuyệt đối của cả bốn bộ dữ liệu cho thấy rất rõ ràng là tất cả các tham số độ dốc *m* của các DEM tái chia mẫu đều nhỏ hơn 1 và tham số chặn *b* lớn hơn 0. Điều này có nghĩa là, tại các vị trí nơi địa hình thấp (thường là đáy của các thung lũng), các pixel của dữ liệu DEM được tạo bởi các phương pháp này có thể cao hơn các pixel tương ứng trong DEM chuẩn. Ngược lại, đối với các vị trí nơi địa hình cao (như đỉnh đồi hoặc rặng núi), độ cao của các pixel trong dữ liệu DEM tái chia mẫu có thể thấp hơn các pixel tương ứng trong DEM chuẩn. Khi sử dụng mô hình HNN để tăng độ phân giải hiện tượng này có thể được hạn chế. Điều này có thể được chứng minh khi trong các giá trị của bốn cặp tham số *m* và *b* cho phương pháp tăng độ phân giải DEM bằng mô hình HNN cho dữ liệu DEM trong bộ dữ liệu thực 5m (*m* = 1.0195, *b* = -5.9080) (S1) và 30m (*m* = 1.0043, *b* = -4.1179) (S2), tương ứng với việc các đường khớp này rất gần với đường 1: 1 với *m* = 1 và *b* = 0.

Thậm chí đối với bộ dữ liệu DEM 5m (bộ dữ liệu S1), phương pháp tăng độ phân giải bằng mô hình HNN có xu hướng tạo ra các giá trị độ cao ở các khu vực có

độ cao thấp sẽ thấp hơn một chút và ở các khu vực có độ lớn cao sẽ cao hơn một chút so với DEM chuẩn. Điều này có thể được giải thích bởi hiệu ứng của hàm điều kiện ràng buộc về độ cao của mô hình tăng độ phân giải HNN. Hiệu ứng này rất quan trọng vì nó cho thấy cách tiếp cận khi sử dụng mô hình HNN không chỉ là làm trơn dữ liệu như các thuật toán song tuyến, Bi-cubic và Kriging.

Nói cách khác, cấu trúc của phương pháp HNN (được xây dựng với mục tiêu làm trơn ở mức trên pixel là hàm mục tiêu và điều kiện ràng buộc mức pixel) đã mang lại một ưu điểm vượt trội mà các phương pháp nội suy và tái chia mẫu phổ biến khác không làm được.

3.5 So sánh độ chính xác về độ cao giữa các DEM sau khi tăng độ phân giải bằng thuật toán mạng neuron Hopfiled và các phương pháp tái chia mẫu với các điểm độ cao kiểm tra được đo bằng máy toàn đạc điện tử

Để so sánh độ chính xác về độ cao, NCS đã tiến hành so sánh các giá trị độ cao của các DEM sau khi tăng độ phân giải bằng thuật toán mạng neuron Hopfiled và các phương pháp tái chia mẫu với các điểm độ cao kiểm tra được đo trực tiếp trên thực địa bằng máy toàn đạc điện tử (TĐĐT) - đối với khu đo Mai Pha, Lạng Sơn. Vị trí các điểm đo (236 điểm đo) trên thực địa được minh họa trên Hình 3-16.

Sau khi tiến hành rải điểm đo thực địa bằng máy toàn đạc điện tử lên nền grid DEM của cùng khu đo đó trên ArcGis (Hình 3-16), dùng công cụ "Extract Multi Value to Point" có sẵn trong phần mềm ArcGIS để trích xuất giá trị độ cao trên DEM tại các vị trí có điểm đo để so sánh các giá trị chênh cao giữa độ cao đo bằng TĐĐT và độ cao xác định được trên DEM (tại cùng một vị trí).



Hình 3-16. Rải các điểm đo thực lên DEM 5m Lạng Sơn

Bảng so sánh chênh cao giữa DEM sau khi tăng độ phân giải bằng thuật toán mạng neuron Hopfiled (DEM HNN), DEM đầu vào (DEM Input) và các DEM sau khi tái chia mẫu song tuyến (DEM Bilinear), Bi-cubic (DEM Bi-cubic), Kriging (DEM Kriging) – đối với bộ dữ liệu DEM 5m Lạng Sơn với 236 điểm độ cao được đo bằng máy toàn đạc điện tử ở cùng khu vực được trình bày trong phần Phụ lục 2.

Kết quả so sánh các giá trị chênh lệch về độ cao (tính theo giá trị tuyệt đối của các chênh lệch độ cao) và thống kê các sai số của các DEM sau khi tăng độ phân giải và các độ cao đo thực địa được thể hiện trong Bảng 3-3.

Bảng 3-3. Bảng thống kê các chênh lệch độ cao, các sai số giữa các DEM tăng độ phân giải và các điểm đo bằng toàn đạc điện tử

	DEM	DEM	DEM	DEM	DEM
	HNN	Input	Bilinear	Bi-cubic	Kriging
Max (m)	3.637	7.240	4.878	5.076	3.609
Min (m)	0.001	0.000	0.000	0.000	0.000
Mean (m)	0.783	1.786	1.168	1.208	0.957
RMSE (m)	1.097	2.378	1.518	1.623	1.190

Kết quả trong Bảng 3-3 cho thấy: giá trị sai số trung phương RMSE của các DEM khi so sánh với các điểm đo lớn hơn RMSE khi so sánh với bề mặt DEM chuẩn, nguyên nhân là do các điểm đo ở vị trí bất kỳ trong một ô pixel của DEM, nhưng giá trị độ cao của tất cả các điểm trong một ô pixel của DEM thì chỉ có một giá trị, giá trị này được coi là giá trị của điểm ở trung tâm của pixel đó, vì thế nếu địa hình không bằng phẳng thì độ cao trên thực địa của các điểm nằm trong một ô pixel sẽ có các giá trị khác nhau, mặc dù trên DEM thì các điểm này sẽ có cùng một giá trị độ cao.

Tuy nhiên, có thể thấy rằng xu hướng giảm giá trị RMSE cũng tương đồng như trong trường hợp so sánh DEM tăng độ phân giải với DEM chuẩn. So sánh các giá trị sai số trung phương RMSE của các DEM sau khi tăng độ phân giải bằng phương pháp sử dụng mô hình mạng neuron Hopfiled (HNN) và các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic và Kriging, ta thấy giá trị sai số trung phương RMSE của phương pháp HNN là nhỏ nhất trong các phương pháp trên, có nghĩa là độ chính xác của phương pháp sử dụng mô hình mạng neuron Hopfiled là cao nhất.

3.6 Kết luận chương 3

Trong chương này đã đề xuất một phương pháp mới để tăng độ phân giải không gian và độ chính xác của các DEM dạng grid và thử nghiệm, đánh giá bằng các tham số khác nhau bằng cách sử dụng các dữ liệu DEM với độ phân giải và đặc tính khác nhau. Thuật toán tăng độ phân giải mới đề xuất được xây dựng dựa trên mạng neuron Hopfield (HNN) với hàm mục tiêu là tối đa hóa sự phụ thuộc không gian và điều kiện ràng buộc về độ cao. Các thực nghiệm được thực hiện trên hai loại dữ liệu độ cao: các bộ dữ liệu DEM giảm độ phân giải 20m (D1) và 30m (D2) ở tỉnh Nghệ An, Việt Nam; các bộ dữ liệu DEM dựa trên dữ liệu thực (sampled DEMs) 5m (S1) ở tỉnh Lạng Sơn (được xây dựng từ dữ liệu đo đạc mặt đất) và DEM 30m (S2) ở Đắc Hà, tỉnh Kontum, Việt Nam (được tạo ra từ các đường bình độ). Phương pháp đề xuất được đánh giá theo ba phương pháp tái chia mẫu phổ biến

thường được sử dụng hiện nay là phương pháp song tuyến (Bilinear), Bi-cubic và Kriging.

Các kết quả thử nghiệm cho thấy độ chính xác của phương pháp tăng độ phân giải không gian của DEM sử dụng mô hình HNN cao hơn hẳn so với các phương pháp tái chia mẫu song tuyến, Bi-cubic và Kriging.

Đánh giá trực quan cho thấy các DEM sau khi tăng độ phân giải sử dụng mô hình HNN khớp với các DEM tham chiếu hơn so với các DEM được tạo ra bằng các phương pháp tái chia mẫu song tuyến, Bi-cubic và Kriging.

Đánh giá độ chính xác định lượng dựa trên sai số trung phương cho thấy độ chính xác của các DEM sau khi tăng độ phân giải theo thuật toán sử dụng mô hình HNN cao hơn so với các phương pháp tái chia theo mô hình song tuyến, Bi-cubic và Kriging. Sai số trung phương của các DEM sau khi tăng độ phân giải đã giảm khoảng 39,9%, 5,2%, 43,9% và 10,0% tương ứng với các DEM giảm độ phân giải 20m và 30m ở tỉnh Nghệ An (bộ dữ liệu D1 và D2), DEM lấy mẫu 5m ở tỉnh Lạng Sơn (bộ dữ liệu S1) và DEM lấy mẫu 30m ở Đắc Hà, Việt Nam (bộ dữ liệu S2). Các giá trị sai số trung phương tổng thể của DEM sau khi tăng độ phân giải theo mô hình HNN nhỏ hơn so với các DEM được tạo ra theo phương pháp tái chia mẫu song tuyến, Bi-cubic và Kriging, đặc biệt là đối với các bộ dữ liệu DEM 5m và 20m.

Việc đánh giá cũng được thực hiện thêm bằng cách sử dụng hồi quy tuyến tính của độ phân giải không gian của DEM chuẩn so với DEM tăng độ phân giải theo mô hình HNN và DEM sau khi tái chia mẫu sử dụng các điểm chuẩn, đặc biệt tập trung vào các hệ số *m*, *b và* R^2 . Phân tích các tham số này cho thấy rằng, các DEM sau khi tăng độ phân giải theo mô hình HNN thì gần với các DEM tham chiếu hơn so với các DEM được tạo ra bằng các phương pháp tái chia mẫu thông thường.

Kết quả so sánh các giá trị chênh cao (tính theo giá trị tuyệt đối của các chênh cao) và thống kê các sai số của các DEM sau khi tăng độ phân giải (bằng phương pháp sử dụng mô hình mạng neuron Hopfiled (HNN) và các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic và Kriging) với 236 điểm độ cao đo thực địa bằng máy toàn đạc điện tử ở cùng khu vực Mai Pha - Lạng Sơn, cũng cho thấy sai số trung phương RMSE của phương pháp HNN là nhỏ nhất trong các phương pháp trên, có nghĩa là độ chính xác của phương pháp sử dụng mô hình mạng neuron Hopfiled là cao hơn các phương pháp tái chia mẫu thông thường.

Đánh giá trực quan và định lượng đã cho thấy thuật toán tăng độ phân giải không gian của DEM sử dụng mô hình HNN thực hiện chính xác hơn đối với một số đặc điểm địa hình đặc trưng như đáy thung lũng hoặc đỉnh của các rặng núi. Sai số trung phương của các mặt cắt nằm chủ yếu trong các đặc điểm địa hình này giảm khoảng 20% (tức là, được cải thiện nhiều hơn) so với các mặt cắt xảy ra chủ yếu ở sườn núi hoặc các khu vực bằng phẳng. Sự cải thiện này có thể được coi là do hiệu ứng của sự ràng buộc độ cao với các hàm tối đa hóa sự phụ thuộc không gian trong cách tiếp cận sử dụng mô hình HNN này. Đó là, cách thức đặc trưng của phương pháp sử dụng mô hình HNN mang lại lợi thế về cấu trúc khi tăng độ phân giải của DEM mà các phương pháp tái chia mẫu và nội suy không gian phổ biến hiện nay không làm được.

Mặc dù mô hình HNN được đề xuất nhằm tăng độ phân giải không gian cho grid DEM đã được chứng minh là chính xác hơn so với các phương pháp tái chia mẫu thông thường nhưng thuật toán vẫn còn có một số hạn chế. Cụ thể, do mô hình HNN sử dụng hàm kích hoạt là hàm tuyến tính, không cho phép nới lỏng điều kiện ràng buộc như trong mô hình HNN cho siêu phân giải bản đồ. Do đó, nếu DEM ban đầu chứa các sai số lớn thì sai số này có thể sẽ được chuyển hoàn toàn sang DEM tăng độ phân giải. Để giải quyết vấn đề này, cần nghiên cứu thêm và sửa đổi mô hình HNN, đặc biệt là các hàm kích hoạt.

KẾT LUẬN VÀ KIẾN NGHỊ

Kết luận

Kết quả nghiên cứu của luận án đã khẳng định và chứng minh được các luận điểm khoa học của đề tài, từ đó có thể đưa ra các kết luận như sau:

1. Các mô hình đánh giá độ chính xác cho DEM sử dụng kết hợp các tham số hồi quy (m, b) khi xây dựng mối quan hệ tương quan tuyến tính giữa dữ liệu chuẩn, dữ liệu đối sánh và hệ số tương quan R là phù hợp để đánh giá độ chính xác mô hình số độ cao DEM dạng grid. Việc sử dụng đồng thời các tham số trên cho phép đánh giá được sự tồn tại của các thành phần sai số ngẫu nhiên và sai số hệ thống trong các dữ liệu mô hình số độ cao.

2. Theo kết quả thử nghiệm được đánh giá độ chính xác theo cách tiếp cận mới, có sử dụng các tham số hồi quy tuyến tính (*m*, *b*) và hệ số tương quan *R* cho thấy, các thuật toán phổ biến bao gồm Bilinear, Bi-cubic, đặc biệt là phương pháp Kriging có thể cải thiện một cách nhất định về độ chính xác của mô hình số độ cao dạng grid khi tiến hành chia nhỏ các điểm ảnh. Tuy nhiên, qua phân tích cho thấy các thuật toán này gây ra một lượng sai số hệ thống trong kết quả DEM do hiệu ứng làm trơn. Các giá trị độ cao trên DEM có xu thế thấp đi tại các điểm đỉnh đồi, núi và phân thủy cũng như cao hơn ở các điểm đáy thung lũng hoặc tụ thủy.

3. Phương pháp mới để nâng cao độ chính xác cho dữ liệu DEM dạng grid bằng mạng neuron Hopfield mà luận án đề xuất có thể cải thiện độ chính xác cho DEM dạng grid so với các phương pháp tái chia mẫu khác. Mô hình nâng cao độ chính xác này là kết hợp của việc làm trơn DEM thông qua hàm mục tiêu được xác định bằng giá trị *semi-variance* min và hàm điều kiện ràng buộc về độ cao. Kết quả đánh giá độ chính xác cả bằng quan sát trực quan và các dữ liệu thống kê cho thấy phương pháp được đề xuất cho kết quả tốt hơn các phương pháp đang được sử dụng hiện nay sử dụng để chia nhỏ và làm trơn DEM. Đặc biệt là, hàm điều kiện trong mô hình đã cho phép giải quyết ảnh hưởng của hiệu ứng làm trơn.

Kiến nghị và hướng nghiên cứu tiếp theo

1. Dựa trên kết quả của thuật toán, NCS có mong muốn được tiếp tục hỗ trợ để có thể xây dựng các module phần mềm cho phép ứng dụng thuật toán HNN trên thực tế trong việc nâng cao độ chính xác dữ liệu mô hình số độ cao dạng grid và các dữ liệu độ cao có dạng tương tự.

2. Nghiên cứu tiếp tục hoàn thiện thuật toán, trong đó xây dựng hàm điều kiện ràng buộc độ cao mềm cho phép khử bớt các sai số tồn tại trong dữ liệu đầu vào mô hình số độ cao.

3. Nghiên cứu mở rộng thuật toán khi có thêm các nguồn thông tin khác có thể hỗ trợ việc hiệu chỉnh độ cao của mô hình mới, hoặc cho phép trộn các nguồn dữ liệu độ cao khác nhau để tạo ra mô hình số độ cao có độ chính xác cao hơn dữ liệu ban đầu, đủ độ chính xác cho các lĩnh vực có sử dụng mô hình số độ cao như nghiên cứu trượt lở, xây dựng mô hình dòng chảy thủy văn, v.v.

DANH MỤC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA TÁC GIẢ

Bài báo quốc tế ISI:

 Nguyen Quang Minh, Nguyen Thi Thu Huong, La Phu Hien, Hugh Lewis, P. Atkinson. *Downscaling gridded DEM using the Hopfield Neural Network*. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, ID JSTARS-2018-00862, 2018.

Bài báo Scopus:

- Nguyen Quang Minh, Nguyen Thi Thu Huong. Increasing Spatial Resolution of Remotely Sensed Image using HNN Super-resolution Mapping Combined with a Forward Model. Journal of the Korean of Surveying, Geodesy, Photogrammetry and Cartography. Vol.31, No. 6-2, 559-565, 2013, ISSN 1598-4850(Print), ISSN 2288-260X (Online), 2013.
- Nguyen Quang Minh, Nguyen Thi Thu Huong, La Phu Hien, Duong Thi Tuyet Nhung. Comparison of the conventional resampling methods for gridded DEM downscaling. Journal of the Polish Mineral Engineering Society Pol-Viet 2019. Doi: <u>http://doi.org/10.29227/IM-2019-01-77</u>.

Bài báo Hội nghị quốc tế:

- 4. Nguyen Quang Minh, Nguyen Thi Thu Huong, La Phu Hien, Pham Thanh Thao, P.M. Atkinson. *Increasing the Grid DEM Resolusion Using Hopfield Neural Network, A Test For Data in LangSon Province, VietNam*. International symposium on Geo-Spatial and Mobile mapping technologies and summer school for Mobile mapping technology (GMMT2016). ISBN 978-604-76-914-4, 2016.
- Nguyen Quang Minh, Nguyen Thi Thu Huong, La Phu Hien. A Quantitative Assessment of Algorithm for Increasing Gridded DEM resolution using the Hopfield Neural Network. International Conference on Earth Sciences And Sustainable Geo-Resources Development (ESASGD 2016). ISBN 978-604-76-1171-3, 2016.

Bài báo Tạp chí trong nước:

- 6. Nguyễn Thị Thu Hương, Nguyễn Quang Minh. Comparison of the resampling methods for gridded dem downscaling. Tạp chí Khoa học Kỹ thuật Mỏ - Địa Chất. Số tiếng Anh, tập 59, kỳ 6, ISSN 1859-1469, 2018.
- 7. Nguyễn Thị Thu Hương. Một phương pháp nhằm tăng độ phân giải không gian của mô hình số độ cao dạng grid sử dụng mạng nơ-ron Hopfiled. Tạp chí Khoa học Kỹ thuật Mỏ Địa Chất. Tập 60, kỳ 2, ISSN 1859-1469, 2019.

Bài báo Hội nghị trong nước:

 Nguyễn Thị Thu Hương. Một cách tiếp cận mới nhằm tăng độ phân giải không gian của mô hình số độ cao dạng grid bằng phương pháp sử dụng mạng nơ-ron Hopfield. Kỳ yếu hội thảo khoa học Trái đất – Mỏ - Môi trường bền vững (EME 2018). P. 238-246, ISBN: 978-604-913-687-0, 2018.

Đề tài nghiên cứu khoa học:

9. Nguyễn Thị Thu Hương, Lã Phú Hiến, Lê Ngọc Giang. Nghiên cứu nâng cao độ chính xác của mô hình số độ cao DEM dạng grid bằng mạng nơ-ron Hopfiled. Đề tài NCKH cấp Trường, mã số T18-11, 2018.

TÀI LIỆU THAM KHẢO

Tiếng Việt

- [83] Bùi Công Cường, "Kiến thức cơ sở của hệ mờ", in Hệ mờ, mạng nơ ron và ứng dụng, B. C. C. N. D. Phước, Ed. Hà Nội: Nhà xuất bản khoa học kỹ thuật, 2001, pp. 9-35.
- [84] N. Q. Minh, "Nghiên cứu ứng dụng mạng nơ ron Hopfiled nhằm tăng cường độ phân giải bản đồ lớp phủ", Báo cáo đề tài Nghiên cứu Khoa học Công nghệ Cấp Bộ, Hà Nội, 2013.
- [85] T. Đ. T. Lê Tuấn Anh, Phạm Văn Tuân, Lê Đình Hiển, "Ứng dụng công nghệ tích hợp Lidar và chụp ảnh hàng không (Citymapper - Leica) trong thu nhận, xử lý và thành lập dữ liệu không gian địa lý", Báo cáo tại Hội nghị khoa học, công nghệ toàn quốc ngành Đo đạc và Bản đồ, Hà Nội, 2018.
- [86] T. A. Kiệt, "Phương pháp đo ảnh giải tích và đo ảnh số", Đại học Mỏ-Địa chất, Hà Nội, 2000.
- [87] T. Q. Cương, "Nghiên Cứu Cơ Sở Khoa Học Xây Dựng Mô Hình Số Độ Cao Phục Vụ Quản Lý Tài Nguyên Thiên Nhiên", Báo cáo tổng kết đề tài khoa học và kỹ thuật Bộ Tài nguyên và môi trường, Hà Nội, 2006.
- [89] T. B. Dieu et al., "Xây dựng mô hình số bề mặt và bản đồ trực ảnh sử dụng công nghệ đo ảnh máy bay không người lái (UAV)", Báo cáo tại Hội nghị khoa học Đo đạc-Bản đồ với ứng phó biến đổi khí hậu, Hà Nội, 7/2016.

Tiếng Anh

- [1] G. Droj, "Improving the accuracy of digital terrain models," *Studia Universitatis Babes-Bolyai: Series Informatica, LIII*, vol. 1, pp. 65-72, 2008.
- [2] L. G. Miles, "Global Digital Elevation Model Accuracy Assessment in the Himalaya, Nepal," 2013.
- [3] D. Tank and J. J. Hopfield, "Simple'neural'optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit," *IEEE transactions on circuits and systems*, vol. 33, no. 5, pp. 533-541, 1986.
- [4] J. P. Walker and G. R. Willgoose, "On the effect of digital elevation model accuracy on hydrology and geomorphology," *Water Resources Research*, vol. 35, no. 7, pp. 2259-2268, 1999.
- [5] J. Höhle and M. Höhle, "Accuracy assessment of digital elevation models by means of robust statistical methods," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 64, no. 4, pp. 398-406, 2009.
- [6] M. Mokarrama and M. Hojati, "Landform classification using a sub-pixel spatial attraction model to increase spatial resolution of digital elevation model (DEM)," *The Egyptian Journal of Remote Sensing and Space Science*, vol. 21, no. 1, pp. 111-120, 2018.
- [7] J. Titus and S. Geroge, "A comparison study on different interpolation methods based on satellite images," *International Journal of Engineering Research & Technology*, vol. 2, no. 6, pp. 82-85, 2013.

- [8] F. J. Aguilar, M. A. Aguilar, and F. Agüera, "Accuracy assessment of digital elevation models using a non-parametric approach," *International Journal of Geographical Information Science*, vol. 21, no. 6, pp. 667-686, 2007.
- [9] M. N. Ikechukwu, E. Ebinne, U. Idorenyin, and N. I. Raphael, "Accuracy Assessment and Comparative Analysis of IDW, Spline and Kriging in Spatial Interpolation of Landform (Topography): An Experimental Study," *Journal of Geographic Information System*, vol. 09, no. 03, pp. 354-371, 2017.
- [10] Y. Tang, J. Zhang, H. Li, H. Ding, and L. Jing, "Terrain data conflation using an improved pattern-based multiple-point geostatistical approach," vol. 9263, p. 926304: International Society for Optics and Photonics.
- [11] C. A. Felgueiras, J. O. Ortiz, E. C. G. Camargo, L. M. Namikawa, and S. Rosim, "Fusion Based on Geostatistics to Improve the Altimetry Accuracies of Digital Elevation Models. Geoinfor Geostat: An Overview 6: 2," *of*, vol. 7, p. 2, 2018.
- [12] V. Chaplot, F. Darboux, H. Bourennane, S. Leguédois, N. Silvera, and K. Phachomphon, "Accuracy of interpolation techniques for the derivation of digital elevation models in relation to landform types and data density," *Geomorphology*, vol. 77, no. 1-2, pp. 126-141, 2006.
- [13] C. R. I. P. (CRIP). (2014, 26/7). *Introduction for Digital Elevation Models*. Available: <u>http://charim.net/datamanagement/32</u>
- [14] P. Holmes, "Correlation: From picture to formula," (in English), *Teaching Statistics*, vol. 23, no. 3, pp. 67-71, 2001.
- [15] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," (in English), *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, p. 3088, 1984.
- [16] J.-W. Lin, "Artificial Neural Network Related to Biological Neuron Network: A Review," (in English), Advanced Studies in Medical Sciences, vol. 5, no. 1, pp. 55-62, 2017.
- [17] J. A. Anderson, "Neural models with cognitive implications," (in English), *Basic processes in reading: Perception and comprehension*, pp. 27-90, 1977.
- [18] B. E. Vieux, "DEM aggregation and smoothing effects on surface runoff modeling," (in English), *Journal of Computing in Civil Engineering*, vol. 7, no. 3, pp. 310-338, 1993.
- [19] T. Kohonen, *Associative memory: A system-theoretical approach*. Springer Science & Business Media, 2012.
- [20] V. Florinsky Igor, "Digital Terrain Analysis in soil science and geology," ed: Amsterdam: ACADEMIC PRESS/Elsevier, 2012.
- [21] K. Suzuki, Artificial neural networks: methodological advances and biomedical applications. BoD–Books on Demand, 2011.
- [22] F. M. Ziadat, "Effect of Contour Intervals and Grid Cell Size on the Accuracy of DEMs and Slope Derivatives," (in English), *Transactions in GIS*, vol. 11, no. 1, pp. 67-81, 2007/02/01 2007.

- [23] ASPRS, "ASPRS positional accuracy standards for digital geospatial data," (in English), *Photogrammetric Engineering and Remote Sensing*, vol. 81, no. 3, pp. 1-26, 2015.
- [24] N. Q. Minh, "Image smoothing of multispectral imagery based on the HNN and geo-statistics," (in English), *Yaogan Xuebao- Journal of Remote Sensing*, vol. 15, no. 3, pp. 640-644, 2011.
- [25] G. van der Veer, "4 Development and application of geospatial models for verifying the geographical origin of food," in *New Analytical Approaches for Verifying the Origin of Food*, P. Brereton, Ed.: Woodhead Publishing, 2013, pp. 60-80.
- [26] J. Seo, "Study on Geographic information production using airborne laser scanning," presented at the FIG Working Week 2001, Seoul, Korea, 2001.
- [27] F. Ackermann, "Techniques and strategies for DEM generation," (in English), *Digital photogrammetry: An addendum to the manual of photogrammetry*, pp. 135-141, 1996.
- [28] USGS, "Digital elevation models, data users guide 5," (in English), *Reston, Virginia, US Geological Survey,* pp. 1-50, 1993.
- [29] A. O. Altunel, "Evaluation of TanDEM-X 90 m Digital Elevation Model," (in English), *International Journal of Remote Sensing*, vol. 40, no. 7, pp. 2841-2854, 2019/04/03 2019.
- [30] P. M. Atkinson, "Mapping sub-pixel boundaries from remotely sensed images," (in English), *Innovations in GIS*, vol. 4, pp. 166-180, 1997.
- [31] FGDC, "Geospatial Positioning Accuracy Standards Part 3: National Standard for Spatial Data Accuracy," (in English), *National Aeronautics and Space Administration: Virginia, NV, USA,* 1998.
- [32] Q. Weng, "An evaluation of spatial interpolation accuracy of elevation data," in *Progress in spatial data handling*: Springer, 2006, pp. 805-824.
- [33] S. Fadnavis, "Image interpolation techniques in digital image processing: an overview," (in English), *International Journal of Engineering Research and Applications*, vol. 4, no. 10, pp. 70-73, 2014.
- [34] Humboldt State University. (2017, 26/7). *Geospatial Lessions*. Available: <u>http://gsp.humboldt.edu/</u>
- [35] M. Crosetto and F. Pérez Aragues, "Radargrammetry and SAR interferometry for DEM generation: validation and data fusion," 2000, vol. 450, p. 367.
- [36] G. J. P. Schumann and P. D. Bates, "The need for a high-accuracy, openaccess global DEM," (in English), *Frontiers in Earth Science*, vol. 6, p. 225, 2018.
- [37] Y. Gorokhovich and A. Voustianiouk, "Accuracy assessment of the processed SRTM-based elevation data by CGIAR using field data from USA and Thailand and its relation to the terrain characteristics," (in English), *Remote sensing of Environment*, vol. 104, no. 4, pp. 409-415, 2006.

- [38] F. Dell'Acqua and P. Gamba, "Preparing an urban test site for SRTM data validation," (in English), *IEEE transactions on geoscience and remote sensing*, vol. 40, no. 10, pp. 2248-2256, 2002.
- [39] L. Bian and R. Butler, "Comparing effects of aggregation methods on statistical and spatial properties of simulated spatial data," (in English), *Photogrammetric engineering and remote sensing*, vol. 65, pp. 73-84, 1999.
- [40] K.-t. Chang and B.-w. Tsai, "The effect of DEM resolution on slope and aspect mapping," (in English), *Cartography and geographic information systems*, vol. 18, no. 1, pp. 69-77, 1991.
- [41] P. V. Bolstad and T. Stowe, "An evaluation of DEM accuracy: elevation, slope, and aspect," (in English), *Photogrammetric Engineering & Remote Sensing*, vol. 60, no. 11, pp. 1327-1332, 1994.
- [42] P. P. Shingare and S. S. Kale, "Review on digital elevation model," (in English), *International Journal of Modern Engineering Research (IJMER)*, vol. 3, no. 4, pp. 2412-2418, 2013.
- [43] C. H. Genitha and K. Vani, "Super resolution mapping of satellite images using hopfield neural networks," 2010, pp. 114-118: IEEE.
- [44] L. Renouard and F. Lehmann, "Digital Aerial Survey Data for Telecoms Network Planning: Practical Experience with a High-resolution Three-view Stereo Camera Discussion," (in English), OFFICIAL PUBLICATION-EUROPEAN ORGANISATION FOR EXPERIMENTAL PHOTOGRAMMETRIC RESEARCH, pp. 97-108, 1999.
- [45] B. Petzold and L. Nordrhein-Westfalen, "DTM Determination by Laser-Scanning–An Efficient Alternative," (in English), *OEEPE*, p. 215, 1999.
- [46] J. W. McKean and S. J. Sheather, "Statistics, Nonparametric," in Encyclopedia of Physical Science and Technology (Third Edition), R. A. Meyers, Ed. New York: Academic Press, 2003, pp. 891-914.
- [47] R. F. Vázquez and J. Feyen, "Assessment of the effects of DEM gridding on the predictions of basin runoff using MIKE SHE and a modelling resolution of 600 m," (in English), *Journal of Hydrology*, vol. 334, no. 1-2, pp. 73-87, 2007.
- [48] L. De Floriani and P. Magillo, "Digital Elevation Models," in *Encyclopedia of Database Systems*, L. Liu and M. T. Özsu, Eds. New York, NY: Springer New York, 2018, pp. 1078-1083.
- [49] C. H. Grohmann and S. S. Steiner, "SRTM resample with short distance-low nugget kriging," (in English), *International Journal of Geographical Information Science*, vol. 22, no. 8, pp. 895-906, 2008.
- [50] Q. M. Nguyen, P. M. Atkinson, and H. G. Lewis, "Super-resolution mapping using Hopfield neural network with panchromatic imagery," (in English), *International journal of remote sensing*, vol. 32, no. 21, pp. 6149-6176, 2011.
- [51] J. M. Schoorl, M. P. W. Sonneveld, and A. Veldkamp, "Three-dimensional landscape process modelling: the effect of DEM resolution," (in English),

Earth Surface Processes and Landforms: The Journal of the British Geomorphological Research Group, vol. 25, no. 9, pp. 1025-1034, 2000.

- [52] Q. M. Nguyen, P. M. A. Van Duong Do, and H. G. Lewis, "Downscaling multispectral imagery based on the HNN using forward model," (in English), 2009.
- [53] M. W. Thornton, P. M. Atkinson, and D. A. Holland, "Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping," (in English), *International Journal* of Remote Sensing, vol. 27, no. 3, pp. 473-491, 2006.
- [54] T. Kasetkasem, M. K. Arora, and P. K. Varshney, "Super-resolution land cover mapping using a Markov random field based approach," (in English), *Remote Sensing of Environment*, vol. 96, no. 3-4, pp. 302-314, 2005.
- [55] D. J. Amit, H. Gutfreund, and H. Sompolinsky, "Spin-glass models of neural networks," (in English), *Physical Review A*, vol. 32, no. 2, p. 1007, 1985.
- [56] W. Shi, B. Wang, and Y. Tian, "Accuracy analysis of digital elevation model relating to spatial resolution and terrain slope by bilinear interpolation," (in English), *Mathematical Geosciences*, vol. 46, no. 4, pp. 445-481, 2014.
- [57] J. S. Deems, T. H. Painter, and D. C. Finnegan, "Lidar measurement of snow depth: a review," (in English), *Journal of Glaciology*, vol. 59, no. 215, pp. 467-479, 2013.
- [58] D. F. Maune, J. B. Maitra, and E. J. McKay, "Accuracy standards," (in English), *Digital Elevation Models and Applications: The DEM Users Manual*, pp. 61-82, 2001.
- [59] P. C. Kyriakidis, A. M. Shortridge, and M. F. Goodchild, "Geostatistics for conflation and accuracy assessment of digital elevation models," (in English), *International Journal of Geographical Information Science*, vol. 13, no. 7, pp. 677-707, 1999.
- [60] A. Hirano, R. Welch, and H. Lang, "Mapping from ASTER stereo image data: DEM validation and accuracy assessment," (in English), *ISPRS Journal* of Photogrammetry and Remote Sensing, vol. 57, no. 5, pp. 356-370, 2003/04/01/ 2003.
- [61] D. Kidner, M. Dorey, and D. Smith, "What's the point? Interpolation and extrapolation with a regular grid DEM," 1999.
- [62] H. Niroumand, M. F. M. Zain, and M. Jamil, "Statistical Methods for Comparison of Data Sets of Construction Methods and Building Evaluation," (in English), *Procedia - Social and Behavioral Sciences*, vol. 89, pp. 218-221, 2013.
- [63] L. Zhilin, Z. Qing, and C. Gold, *Digital terrain modeling Principles and methodology*. CRC-Press, 2005.
- [64] K. Mehrotra, C. K. Mohan, and S. Ranka, *Elements of artificial neural networks*. MIT press, 1997.
- [65] X. Li, Y. Du, and F. Ling, "Spatially adaptive smoothing parameter selection for Markov random field based sub-pixel mapping of remotely sensed

images," (in English), *International journal of remote sensing*, vol. 33, no. 24, pp. 7886-7901, 2012.

- [66] Y. Kim, N. Kang, S. Kim, and H. Kim, "Evaluation for Snowfall Depth Forecasting using Neural Network and Multiple Regression Models," (in English), *Journal of Korean Society of Hazard Mitigation*, vol. 13, 04/30 2013.
- [67] I. Puente, H. González-Jorge, J. Martínez-Sánchez, and P. Arias, "Review of mobile mapping and surveying technologies," (in English), *Measurement*, vol. 46, no. 7, pp. 2127-2145, 2013.
- [68] M. P. Smith, A. X. Zhu, J. E. Burt, and C. Stiles, "The effects of DEM resolution and neighborhood size on digital soil survey," (in English), *Geoderma*, vol. 137, no. 1-2, pp. 58-69, 2006.
- [69] P. A. Burrough, R. McDonnell, R. A. McDonnell, and C. D. Lloyd, *Principles of geographical information systems*. Oxford university press, 2015.
- [70] D. F. Maune, S. M. Kopp, C. A. Crawford, and C. E. Zervas, "Introduction to digital elevation models (DEM)," (in English), *Digital photogrammetry: an addendum to the manual of photogrammetry*, 1996.
- [71] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan, *Neural network design*. Martin Hagan, 2014.
- [72] A. Castrignanò, G. Buttafuoco, R. Comolli, and C. Ballabio, "Accuracy assessment of digital elevation model using stochastic simulation," 2006, pp. 490-498: Citeseer.
- [73] A. J. Tatem, H. G. Lewis, P. M. Atkinson, and M. S. Nixon, "Multiple-class land-cover mapping at the sub-pixel scale using a Hopfield neural network," (in English), *International Journal of Applied Earth Observation and Geoinformation*, vol. 3, no. 2, pp. 184-190, 2001.
- [74] F. J. Aguilar, F. Agüera, M. A. Aguilar, and F. Carvajal, "Effects of terrain morphology, sampling density, and interpolation methods on grid DEM accuracy," (in English), *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 7, pp. 805-816, 2005.
- [75] I. Chaubey, A. S. Cotter, T. A. Costello, and T. S. Soerens, "Effect of DEM data resolution on SWAT output uncertainty," (in English), *Hydrological Processes: An International Journal*, vol. 19, no. 3, pp. 621-628, 2005.
- [76] N. A. Muhadi, A. F. Abdullah, S. K. Bejo, M. R. Mahadi, and A. Mijic, "The Use of LiDAR-Derived DEM in Flood Applications: A Review," (in English), *Remote Sensing*, vol. 12, no. 14, p. 2308, 2020.
- [77] F. Ling, Y. Du, F. Xiao, H. Xue, and S. Wu, "Super-resolution land-cover mapping using multiple sub-pixel shifted remotely sensed images," (in English), *International Journal of Remote Sensing*, vol. 31, no. 19, pp. 5023-5040, 2010.
- [78] C. U. Paredes-Hernández, N. J. Tate, K. J. Tansey, P. F. Fisher, and W. E. Salinas-Castillo, "Increasing the accuracy of low spatial resolution digital elevation models using geostatistical conflation," 2010, pp. 413-416.

- [79] T. Tadono, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto, "Precise global DEM generation by ALOS PRISM," (in English), *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, no. 4, p. 71, 2014.
- [80] B. Atkins, C. A. Bouman, J. P. Allebach, J. S. Gondek, M. T. Schramm, and F. W. Sliz, "Computerized method for improving data resolution," ed: Google Patents, 2000.
- [81] Z. Zhao, G. Benoy, T. L. Chow, H. W. Rees, J.-L. Daigle, and F.-R. Meng, "Impacts of accuracy and resolution of conventional and LiDAR based DEMs on parameters used in hydrologic modeling," (in English), *Water resources management*, vol. 24, no. 7, pp. 1363-1380, 2010.
- [82] K. F. Noltimier *et al.*, "RADARSAT Antarctic mapping project-mosaic construction," 1999, vol. 5, pp. 2349-2351: IEEE.
- [88] S. Mukherjee, P. K. Joshi, S. Mukherjee, A. Ghosh, R. D. Garg, and A. Mukhopadhyay, "Evaluation of vertical accuracy of open source Digital Elevation Model (DEM)," (in Vietnamese), *International Journal of Applied Earth Observation and Geoinformation*, vol. 21, pp. 205-217, 2013/04/01/2013.

PHỤ LỤC 1

MỘT SỐ ĐOẠN CODE CỦA CHƯƠNG TRÌNH NÂNG CAO ĐỘ CHÍNH XÁC CỦA GRID DEM SỬ DỤNG MÔ HÌNH MẠNG NEURON HOPFIELD

import gdal, osr

import numpy as np

Định nghĩa các hàm # Hàm mở file và gán vào một đối tượng thuộc lớp gdal def openRaster(fn, access): ds = gdal.Open(fn, access) if ds is None: print("Error opening raster dataset") return ds

#load một band từ file đối tượng thuộc lớp gdal là ds.

#ds được tạo ra bằng cách gọi hàm gdal.Open()

#kết quả trả lại của hàm là một array trong thư viện numpy (band) đặc trưng cho một lớp độ cao

```
def getRasterBand(fn, band = 1, access = 0):
    ds = openRaster(fn, access)
    band = ds.GetRasterBand(1).ReadAsArray()
    return band
```

#load một band từ file đối tượng thuộc lớp gdal là ds. #ds được tạo ra bằng cách gọi hàm gdal.Open() #Lấy giá trị tọa độ của ảnh def getGeoT(fn, access = 0): ds = openRaster(fn, access) geot = ds.GetGeoTransform() return geot

```
#load một band từ file đối tượng thuộc lớp gdal là ds.
#ds được tạo ra bằng cách gọi hàm gdal.Open()
#Lấy giá trị phép chiếu sử dụng cho ảnh này
def getSRC(fn, access = 0):
    ds = openRaster(fn, access)
    proj = ds.GetProjection()
    return proj
```

#Khởi tạo dữ liệu downscaling,

#giá trị độ cao ban đầu của các sub-pixel đúng bằng giá trị độ cao của pixel gốc def initialize(data, zoom):

```
band = np.repeat(data, zoom, axis = 0)
```

```
band = np.repeat(band, zoom, axis = 1)
```

```
# for i in range(0, band.shape[0]):
```

```
# for j in range(0, band.shape[1]):
```

```
# print(band[i][j])
```

```
return band
```

#Tính giá trị của hàm spatial dependence maximisation def sd(dtin):

```
width = dtin.shape[0]
height = dtin.shape[1]
usd = np.zeros((width, height))
```

```
# print(width, height)
```

```
for i in range(0, width):
for j in range(0, height):
#set window
```

count = 0 sum = 0.0if i == 0: swr = 0else: swr = i - 1

if i == width - 1: ewr = i + 1

else:

ewr = i + 2

if j == 0: swc = 0else: swc = j - 1if j == height - 1: ewc = j + 1else: ewc = j + 2for 1 in range(swr, ewr): for m in range(swc, ewc): count += 1 sum += dtin[1][m] $v_current = dtin[i][j]$ $vexp = (sum - v_current)/(count - 1)$ $usd[i][j] = vexp - v_current$

return usd

#Điều kiện ràng buộc về độ cao, Elevation Function def elc(dtin, goc):

Xác định kích thước của DEM gốc và DEM được downscaled

```
width = dtin.shape[0]
height = dtin.shape[1]
goc_w = goc.shape[0]
goc_h = goc.shape[1]
zoom = int(width/goc_w)
uec = np.zeros((width, height))
```

print(width, height, zoom)

```
for i in range(0, goc_w):
```

```
for j in range(0, goc_h):
```

#determine the average elevation

#Determine the range of a pixel in the original DEM

```
sum = 0.0
```

```
swr = i*zoom
```

```
ewr = swr + zoom
```

```
swc = j*zoom
```

ewc = swc + zoom

Tính tổng của các điểm độ cao của các sub-pixel trên một pixel

```
for l in range(swr, ewr):
```

for m in range(swc, ewc):

sum += dtin[1][m]

Độ cao gốc Elevation để tính du_elevation của một pixel gốc

elev = goc[i][j]

Độ cao trung bình

```
velc = sum /(zoom*zoom)
```

Tính giá trị hiệu chỉnh của hàm độ cao cho từng sub-pixel

```
for l in range(swr, ewr):
    for m in range(swc, ewc):
        uec[l][m] = elev - velc
```

return uec

#Ghi kết quả ra file

```
def createRaster(fn, data, geot, proj, driverFmt = "GTiff"):
    driver = gdal.GetDriverByName(driverFmt)
    outds = driver.Create(fn, xsize = data.shape[1], ysize = data.shape[0], bands = 1,
    eType = gdal.GDT_Float32)
    outds.SetGeoTransform(geot)
    outds.SetProjection(proj)
    outds.GetRasterBand(1).WriteArray(data)
    outds = None
```

tạo đối tượng str là fn1 chỉ đến file độ cao

fn1 = "E:/AAATHESIS/AAA_Luận án/QGIS 3.16/ned10m_cut.tif"
fnout1 = "E:/ AAATHESIS/AAA_Luận án/QGIS 3.16 /test1.tif"
fnout2 = "E:/ AAATHESIS/AAA_Luận án/QGIS 3.16 /test2.tif"
goc = getRasterBand(fn1)
#Lấy các tham số vị trí của DEM gốc
geotgoc = getGeoT(fn1)
print(geotgoc)
#lấy thông tin phép chiếu của DEM gốc
projgoc = getSRC(fn1)

z = 4 #zoom factor = 4
#Xác lập thông tin vị trí cho DEM được chia nhỏ bằng HNN

```
geotnew = [geotgoc[0], geotgoc[1]/z, geotgoc[2], geotgoc[3], geotgoc[4],
geotgoc[5]/z]
dscal = initialize(goc, z) #khởi tạo bằng giá trị DEM gốc, OK
Energy_old = 10000000000.0
Threshold = 0.001
Energy_dif = 10000000.0
iteration = 0
while abs(Energy_dif) > Threshold:
  usd = sd(dscal)
  uec = elc(dscal, goc)
  u = usd + uec
  Energy_new = abs(usd).sum()+abs(uec).sum()
  dscal = dscal + u
  Energy_dif = Energy_old - Energy_new
  Energy_old = Energy_new
  iteration +=1
```

```
print("Vòng lặp số: ", iteration, Energy_new)
```

createRaster(fnout1, dscal, geotnew, projgoc)

iface.addRasterLayer(fnout1)
#iface.addRasterLayer(fn1)
#iface.addRasterLayer(fn1)

PHỤ LỤC 2

SO SÁNH CÁC GIÁ TRỊ CHÊNH CAO GIỮA CÁC DEM SAU KHI TĂNG ĐỘ PHÂN GIẢI BẰNG PHƯƠNG PHÁP SỬ DỤNG MẠNG NEURON HOPFILED, CÁC PHƯƠNG PHÁP TÁI CHIA MÃU VỚI CÁC ĐIỂM ĐỘ CAO ĐƯỢC ĐO BẰNG MÁY TOÀN ĐẠC ĐIỆN TỬ Ở CÙNG KHU VỰC

Bảng PL2-1. Bảng so sánh chênh cao giữa DEM sau khi tăng độ phân giải bằng thuật toán mạng neuron Hopfiled (DEM HNN), DEM đầu vào (DEM Input) và các DEM sau khi tái chia mẫu song tuyến (DEM Bilinear), Bi-cubic (DEM Bi-cubic), Kriging (DEM Kriging) – đối với bộ dữ liệu DEM 5m Lạng Sơn với 236 điểm độ cao được đo bằng máy toàn đạc điện tử ở cùng khu vực

	Độ cao	DEM		DEM		DEM		DEM		DEM	
STT	đo	5m	Chênh	20m	Chênh	5m	Chênh	5m	Chênh	5m	Chênh
		HNN	cao	Input	cao	Bilinear	cao	Bicubic	cao	Kriging	cao
1	321.5	320.19	1.31	317.13	4.37	319.12	2.38	318.39	3.11	319.47	2.03
2	319.9	318.44	1.46	317.13	2.77	317.32	2.58	317.25	2.65	317.97	1.93
3	319.8	317.82	1.98	317.13	2.67	317.12	2.68	317.13	2.67	317.13	2.67
4	313.5	313.83	-0.33	317.13	-3.63	313.19	0.31	313.32	0.18	315.33	-1.83
5	330.1	329.06	1.04	329.90	0.20	328.95	1.15	329.27	0.83	328.86	1.24
6	329.6	329.19	0.41	328.68	0.92	328.51	1.09	328.57	1.03	329.14	0.46
7	328.7	328.67	0.03	327.98	0.72	328.15	0.55	328.09	0.61	328.48	0.22
8	328.4	328.84	-0.44	328.68	-0.28	327.58	0.82	327.97	0.43	328.10	0.30
9	328.4	328.84	-0.44	328.68	-0.28	327.58	0.82	327.97	0.43	328.10	0.30
10	326.9	327.13	-0.23	325.31	1.59	325.98	0.92	325.73	1.17	326.94	-0.04
11	326.6	326.81	-0.21	327.98	-1.38	325.62	0.98	326.01	0.59	326.69	-0.09
12	326.9	327.14	-0.24	324.60	2.30	325.62	1.28	325.24	1.66	327.22	-0.32
13	326.5	327.49	-0.99	324.60	1.90	325.71	0.79	325.32	1.18	326.37	0.13
14	325	324.84	0.16	325.31	-0.31	323.27	1.73	324.04	0.96	323.57	1.43
15	317.4	317.78	-0.38	317.05	0.35	317.60	-0.20	318.07	-0.67	317.41	-0.01
16	315.5	316.31	-0.81	322.74	-7.24	319.18	-3.68	320.49	-4.99	316.94	-1.44
17	312.9	315.02	-2.12	314.00	-1.10	317.30	-4.40	317.78	-4.88	314.19	-1.29
18	314	311.74	2.26	314.00	0.00	314.00	0.00	314.00	0.00	314.00	0.00

19	313.4	310.83	2.57	314.00	-0.60	313.01	0.39	313.38	0.02	311.48	1.92
20	325.9	326.50	-0.60	326.68	-0.78	326.77	-0.87	326.70	-0.80	325.34	0.56
21	320.2	319.34	0.86	323.13	-2.93	320.85	-0.65	321.71	-1.51	319.67	0.53
22	326	326.82	-0.82	329.05	-3.05	327.57	-1.57	328.13	-2.13	326.02	-0.02
23	324	324.87	-0.87	329.05	-5.05	325.82	-1.82	327.07	-3.07	322.45	1.55
24	319.9	317.63	2.27	314.00	5.90	320.40	-0.50	319.71	0.19	318.98	0.92
25	322.5	320.05	2.45	322.74	-0.24	322.31	0.19	322.43	0.07	320.69	1.81
26	322.5	322.52	-0.02	322.74	-0.24	324.32	-1.82	323.73	-1.23	320.69	1.81
27	317.5	315.41	2.09	317.74	-0.24	316.78	0.72	317.07	0.43	317.67	-0.17
28	317.5	318.94	-1.44	317.05	0.45	317.95	-0.45	317.61	-0.11	317.05	0.45
29	317.4	320.28	-2.88	317.05	0.35	319.15	-1.75	318.33	-0.93	319.34	-1.94
30	317.1	318.51	-1.41	321.28	-4.18	317.85	-0.75	318.35	-1.25	317.90	-0.80
31	317.2	319.83	-2.63	321.28	-4.08	319.16	-1.96	319.16	-1.96	319.35	-2.15
32	323.8	324.96	-1.16	321.44	2.36	324.35	-0.55	324.24	-0.44	323.43	0.37
33	321	320.92	0.08	317.87	3.13	319.09	1.91	318.58	2.42	319.54	1.46
34	310.6	312.31	-1.71	311.49	-0.89	310.70	-0.10	310.70	-0.10	311.37	-0.77
35	321.3	322.81	-1.51	321.44	-0.14	322.75	-1.45	322.26	-0.96	320.32	0.98
36	312.5	312.55	-0.05	310.03	2.47	314.03	-1.53	313.27	-0.77	313.08	-0.58
37	319.7	318.77	0.93	317.00	2.70	319.59	0.11	318.63	1.07	319.87	-0.17
38	316.9	316.77	0.13	317.00	-0.10	318.06	-1.16	318.81	-1.91	315.00	1.90
39	317	317.13	-0.13	317.00	0.00	316.54	0.46	316.54	0.46	316.84	0.16
40	295.4	292.70	2.70	291.36	4.04	294.88	0.52	293.57	1.83	294.64	0.76
41	314.2	310.62	3.58	309.90	4.30	313.52	0.68	312.20	2.00	313.75	0.45
42	294.4	294.96	-0.56	296.60	-2.20	296.18	-1.78	296.34	-1.94	293.69	0.71
43	305.5	304.52	0.98	303.00	2.50	305.64	-0.14	304.64	0.86	305.71	-0.21
44	306.5	306.66	-0.16	303.00	3.50	307.77	-1.27	307.58	-1.08	308.48	-1.98
45	301.5	300.58	0.92	301.96	-0.46	300.88	0.62	301.26	0.24	301.03	0.47
46	302.1	301.49	0.61	301.96	0.14	302.22	-0.12	302.12	-0.02	301.19	0.91
47	302.6	301.57	1.03	303.00	-0.40	302.48	0.12	302.48	0.12	301.31	1.29
48	301.8	300.58	1.22	301.96	-0.16	301.12	0.68	301.44	0.36	300.16	1.64
49	303.3	302.01	1.29	303.00	0.30	303.00	0.30	303.00	0.30	303.56	-0.26
50	302.6	300.38	2.22	303.00	-0.40	301.33	1.27	301.91	0.69	301.31	1.29
51	305.8	307.09	-1.29	308.11	-2.31	305.95	-0.15	306.76	-0.96	304.94	0.86
52	293.4	293.67	-0.27	296.60	-3.20	295.65	-2.25	295.98	-2.58	293.53	-0.13

53	293.8	295.40	-1.60	298.62	-4.82	296.90	-3.10	297.57	-3.77	295.13	-1.33
54	294.9	296.93	-2.03	298.62	-3.72	297.78	-2.88	298.05	-3.15	295.87	-0.97
55	297.5	298.54	-1.04	298.62	-1.12	298.69	-1.19	298.67	-1.17	297.15	0.35
56	297.4	297.46	-0.06	298.62	-1.22	298.12	-0.72	298.31	-0.91	297.88	-0.48
57	297.1	297.59	-0.49	296.60	0.50	298.20	-1.10	297.60	-0.50	296.60	0.50
58	300.7	299.43	1.27	296.60	4.10	299.92	0.78	299.88	0.82	299.82	0.88
59	303.7	303.14	0.56	303.00	0.70	304.31	-0.61	303.62	0.08	302.77	0.93
60	309.3	309.74	-0.44	311.50	-2.19	309.80	-0.50	309.80	-0.50	308.81	0.49
61	311.1	311.64	-0.54	311.50	-0.39	311.02	0.08	310.57	0.53	311.41	-0.31
62	314.1	312.93	1.17	311.50	2.61	313.09	1.01	312.49	1.61	314.22	-0.12
63	289.3	290.23	-0.93	293.65	-4.35	292.94	-3.64	293.22	-3.92	291.85	-2.55
64	287.7	288.85	-1.15	291.36	-3.66	291.79	-4.09	292.06	-4.36	290.60	-2.90
65	306.2	306.75	-0.55	305.00	1.20	307.17	-0.97	306.36	-0.16	304.91	1.29
66	306.2	306.75	-0.55	305.00	1.20	307.17	-0.97	306.36	-0.16	304.91	1.29
67	304.4	305.54	-1.14	305.00	-0.60	306.15	-1.75	305.34	-0.94	306.78	-2.38
68	305.9	308.68	-2.78	305.00	0.90	309.35	-3.45	309.35	-3.45	307.28	-1.38
69	307.6	308.86	-1.26	305.04	2.56	309.27	-1.67	309.26	-1.66	309.19	-1.59
70	308.5	309.37	-0.87	313.46	-4.96	310.58	-2.08	311.63	-3.13	310.06	-1.56
71	305.1	306.17	-1.07	305.00	0.10	306.66	-1.56	306.04	-0.94	305.00	0.10
72	307.4	306.70	0.70	305.04	2.36	307.15	0.25	306.35	1.05	307.12	0.28
73	308.2	307.89	0.31	305.04	3.16	308.59	-0.39	308.83	-0.63	308.71	-0.51
74	307.7	308.48	-0.78	305.04	2.66	309.25	-1.55	309.25	-1.55	307.12	0.58
75	297.6	296.16	1.44	294.93	2.67	297.64	-0.04	296.68	0.92	297.22	0.38
76	294	293.46	0.54	294.93	-0.93	294.93	-0.93	294.93	-0.93	294.75	-0.75
77	298.8	298.43	0.37	294.93	3.87	299.81	-1.01	299.87	-1.07	299.49	-0.69
78	301.3	301.21	0.09	305.00	-3.70	301.94	-0.64	302.71	-1.41	299.71	1.59
79	298.2	297.56	0.64	293.65	4.55	299.34	-1.14	299.34	-1.14	298.99	-0.79
80	302.5	301.52	0.98	305.00	-2.50	302.41	0.09	303.40	-0.90	302.25	0.25
81	305.9	308.68	-2.78	305.00	0.90	309.35	-3.45	309.35	-3.45	309.51	-3.61
82	307.1	310.74	-3.64	313.46	-6.36	311.40	-4.30	312.18	-5.08	309.46	-2.36
83	310.9	311.10	-0.20	313.46	-2.56	312.57	-1.67	312.91	-2.01	312.94	-2.04
84	310.4	310.58	-0.18	313.69	-3.29	311.52	-1.12	312.33	-1.93	311.72	-1.32
85	318.2	317.21	0.99	313.69	4.51	317.29	0.91	317.61	0.59	317.74	0.46
86	312.7	312.75	-0.05	313.46	-0.76	313.52	-0.82	313.50	-0.80	313.46	-0.76

87	321.1	319.94	1.16	322.48	-1.38	320.31	0.79	321.15	-0.05	320.75	0.35
88	316.9	316.69	0.21	313.46	3.44	317.97	-1.07	317.97	-1.07	315.78	1.12
89	321.4	321.29	0.11	322.48	-1.08	322.22	-0.82	322.32	-0.92	320.05	1.35
90	318	318.51	-0.51	321.44	-3.44	319.39	-1.39	320.36	-2.36	316.55	1.45
91	306.2	306.75	-0.55	305.00	1.20	307.17	-0.97	306.36	-0.16	304.91	1.29
92	306.2	306.75	-0.55	305.00	1.20	307.17	-0.97	306.36	-0.16	304.91	1.29
93	309	309.42	-0.42	309.90	-0.90	311.68	-2.68	311.68	-2.68	309.20	-0.20
94	322.1	322.64	-0.54	322.62	-0.52	322.62	-0.52	322.62	-0.52	322.65	-0.55
95	321.4	322.13	-0.73	322.62	-1.22	321.43	-0.03	321.88	-0.48	321.36	0.04
96	293.1	292.17	0.93	287.85	5.25	295.66	-2.56	295.66	-2.56	291.40	1.70
97	297.5	296.85	0.65	300.13	-2.63	298.68	-1.18	299.82	-2.32	295.23	2.27
98	293.3	292.37	0.93	294.32	-1.02	294.78	-1.48	295.70	-2.40	294.40	-1.10
99	297.5	297.23	0.27	300.13	-2.63	298.68	-1.18	299.23	-1.73	297.10	0.40
100	312.3	310.10	2.20	309.90	2.40	312.79	-0.49	311.70	0.60	312.64	-0.34
101	318.8	319.61	-0.81	321.44	-2.64	320.33	-1.53	320.75	-1.95	317.12	1.68
102	316.3	315.79	0.51	317.00	-0.70	316.82	-0.52	317.72	-1.42	316.19	0.11
103	310.3	310.76	-0.46	309.28	1.02	312.00	-1.70	311.10	-0.80	311.92	-1.62
104	309.8	310.09	-0.29	309.90	-0.10	312.39	-2.59	311.51	-1.71	310.12	-0.32
105	310.2	308.66	1.54	309.28	0.92	309.59	0.61	309.59	0.61	309.77	0.43
106	311.1	313.93	-2.83	317.00	-5.90	315.07	-3.97	315.80	-4.70	311.90	-0.80
107	314.1	315.65	-1.55	317.00	-2.90	317.00	-2.90	317.00	-2.90	313.69	0.41
108	312.1	310.72	1.38	310.03	2.07	311.77	0.33	311.12	0.98	311.82	0.28
109	315.8	312.92	2.88	309.90	5.90	315.04	0.76	315.28	0.52	315.35	0.45
110	316.4	315.65	0.75	317.00	-0.60	317.00	-0.60	317.00	-0.60	317.01	-0.61
111	317.1	316.26	0.84	317.00	0.10	315.28	1.82	315.97	1.13	316.59	0.51
112	289.7	289.66	0.04	290.31	-0.61	293.21	-3.51	293.30	-3.60	291.08	-1.38
113	289.8	289.66	0.14	290.31	-0.51	293.21	-3.41	293.30	-3.50	291.17	-1.37
114	292.6	289.41	3.19	290.31	2.29	292.74	-0.14	291.86	0.74	291.94	0.66
115	293.7	292.92	0.78	294.22	-0.52	295.32	-1.62	294.91	-1.21	294.86	-1.16
116	289.3	288.43	0.87	289.81	-0.51	291.91	-2.61	291.84	-2.54	290.56	-1.26
117	287.9	287.76	0.14	289.81	-1.91	291.40	-3.50	291.52	-3.62	290.06	-2.16
118	289.2	289.02	0.18	290.00	-0.80	292.28	-3.08	292.35	-3.15	290.54	-1.34
119	289.2	289.41	-0.21	290.00	-0.80	292.71	-3.51	292.62	-3.42	291.76	-2.56
120	293.2	291.49	1.71	291.96	1.24	291.16	2.04	291.50	1.70	293.00	0.20

121	305.7	306.23	-0.53	303.00	2.70	303.55	2.15	303.34	2.36	304.87	0.83
122	305.2	306.23	-1.03	303.00	2.20	303.55	1.65	303.34	1.86	304.12	1.08
123	301.7	302.47	-0.77	303.00	-1.30	299.52	2.18	300.81	0.89	299.63	2.07
124	305.6	305.76	-0.16	305.77	-0.17	305.77	-0.17	305.77	-0.17	305.66	-0.06
125	306.2	306.03	0.17	305.77	0.43	304.95	1.25	305.26	0.94	305.47	0.73
126	301.2	302.47	-1.27	303.00	-1.80	299.52	1.68	300.81	0.39	299.35	1.85
127	304.4	304.50	-0.10	304.12	0.28	304.15	0.25	304.12	0.28	304.00	0.40
128	303.4	303.45	-0.05	304.09	-0.69	301.80	1.60	301.71	1.69	302.35	1.05
129	304.9	304.78	0.12	304.77	0.13	304.43	0.47	304.43	0.47	304.52	0.38
130	304.7	304.04	0.66	304.00	0.70	304.00	0.70	304.00	0.70	304.26	0.44
131	304.3	304.69	-0.39	304.09	0.21	304.20	0.10	304.17	0.13	304.46	-0.16
132	304.3	304.06	0.24	304.00	0.30	304.02	0.28	304.02	0.28	304.38	-0.08
133	304.3	304.02	0.28	304.00	0.30	303.14	1.16	303.49	0.81	303.91	0.39
134	303.3	303.36	-0.06	304.09	-0.79	302.02	1.28	302.02	1.28	302.32	0.98
135	304	304.08	-0.08	304.09	-0.09	302.82	1.18	303.30	0.70	303.24	0.76
136	301.8	302.00	-0.20	299.01	2.79	300.28	1.52	299.80	2.00	300.70	1.10
137	302.3	302.29	0.01	299.01	3.29	301.01	1.29	300.63	1.67	301.10	1.20
138	301.7	302.06	-0.36	301.00	0.70	301.00	0.70	301.00	0.70	301.90	-0.20
139	296.2	298.08	-1.88	299.01	-2.81	297.45	-1.25	298.01	-1.81	296.77	-0.57
140	296.8	295.92	0.88	299.01	-2.21	296.13	0.67	296.13	0.67	297.68	-0.88
141	293.3	292.53	0.77	290.31	2.99	294.20	-0.90	293.47	-0.17	294.97	-1.67
142	297.7	297.93	-0.23	301.00	-3.30	297.61	0.09	297.61	0.09	297.20	0.50
143	294.6	293.26	1.34	294.22	0.38	294.22	0.38	294.22	0.38	294.85	-0.25
144	304.9	305.05	-0.15	304.12	0.78	304.35	0.55	304.38	0.52	304.54	0.36
145	304.9	305.05	-0.15	304.12	0.78	304.35	0.55	304.38	0.52	304.54	0.36
146	303.9	302.83	1.07	304.12	-0.22	303.12	0.78	303.42	0.48	302.35	1.55
147	303.1	302.65	0.45	304.12	-1.02	302.83	0.27	303.31	-0.21	301.97	1.13
148	304.9	304.40	0.50	304.77	0.13	303.24	1.66	303.69	1.21	303.72	1.18
149	305.3	305.20	0.10	304.77	0.53	304.91	0.39	304.86	0.44	304.96	0.34
150	304.9	305.68	-0.78	302.37	2.53	303.11	1.79	302.83	2.07	304.41	0.49
151	304.9	304.90	0.00	302.37	2.53	302.27	2.63	301.92	2.98	304.09	0.81
152	304.7	305.14	-0.44	304.12	0.58	304.45	0.25	304.45	0.25	304.88	-0.18
153	304.9	305.05	-0.15	304.12	0.78	304.35	0.55	304.38	0.52	304.54	0.36
154	304.9	305.05	-0.15	304.12	0.78	304.35	0.55	304.38	0.52	304.54	0.36

155	305.5	305.43	0.07	305.77	-0.27	305.26	0.24	305.37	0.13	305.61	-0.11
156	301	298.33	2.67	302.37	-1.37	296.12	4.88	296.11	4.89	299.11	1.89
157	292.1	289.62	2.48	289.81	2.29	289.72	2.38	289.75	2.35	292.87	-0.77
158	296.8	297.97	-1.17	302.37	-5.57	295.85	0.95	295.94	0.86	295.67	1.13
159	295.9	296.51	-0.61	300.31	-4.41	295.16	0.74	295.16	0.74	294.96	0.94
160	296	295.59	0.41	300.31	-4.31	294.91	1.09	294.91	1.09	296.74	-0.74
161	296.3	296.40	-0.10	293.47	2.83	294.67	1.63	294.22	2.08	294.87	1.43
162	302.7	302.35	0.35	304.10	-1.40	302.29	0.41	303.00	-0.30	302.70	0.00
163	302.9	303.38	-0.48	302.23	0.67	301.63	1.27	301.88	1.02	301.65	1.25
164	302.9	303.38	-0.48	302.23	0.67	301.63	1.27	301.88	1.02	301.65	1.25
165	298.8	299.74	-0.94	302.23	-3.43	297.85	0.95	297.85	0.95	297.09	1.71
166	305	304.97	0.03	304.56	0.44	305.17	-0.17	305.17	-0.17	305.12	-0.12
167	304.5	304.03	0.47	302.45	2.05	302.73	1.77	302.73	1.77	303.65	0.85
168	304.7	304.73	-0.03	304.56	0.14	304.85	-0.15	304.75	-0.05	304.96	-0.26
169	300.9	301.93	-1.03	302.45	-1.55	299.72	1.18	300.85	0.05	299.54	1.36
170	304.4	304.58	-0.18	302.23	2.17	302.89	1.51	302.68	1.72	303.27	1.13
171	305.3	304.88	0.42	305.00	0.30	304.18	1.12	304.48	0.82	304.62	0.68
172	305.1	304.88	0.22	305.00	0.10	304.94	0.16	304.96	0.14	305.01	0.09
173	305.4	305.27	0.13	305.00	0.40	305.39	0.01	305.39	0.01	305.49	-0.09
174	304.4	304.27	0.13	304.73	-0.33	303.97	0.43	304.27	0.13	303.60	0.80
175	303.9	303.62	0.28	302.77	1.13	303.53	0.37	303.26	0.64	303.61	0.29
176	304.8	304.55	0.25	304.73	0.07	304.37	0.43	304.49	0.31	304.77	0.03
177	304.6	304.79	-0.19	304.73	-0.13	304.69	-0.09	304.70	-0.10	304.73	-0.13
178	305.9	305.72	0.18	305.77	0.13	304.90	1.00	305.22	0.68	305.37	0.53
179	305.9	305.72	0.18	305.77	0.13	304.90	1.00	305.22	0.68	305.37	0.53
180	296.9	297.63	-0.73	293.47	3.43	295.39	1.51	294.65	2.25	297.85	-0.95
181	301.9	301.88	0.02	302.77	-0.87	302.24	-0.34	302.44	-0.54	300.75	1.15
182	303.9	303.97	-0.07	302.77	1.13	303.79	0.11	303.66	0.24	303.06	0.84
183	304.4	304.73	-0.33	304.10	0.30	303.82	0.58	303.93	0.47	303.88	0.52
184	303	303.69	-0.69	302.23	0.77	302.34	0.66	302.34	0.66	302.09	0.91
185	302.8	303.00	-0.20	302.77	0.03	302.81	-0.01	302.78	0.02	301.70	1.10
186	293.9	291.82	2.08	291.96	1.94	290.70	3.20	290.70	3.20	294.15	-0.25
187	294.3	292.72	1.58	291.96	2.34	291.33	2.97	291.57	2.73	291.96	2.34
188	304.9	304.79	0.11	302.45	2.45	303.24	1.66	302.94	1.96	303.62	1.28

189	298.1	298.88	-0.78	298.37	-0.27	298.37	-0.27	298.37	-0.27	296.26	1.84
190	293.4	291.88	1.52	289.18	4.22	292.26	1.14	292.83	0.57	290.00	3.40
191	296.5	296.86	-0.36	298.37	-1.87	294.90	1.60	294.53	1.97	293.70	2.80
192	295	293.55	1.45	289.18	5.82	293.84	1.16	293.82	1.18	293.76	1.24
193	297.8	297.48	0.32	298.37	-0.57	296.14	1.66	296.97	0.83	298.37	-0.57
194	298.3	298.88	-0.58	298.69	-0.39	298.53	-0.23	298.53	-0.23	296.39	1.91
195	298.3	298.88	-0.58	298.69	-0.39	298.53	-0.23	298.53	-0.23	296.39	1.91
196	292.3	290.92	1.38	289.44	2.86	291.68	0.62	290.84	1.46	291.66	0.64
197	294.5	295.10	-0.60	298.69	-4.19	296.38	-1.88	297.24	-2.74	292.86	1.64
198	299.3	298.85	0.45	298.69	0.61	299.34	-0.04	299.07	0.23	299.68	-0.38
199	294.2	294.42	-0.22	295.38	-1.18	296.20	-2.00	295.89	-1.69	295.38	-1.18
200	296.4	296.53	-0.13	295.38	1.02	297.22	-0.82	296.53	-0.13	294.65	1.75
201	297.9	298.57	-0.67	295.38	2.52	298.55	-0.65	298.75	-0.85	296.35	1.55
202	302	301.99	0.01	298.69	3.31	301.12	0.88	301.22	0.78	301.70	0.30
203	304.7	305.14	-0.44	304.00	0.70	304.00	0.70	304.00	0.70	304.00	0.70
204	305.3	305.35	-0.05	304.00	1.30	304.00	1.30	304.00	1.30	304.44	0.86
205	304.4	304.45	-0.05	304.00	0.40	304.00	0.40	304.00	0.40	304.01	0.39
206	302.3	302.70	-0.40	301.28	1.02	302.54	-0.24	302.11	0.19	303.32	-1.02
207	304.2	304.46	-0.26	303.94	0.26	304.25	-0.05	304.25	-0.05	304.21	-0.01
208	302.8	302.65	0.15	300.64	2.16	302.37	0.43	302.12	0.68	301.18	1.62
209	304.3	303.56	0.74	300.64	3.66	303.03	1.27	303.03	1.27	303.60	0.70
210	304.5	304.37	0.13	304.00	0.50	304.18	0.32	304.11	0.39	303.25	1.25
211	298.3	298.01	0.29	293.34	4.96	297.51	0.79	297.32	0.98	297.70	0.60
212	294.3	295.95	-1.65	293.34	0.96	295.68	-1.38	294.80	-0.50	293.67	0.63
213	300.8	300.80	0.00	304.00	-3.20	301.74	-0.94	302.63	-1.83	300.28	0.52
214	303.3	303.29	0.01	301.84	1.46	302.92	0.38	302.92	0.38	303.05	0.25
215	304.8	304.30	0.50	304.00	0.80	303.39	1.41	303.63	1.17	303.30	1.50
216	302.3	303.26	-0.96	300.70	1.60	302.49	-0.19	302.44	-0.14	301.32	0.98
217	302.3	303.26	-0.96	300.70	1.60	302.49	-0.19	302.44	-0.14	301.32	0.98
218	295.6	296.66	-1.06	298.93	-3.33	298.21	-2.61	298.44	-2.84	298.22	-2.62
219	302.2	302.24	-0.04	301.28	0.92	301.22	0.98	301.25	0.95	300.84	1.36
220	291.8	292.21	-0.41	292.92	-1.12	294.42	-2.62	293.86	-2.06	293.91	-2.11
221	301.7	302.04	-0.34	300.92	0.78	300.40	1.30	300.12	1.58	299.36	2.34
222	301.7	302.04	-0.34	300.92	0.78	300.40	1.30	300.12	1.58	299.36	2.34

223	302.8	303.24	-0.44	303.94	-1.14	302.61	0.19	302.61	0.19	302.83	-0.03
224	297	296.44	0.56	300.64	-3.64	297.20	-0.20	298.51	-1.51	295.35	1.65
225	303.7	303.75	-0.05	303.94	-0.24	303.19	0.51	303.47	0.23	303.35	0.35
226	304.1	303.74	0.36	304.00	0.10	302.93	1.17	303.32	0.78	303.31	0.79
227	304	303.79	0.21	303.94	0.06	303.06	0.94	303.42	0.58	303.51	0.49
228	300.2	300.84	-0.64	298.66	1.54	300.39	-0.19	300.74	-0.54	299.40	0.80
229	298.7	299.91	-1.21	300.64	-1.94	299.99	-1.29	300.22	-1.52	297.48	1.22
230	302.8	302.13	0.67	304.00	-1.20	301.77	1.03	302.62	0.18	301.37	1.43
231	304.4	304.20	0.20	304.00	0.40	303.99	0.41	303.99	0.41	302.94	1.46
232	302.8	302.30	0.50	298.66	4.14	301.61	1.19	301.50	1.30	302.22	0.58
233	292.4	288.94	3.46	286.02	6.38	291.93	0.47	291.30	1.10	291.97	0.43
234	292.2	292.31	-0.11	293.34	-1.14	293.41	-1.21	293.37	-1.17	291.54	0.66
235	295.7	296.68	-0.98	293.34	2.36	296.99	-1.29	296.99	-1.29	296.74	-1.04
236	297.2	296.68	0.52	293.34	3.86	296.99	0.21	296.99	0.21	296.74	0.46

PHỤ LỤC 3 MỘT SỐ KẾT QUẢ KHÁC VỀ ĐÁNH GIÁ ĐỘ CHÍNH XÁC CỦA CÁC GRID DEM

PL3.1. So sánh sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả của các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic, Nearest Neighbor và của phương pháp sử dụng mô hình HNN (mạng neuron Hopfiled)



Hình PL3-1. Sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả của các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic, Nearest Neighbor và của phương pháp sử dụng mô hình HNN (mô hình mạng neuron Hopfiled)

Trong đó:

(a)- Sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả sau khi tái chia mẫu song tuyến Bilinear (B-DEM);

(b)- Sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả sau khi tái chia mẫu Bi-cubic (C-DEM);

(c)- Sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả của phương pháp sử dụng mô hình mạng neuron Hopfiled (D-DEM);

(d)- Sự khác nhau giữa DEM tham khảo (Reference DEM) và DEM kết quả của phương pháp nội suy Người láng giềng gần nhất (Nearest Neighbor) (N-DEM).
PL3.2. Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả của các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic, Nearest Neighbor và của phương pháp sử dụng mô hình HNN (mạng neuron Hopfiled)



Hình PL3-2. Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả của các phương pháp tái chia mẫu song tuyến Bilinear, Bi-cubic, Nearest Neighbor và của phương pháp sử dụng mô hình HNN (mạng neuron Hopfiled)

Trong đó:

(a)-Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả sau khi tái chia mẫu song tuyến Bilinear (B-DEM);

(b)-Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả sau khi tái chia mẫu Bi-cubic (C-DEM);

(c)-Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả của phương pháp sử dụng mạng neuron Hopfiled (D-DEM)

(d)-Biểu đồ thống kê số lượng các pixel theo giá trị khác biệt giữa DEM tham khảo (Reference DEM) và DEM kết quả của phương pháp Nearest Neighbor (N-DEM).