

**ĐẠI HỌC QUỐC GIA HÀ NỘI  
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

**TRẦN TUẤN VINH**

**NGHIÊN CỨU VÀ ỨNG DỤNG ẢNH LIDAR TRONG  
GIÁM SÁT SOL KHÍ NÔNG NGHIỆP**

**Chuyên ngành: Khoa học máy tính**

**Mã số: 9480101.01**

**TÓM TẮT LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN**

**Hà Nội – 2022**

Công trình được hoàn thành tại: Trường Đại học Công nghệ,  
Đại học Quốc gia Hà Nội

Người hướng dẫn khoa học:

GS. TS NGUYỄN THANH THUY

PGS. TS NGUYỄN THỊ NHẬT THANH

Phản biện:.....

.....

Phản biện:.....

.....

Phản biện:.....

.....

Luận án sẽ được bảo vệ trước Hội đồng cấp Đại học  
Quốc gia chấm luận án tiến sĩ họp tại .....  
vào hồi     giờ     ngày     tháng     năm 2022

Có thể tìm hiểu luận án tại:

- Thư viện Quốc gia Việt Nam
- Trung tâm Thông tin - Thư viện, Đại học Quốc gia Hà Nội

## MỞ ĐẦU

### 1. Tính cấp thiết của vấn đề nghiên cứu

Ô nhiễm không khí là một vấn đề lớn và cấp thiết mà tất cả các quốc gia trên thế giới đang phải đối mặt. Hậu quả là ô nhiễm không khí đã ảnh hưởng đến sức khỏe và cuộc sống của con người, gây thiệt hại trên diện rộng cho thảm thực vật, cây trồng, động vật hoang dã, vật liệu, công trình xây dựng và khí hậu, đồng thời dẫn đến cạn kiệt nguồn tài nguyên thiên nhiên khan hiếm cần thiết cho phát triển kinh tế lâu dài.

Việt Nam là một nước đang phát triển, do đó, các vấn đề ô nhiễm môi trường nói chung và vấn đề ô nhiễm không khí nói riêng cũng đang xuất hiện ngày càng nhiều. Vì vậy, chúng ta cần thiết phải có giải pháp để giảm thiểu tác động của ô nhiễm không khí.

Kể từ giữa những năm 1990, một hệ thống vệ tinh quan sát Trái đất mới đã có thể phát hiện ô nhiễm không khí tầng đối lưu ở độ phân giải không gian và thời gian ngày càng cao. Hầu hết các chất phát thải ô nhiễm không khí chính có thể được đo bằng một hoặc nhiều thiết bị. Đây là cơ sở để chúng ta có thể liên kết các phép đo vệ tinh với việc định lượng lượng khí thải chính để cải thiện khả năng sử dụng các phép đo tại các trạm quan trắc và các nghiên cứu thực địa trong xây dựng kế hoạch quản lý chất lượng không khí.

Dựa trên sự phân loại sol khí và sự hình thành các loại sol khí này, các nhà nghiên cứu có thể dự đoán được nguồn gốc phát thải các loại sol khí này, qua đó xác định nguồn phát thải chủ yếu của khu vực nghiên cứu.

Cho đến thời điểm hiện nay, chưa có một nghiên cứu nào về sử dụng phương pháp phân loại sol khí của vệ tinh CALIPSO trong việc xác định các nguồn phát thải tại Việt Nam, do đó luận án tập trung vào việc sử dụng nguồn dữ liệu sẵn có này trong việc phát triển phương pháp xác định các nguồn phát thải ô nhiễm không khí tại Việt Nam.

Bên cạnh sự phân loại sol khí để xác định nguồn gốc phát thải chất gây ô nhiễm không khí, thì một thành phần gây ô nhiễm không khí khác cũng rất quan trọng và được nhiều nghiên cứu đề cập đến đó là  $NO_2$ . Các thông số quan trắc chất lượng không khí xung quanh trong những năm gần đây đã cho thấy nồng độ  $NO_2$  cũng có xu hướng tăng trong những năm gần đây tại các đô thị ở Việt Nam.

Giám sát  $NO_2$  trong không gian sử dụng ảnh vệ tinh Sentinel-5P chưa được nghiên cứu tại Việt Nam. Trong luận án này, tôi tiến hành phân tích bản đồ ô nhiễm  $NO_2$  tại Việt Nam theo thời gian để đánh giá ảnh hưởng của dịch COVID-19 đến chất lượng không khí, tập trung phân tích và xác định các nguồn phát thải từ công nghiệp và giao thông với tình trạng ô nhiễm không khí trong giai đoạn này.

Bên cạnh sử dụng dữ liệu viễn thám, phương pháp thực địa cũng được áp dụng một cách hiệu quả. Căn cứ trên các nghiên cứu hiện có, việc ước tính thực địa toàn diện về rơm rạ ở phạm vi cấp tỉnh vẫn chưa được thực hiện cho miền Bắc Việt Nam. Trong nghiên cứu này, tôi đã thực hiện một phương pháp đơn giản để tiến hành thu thập mẫu trên thực địa một cách khoa học và hiệu quả. Kết quả này được sử dụng làm đầu vào để tính toán phát thải do việc đốt rơm rạ khu vực Hà Nội. Phương pháp tiếp cận dựa trên thực địa được áp dụng cho lúa hai vụ điển hình ở Hà Nội.

Riêng đối với các khu vực đô thị và khu công nghiệp, nguồn dữ liệu viễn thám có độ phân giải cao (nhỏ hơn hoặc bằng 30 m) có một ý nghĩa rất quan trọng trong việc xác định mức độ ô nhiễm không khí. Một trong các phương pháp áp dụng để xác định mức độ ô nhiễm của không khí trong khu vực đô thị, khu công nghiệp là ước tính độ sâu quang học của sol khí, đồng thời xác định nồng độ của các hạt  $PM$ . Ở Việt Nam cho tới thời điểm hiện nay vẫn chưa có một nghiên cứu nào về việc phát triển thuật toán ước tính độ sâu quang học của sol khí cho ảnh vệ tinh có độ phân giải cao, đặc biệt là áp dụng cho khu vực đô thị có bề mặt phức tạp. Do đó, tôi đã tập trung vào nghiên cứu và đề xuất được một thuật toán ước tính độ sâu quang học của sol khí đáp ứng các yêu cầu trên.

Hiện nay, các giá trị AOD do vệ tinh ước tính được sử dụng rộng rãi để đánh giá mức độ phơi nhiễm của dân số với  $PM_{2.5}$ , đặc biệt là ở những khu vực thiếu các phép đo dựa trên các trạm mặt đất. Vì ô nhiễm vật chất hạt bao gồm sol khí, nên có thể phát triển mối quan hệ chặt chẽ giữa nồng độ AOD và  $PM_{2.5}$  [31] và do đó sử dụng bộ dữ liệu AOD có nguồn gốc từ vệ tinh để tạo bản đồ phân phối  $PM_{2.5}$ .

Qua phân tích trên, chúng ta thấy được sự cần thiết của việc dự đoán nồng độ  $PM_{2.5}$  dựa trên nguồn dữ liệu AOD độ có độ phân giải cao phù hợp với khu vực đô thị, do đó trong luận án này tôi sử dụng kết quả đạt được trong việc ước tính AOD có độ phân giải 30 m vào việc đề xuất một mô hình dự đoán  $PM_{2.5}$  phù hợp nhất.

## **2. Mục tiêu nghiên cứu của luận án**

*Luận án tập trung vào các mục tiêu chính sau:*

- Nghiên cứu và đề xuất phương pháp xác định nguồn phát thải của các thành phần gây ô nhiễm không khí từ dữ liệu vệ tinh và dữ liệu thực địa.
- Nghiên cứu và phát triển thuật toán ước tính độ sâu quang học của sol khí từ vệ tinh có độ phân giải cao.
- Nghiên cứu và đề xuất mô hình dự đoán nồng độ  $PM_{2.5}$  dựa vào AOD tại khu vực đô thị, khu công nghiệp.

### **3. Cấu trúc của luận án**

Nội dung của luận án, ngoài phần mở đầu và kết luận, bao gồm 4 chương:

#### Chương 1: Cơ sở lý thuyết

Trình bày về cơ sở lý thuyết chung liên quan đến đề tài của nghiên cứu sinh.

#### Chương 2: Xác định nguồn phát thải ô nhiễm không khí

Trong chương này, luận án trình bày các phương pháp xác định nguồn phát thải. Phương pháp sử dụng dữ liệu viễn thám dựa vào việc phân tích dữ liệu vệ tinh CALIPSO với thuật toán phân loại sol khí và phân tích dữ liệu  $NO_2$  từ vệ tinh Sentinel-5P, bên cạnh đó kết quả của phương pháp xác định hệ số phát thải qua dữ liệu thực địa đóng góp một phần rất quan trọng trong việc xác định nguồn phát thải tại khu vực đồng ruộng sau thu hoạch các vụ lúa.

#### Chương 3: Ước tính độ sâu quang học của sol khí

Trình bày phương pháp ước tính độ sâu quang học của sol khí từ dữ liệu vệ tinh có độ phân giải cao 30 m là Landsat 8, tập trung vào việc xây dựng mối quan hệ phản xạ bề mặt mới cho dữ liệu Landsat 8

và áp dụng thuật toán lọc mây Fmask kết hợp với mô hình độ cao số Aster trong việc ước tính AOD của Landsat 8.

#### Chương 4: Dự đoán nồng độ bụi $PM_{2.5}$ của khu vực đô thị

Trình bày các mô hình học máy để dự đoán nồng độ  $PM_{2.5}$  dựa trên dữ liệu độ sâu quang học của Landsat 8 và mô hình khí tượng WRF cho khu vực đô thị, khu công nghiệp điển hình là Bắc Ninh, trên cơ sở các kết quả thu được, luận án đề xuất mô hình phù hợp cho việc dự đoán nồng độ  $PM_{2.5}$ .

#### 4. Đóng góp của luận án

Những đóng góp khoa học chính của luận án bao gồm:

- Nghiên cứu và đề xuất phương pháp xác định nguồn phát thải chất gây ô nhiễm không khí từ dữ liệu vệ tinh CALIPSO và Sentinel-5P và dữ liệu quan trắc thực địa ở Việt Nam [TTV-1, TTV-2, TTV-3, TTV-4];
- Nghiên cứu và phát triển thuật toán ước tính độ sâu quang học của sol khí từ dữ liệu vệ tinh độ phân giải cao Landsat 8 ở khu vực đô thị, khu công nghiệp [TTV-5].
- Nghiên cứu và đề xuất mô hình dự đoán nồng độ  $PM_{2.5}$  từ độ sâu quang học của sol khí thu được từ Landsat 8.

## **CHƯƠNG 1: CƠ SỞ LÝ THUYẾT**

### **1.1. Tổng quan về ô nhiễm không khí**

#### **1.1.1. Ô nhiễm không khí**

#### **1.1.2. Các chất gây ô nhiễm không khí và quá trình gây ô nhiễm**

#### **1.1.3. Nguồn gốc các chất gây ô nhiễm không khí**

#### **1.1.4. Ảnh hưởng của ô nhiễm không khí**

##### **1.1.4.1. Ảnh hưởng tới sức khỏe**

##### **1.1.4.2. Ảnh hưởng tới môi trường**

#### **1.1.5. Hiện trạng ô nhiễm không khí**

##### **1.1.5.1. Trên thế giới**

##### **1.1.5.2. Ở Việt Nam**

### **1.2. Tổng quan về các phương pháp xác định nguồn phát thải**

#### **1.2.1. Xác định nguồn phát thải từ dữ liệu vệ tinh CALIPSO**

#### **1.2.2. Phân tích dữ liệu $NO_2$ của vệ tinh Sentinel-5P để xác định nguồn phát thải do hoạt động công nghiệp và giao thông**

#### **1.2.3. Kiểm kê phát thải từ đốt rơm rạ**

### **1.3. Tổng quan về các phương pháp ước tính độ sâu quang học của sol khí**

#### **1.3.1. Vệ tinh độ phân giải cao [84]**

#### **1.3.2. Thuật toán ước tính độ sâu quang học của sol khí từ ảnh vệ tinh độ phân giải cao**

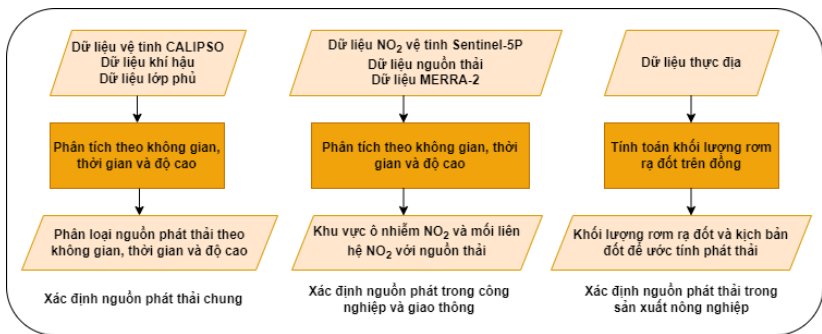
### **1.4. Tổng quan về các phương pháp dự đoán nồng độ bụi $PM_{2.5}$ ở khu vực đô thị**



## CHƯƠNG 2: XÁC ĐỊNH NGUỒN PHÁT THẢI Ô NHIỄM KHÔNG KHÍ

### 2.1. Giới thiệu bài toán xác định nguồn phát thải

Trong chương này, để xác định nguồn phát thải ô nhiễm không khí, tôi đề xuất kết hợp cả hai phương pháp: phương pháp sử dụng phân tích nguồn dữ liệu từ vệ tinh và sử dụng phương pháp bottom-up trong việc xác định nguồn phát thải từ việc đốt rơm rạ sau thu hoạch bằng phương pháp đo đạc thực địa (hình 2.1).



Hình 2.1. Xác định nguồn phát thải ô nhiễm không khí từ dữ liệu thực địa và dữ liệu vệ tinh

### 2.2. Dữ liệu

#### 2.2.1. Dữ liệu vệ tinh

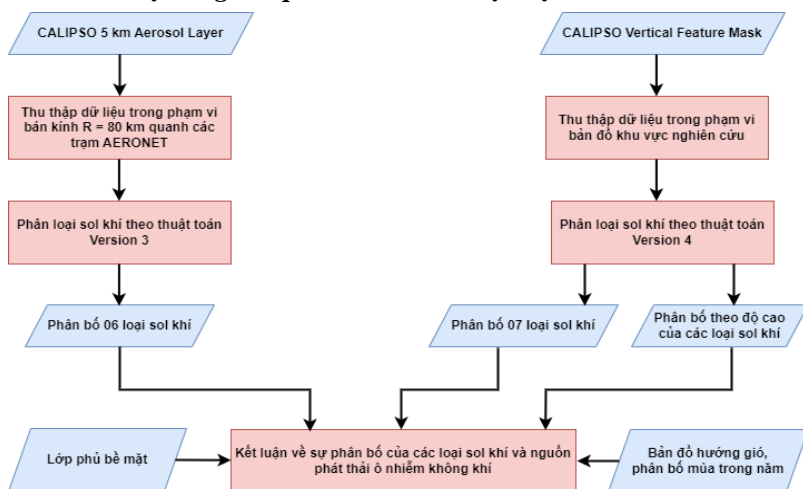
2.2.1.1. CALIPSO (*Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation*)

2.2.1.2. Sentinel-5P

#### 2.2.2. Dữ liệu thực địa

## 2.3. Phương pháp

### 2.3.1. Xác định nguồn phát thải từ dữ liệu vệ tinh CALIPSO

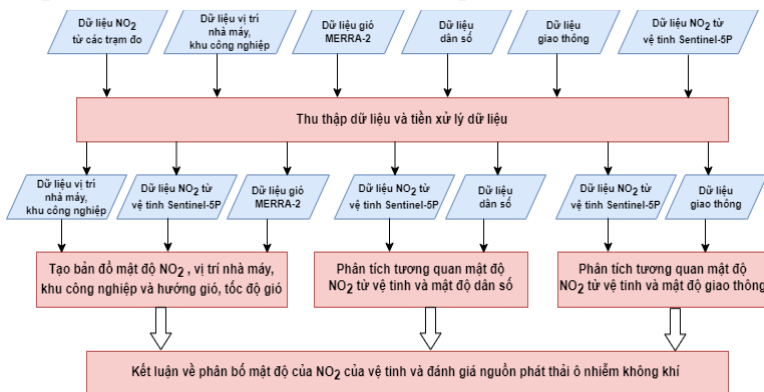


Hình 2.5. Xác định nguồn phát thải ô nhiễm không khí từ vệ tinh CALIPSO

2.3.1.1. Đối với dữ liệu CALIOP Aerosol Layer Level 2 áp dụng thuật toán phân loại sol khí của CALIOP Version 3

2.3.1.2. Đối với dữ liệu CALIOP Vertical Feature Mask Level 2 áp dụng thuật toán phân loại sol khí của CALIOP Version 4

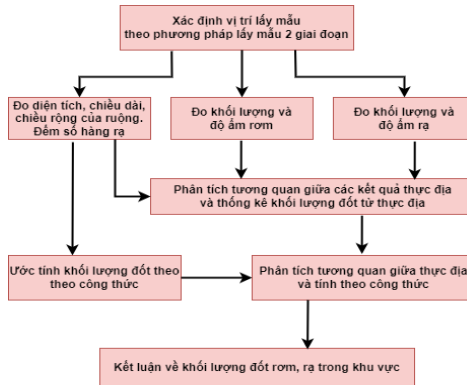
**2.3.2. Phân tích dữ liệu  $NO_2$  của vệ tinh Sentinel-5P để xác định nguồn phát thải do hoạt động công nghiệp và giao thông**



Hình 2.7. Xác định nguồn phát thải  $NO_2$

### 2.3.3. Sử dụng dữ liệu quan trắc thực địa

#### 2.3.3.1. Phương pháp lấy mẫu theo cụm hai giai đoạn (two-stage cluster sampling)



Hình 2.8. Phương pháp xác định khối lượng đốt rơm, rạ từ thực địa

Sử dụng phương pháp lấy mẫu theo cụm hai giai đoạn, thực hiện việc chia khu vực nghiên cứu thành các ô lưới có diện tích bằng nhau 5 km và chọn ngẫu nhiên 13 ô, trong đó bốn cánh đồng thu hoạch bằng máy được chọn ngẫu nhiên lấy mẫu trong mỗi ô cho tổng số 52 mẫu ruộng (C.I. = 9,6%, C.L. = 95%) [115].

Ước tính lượng rơm trên một mét vuông cho một ruộng lúa nhất định dựa trên (phương trình 2.1) sau:

$$Q_{rs} = \frac{(RS_w \times (1 - RS_m)) \times SR \times R_l}{A} \quad (2.1)$$

trong đó  $Q_{rs}$  là lượng rơm rạ khô tính bằng kg/m<sup>2</sup> cho một ruộng lúa,  $RS_w$  là trọng lượng rơm rạ ướt trên mét tính theo một hàng rơm được tính trung bình từ các phép đo mẫu 0,5 m × 0,5 m,  $RS_m$  là độ ẩm tương đối trung bình,  $SR$  là số hàng rơm,  $R_l$  là chiều dài của hàng rơm tính bằng mét, và  $A$  là diện tích ruộng tính bằng m<sup>2</sup>.

Ước tính số lượng gốc rạ cho một thửa ruộng nhất định dựa trên một phương trình tương tự (phương trình 2.2):

$$Q_{sr} = SR_w \times (1 - SR_m) \times A \quad (2.2)$$

trong đó  $Q_{sr}$  là lượng gốc rạ khô tính bằng kg cho một thửa ruộng nhất định,  $SR_w$  là trọng lượng trung bình của gốc rạ trên m<sup>2</sup> từ các phép đo mẫu 0,5 m × 0,5 m,  $SR_m$  là độ ẩm tương đối trung bình của gốc rạ đo được và  $A$  là diện tích ruộng đo bằng m<sup>2</sup>. Kết quả  $Q_{sr}$  mang lại khối lượng đốt gốc rạ tối đa tính bằng kg.

## **2.4. Kết quả**

### **2.4.1. Phương pháp sử dụng dữ liệu vệ tinh CALIPSO**

*2.4.1.1. Sử dụng thuật toán Version 3 với dữ liệu CALIOP Aerosol Layer Level 2 từ 2006 – 2015 phân loại sol khí trên 3 miền của Việt Nam*

Ở miền Bắc Việt Nam, kết quả thu được về sự phân bố loại sol khí hàng tháng trong khu vực bán kính 80 km xung quanh trạm Nghĩa Đô cụ thể như sau: các loại sol khí biomass burning, polluted dust và polluted continental chiếm phần lớn sol khí. Biomass burning xuất hiện thường xuyên hơn vào tháng 3 và tháng 10, do thói quen đốt phụ phẩm nông nghiệp [116]. Polluted continental đặc trưng cho nền ô nhiễm đô thị vào mùa khô cao hơn mùa mưa. Polluted dust, hỗn hợp khói bụi sa mạc hoặc ô nhiễm đô thị, được quan sát thấy từ tháng 4 đến tháng 8 khi nhiệt độ nền cao, nhiệt độ cao cùng với gió làm ô nhiễm không khí di chuyển lên cao và hạn chế phát tán. Clean continental, sol khí nền hoặc vùng nông thôn, xuất hiện trong các khoảng thời gian từ cuối tháng 1 đến đầu tháng 2, cuối tháng 7 và đầu tháng 11, liên quan đến các giai đoạn bón lót và ngâm ủ rơm rạ trong vụ đông xuân và hè thu. Desert dust được quan sát thấy vào tháng 4, tháng 5 và đầu tháng 6, nguyên nhân có thể do bụi sa mạc đến Hà Nội theo quỹ đạo gió kéo dài trở lại các khu vực nội địa phía bắc và tây Trung Quốc và Mông Cổ [117].

Ở miền Trung Việt Nam, sự phân bố các loại sol khí hàng tháng tại khu vực Nha Trang cho thấy trạm Nha Trang nằm ven biển nên sol khí biển có tần suất xuất hiện nhiều hơn từ đầu tháng 10 đến đầu tháng 2 năm sau khi gió mùa Đông Bắc đưa từ biển vào, gió mùa Tây Nam

thời từ đất liền ra biển từ tháng 3 đến tháng 9 nên tỉ lệ sol khí biển thấp hơn. Nền nhiệt độ cao vào tháng 6 và tháng 7 có thể gây ra tần suất polluted dust cao bất thường trong các tháng này. Biomass burning xuất hiện từ tháng 2 và đầu tháng 4 trùng với thời điểm nông dân sống xung quanh (như ở Nam Nóc, Khánh Hòa và ở Sông Hinh, Phú Yên) đốt phụ phẩm nông nghiệp. Các sol khí nông thôn là clean continental do vùng này ít khu vực trồng lúa.

Đối với miền Nam, do Bạc Liêu nằm sâu trong đất liền nên tần suất sol khí biển ít hơn Nha Trang, nhưng tần suất polluted continental và polluted dust cao hơn. Thời tiết nắng nóng trong các tháng 6, 7, 8 kết hợp với mật độ các đô thị cao như Bạc Liêu, Sóc Trăng, Cần Thơ tạo nên hiện tượng đảo nhiệt. Nó gây ra phần lớn polluted dust trong không khí. Biomass burning được quan sát thấy thường xuyên hơn vào cuối tháng 2 đến tháng 3, đầu tháng 6, đầu tháng 11 và tháng 12. Clean continental được đo vào tháng 9 khi tương ứng với thời gian thu hoạch của vụ lúa hè – thu. Thay vì đốt rơm rạ và các phụ phẩm khác, nông dân thường để tại ruộng. Ngoài ra, một diện tích lớn nuôi trồng thủy sản quanh khu vực Bạc Liêu có thể đóng góp vào tổng lượng clean continental trong một năm.

*2.4.1.2. Sử dụng thuật toán Version 3 với dữ liệu CALIOP Aerosol Layer Level 2 từ 2006 – 2015 phân loại sol khí khu vực Hà Nội 2006 – 2015*

*2.4.1.3. Sử dụng thuật toán Version 4 với dữ liệu CALIOP Vertical Feature Mask Level 2 từ 2016 – 2019*

- *Phân loại sol khí theo tháng*

Kết quả thu được cho thấy sự phân bố sol khí hàng tháng ở Hà Nội từ năm 2016 đến năm 2019 theo thuật toán Version 4 của CALIPSO tương tự như thuật toán Version 3.

- *Phân bố loại sol khí theo độ cao*

Kết quả thu được cho thấy các sol khí tập trung chủ yếu dưới 8 km trong khi polluted continental/smoke phân bố ở độ cao dưới 3 km. Điều này phản ánh quy luật ô nhiễm ở các đô thị, trong những điều kiện bất lợi, khí thải từ các hoạt động công nghiệp và giao thông, nông nghiệp. Các loại sol khí dust, polluted dust và elevated smoke ở tầm cao hơn do vận chuyển tầm xa từ nơi khác đến.

Dựa trên các kết quả này, có thể có những dự đoán về sự tồn tại của các loại sol khí cũng như xác định được những nguồn phát thải chủ yếu đối với từng loại ở khu vực Hà Nội.

**2.4.2. Phương pháp sử dụng dữ liệu vệ tinh Sentinel-5P**

*2.4.2.1. Phân tích nguồn khí thải công nghiệp trong nước dựa vào dữ liệu  $NO_2$  của vệ tinh Sentinel-5P*

Các nhà máy này thường tập trung thành cụm nơi có mật độ  $NO_2$  cao hơn các khu vực xung quanh. Tại khu vực Đồng bằng sông Hồng, mức độ ảnh hưởng cao của  $NO_2$  tại các khu vực có nhiều nhà máy đang hoạt động như nhà máy nhiệt điện than, nhà máy thép ở các tỉnh Quảng Ninh, Hải Dương, Hải Phòng và các nhà máy xi măng ở Hà Nam, Ninh Bình và ở trung tâm Hà Nội với mật độ dân số và giao thông cao. Dọc theo các tuyến bờ biển ở các tỉnh Thanh Hóa, Nghệ An và Hà Tĩnh, Quảng Bình, các nhà máy xi măng, nhiệt điện than và các nhà máy thép cũng có mật độ  $NO_2$  cao. Tại Thành phố Hồ Chí Minh, nhiều nhà



máy xi măng và thép đi vào hoạt động cùng với mật độ giao thông cao khiến khu vực đó có mật độ  $\text{NO}_2$  cao hơn các tỉnh lân cận phía Nam.

Ở miền Bắc, gió Đông thổi từ biển trong thời gian này đi qua các nhà máy ở Quảng Ninh, Hải Phòng, Bắc Ninh có thể mang ô nhiễm đến Hải Dương, Hưng Yên và Hà Nội. Trong khi đó, mật độ  $\text{NO}_2$  cao ở Hà Nam, Ninh Bình có thể ảnh hưởng đến Thanh Hóa và Hòa Bình. Ở miền Trung Việt Nam, hướng gió Đông Bắc thịnh hành nên gây ô nhiễm các khu vực đất liền. Tại thành phố Hồ Chí Minh, gió Đông đi qua các cụm nhà máy và trung tâm thành phố có thể làm gia tăng ô nhiễm không khí ở Long An.

*2.4.2.2. Phân tích phát thải do giao thông trong nước và phát thải do con người gây ra dựa vào dữ liệu  $\text{NO}_2$  của vệ tinh Sentinel-5P [TTV-3]*

Tại khu vực đồng bằng Sông Hồng có lưu lượng giao thông và dân số cao nhất Việt Nam, mật độ TROPOMI  $\text{NO}_2$  giảm hầu hết trong thời gian cách ly ( $18,5 \mu\text{mol}/\text{m}^2$ ), điều này có thể được giải thích là do việc giảm giao thông và các hoạt động ngoài trời của người dân do quy định cách ly.

#### ***2.4.3. Phương pháp sử dụng dữ liệu quan trắc thực địa***

Khối lượng đốt trung bình được đo là: rom  $0,27 \text{ kg}/\text{m}^2$ , gốc rạ  $0,61 \text{ kg}/\text{m}^2$ , và tổng dư lượng sau thu hoạch  $0,88 \text{ kg}/\text{m}^2$ . Tôi cũng xem xét các nghiên cứu khác bên ngoài Việt Nam cho thấy khối lượng đốt cao hơn bao gồm  $3600 \text{ kg}/\text{ha}$  và  $5800 \text{ kg}/\text{ha}$  trong hai nghiên cứu riêng biệt ở Thái Lan [111, 112], trong khi nghiên cứu khác ở Đồng bằng sông Cửu Long, Việt Nam có mức trung bình thấp hơn là  $3470 \text{ kg}/\text{ha}$

[103]. Sự khác biệt này cho thấy khối lượng đốt cụ thể của khu vực là cơ sở quan trọng để ước tính lượng khí thải.

Ước tính của tôi kết hợp khối lượng đốt cho rơm, gốc rạ và tổng sinh khối lúa rất hữu ích trong việc xác định khí thải ở Hà Nội và toàn bộ Việt Nam.

## CHƯƠNG 3: ƯỚC TÍNH ĐỘ SÂU QUANG HỌC CỦA SOL KHÍ

### 3.1. Giới thiệu bài toán ước tính độ sâu quang học của sol khí cho khu vực đô thị

Việc phát triển một thuật toán ước tính AOD từ vệ tinh có độ phân giải cao, đặc biệt trong các khu vực đô thị, để giám sát chất lượng không khí là cần thiết và là một thách thức đối với các nhà khoa học.

Các thuật toán ước tính AOD được phát triển dựa trên nguyên lý cơ bản sau: Phản xạ phổ theo góc của TOA (Top Of Atmosphere)  $\rho_{\lambda}(\theta_0, \theta, \phi)$  với  $\theta_0$  là góc thiên đỉnh mặt trời (solar zenith angle),  $\theta$  là góc thiên đỉnh của cảm biến trên vệ tinh (view zenith angle), và  $\phi$  là góc phương vị tương đối (relative azimuth angle) tại bước sóng  $\lambda$  là kết quả của: sự tán xạ của bức xạ trong khí quyển không có tương tác với bề mặt (atmospheric path reflectance), phản xạ của bức xạ ngoài bề mặt truyền trực tiếp tới TOA (surface function) và phản xạ của bức xạ từ bên ngoài trường nhìn của cảm biến vệ tinh (environment function). Hàm môi trường được bỏ qua để có kết quả gần đúng:

$$\rho_{\lambda}^*(\theta_0, \theta, \phi) = \rho_{\lambda}^a(\theta_0, \theta, \phi) + \frac{F_{\lambda}(\theta_0)T_{\lambda}(\theta)\rho_{\lambda}^s(\theta_0, \theta, \phi)}{1 - s_{\lambda}\rho_{\lambda}^s(\theta_0, \theta, \phi)} \quad (3.1)$$

Trong đó:

- $F_{\lambda}$  là thông lượng xuống (downward) chuẩn hóa cho phản xạ bề mặt bằng 0
- $T_{\lambda}$  là hệ số truyền tổng cộng hướng lên vào trong trường nhìn của cảm biến (upward total transmission)
- $s_{\lambda}$  là tỉ số tán xạ ngược trong khí quyển

$\rho_{\lambda}^s$  là phản xạ bề mặt của phổ theo góc.

### 3.2. Dữ liệu

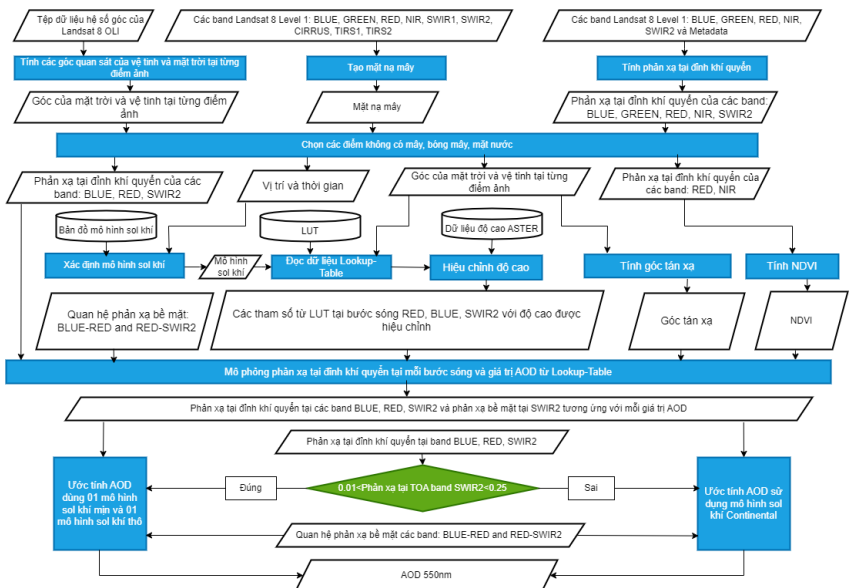
#### 3.2.1. Dữ liệu AERONET Level 2.0 - Version 3

#### 3.2.2. Dữ liệu Landsat 8 Level 1

#### 3.2.3. Dữ liệu MODIS AOD

#### 3.2.4. Dữ liệu DEM Aster

### 3.3. Thuật toán ước tính AOD



Hình 3.2. Sơ đồ thuật toán ước tính AOD từ Landsat 8 OLI

### 3.3.1. Tiền xử lý dữ liệu

#### 3.3.1.1. Tính mặt nạ mây

Sử dụng Fmask

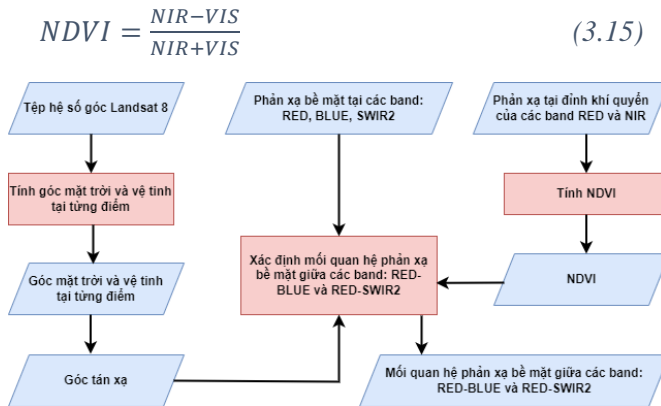
#### 3.3.1.2. Tính các góc của cảm biến vệ tinh và mặt trời

Sử dụng công cụ tính góc của Landsat 8.

#### 3.3.1.3. Tính phản xạ đỉnh khí quyển

### 3.3.2. Xây dựng mối quan hệ phản xạ bề mặt giữa các band BLUE-RED-SWIR2

Trong luận án này, tôi áp dụng tính NDVI cho Landsat 8 dùng band 4 (RED VIS: 0.636 – 0.673) và band 5 (NIR: 0.851 – 0.879) thay vì các band SWIR1 và SWIR2 như ở MODIS, do trong điều kiện không có sol khí thì hai NDVI này có tương quan cao [132].



Hình 3.3. Sơ đồ xác định mối quan hệ phản xạ bề mặt tại các band BLUE, RED và SWIR2

### 3.3.3. Hiệu chỉnh độ cao

### 3.3.4. Mô phỏng TOA Reflectance tại từng bước sóng ứng với từng giá trị AOD từ LUT

### 3.3.5. Ước tính AOD

Theo thuật toán ước tính AOD của MODIS, tôi thực hiện phép đảo ngược phản xạ đa kênh để ước tính các đặc tính của sol khí trên đất liền, kết hợp giữa các mô hình sol khí mịn và thô để phù hợp với độ phản xạ quang phổ quan sát được.

Đối với bề mặt sáng, việc thực hiện ước tính áp dụng với mô hình sol khí Continental.

## 3.4. Kết quả

### 3.4.1. Mất nọ mây

### 3.4.2. Mối quan hệ phản xạ bề mặt giữa các band BLUE-RED-SWIR2

Với kết quả của thực nghiệm, tôi xác định mối quan hệ của hệ số phản xạ bề mặt của các band BLUE-RED-SWIR2 như sau:

$$\text{slope}_{RED/SWIR2} = -0.606570 * NDVI - 0.023431 * \theta + 3.871518 \quad (3.26)$$

$$\text{intercept}_{RED/SWIR2} = 0.002943 * \theta + 0.388459 \quad (3.27)$$

$$\text{slope}_{BLUE/RED} = -0.019279 * \theta + 3.230486 \quad (3.28)$$

$$\text{intercept}_{BLUE/RED} = 0.000223 * \theta - 0.027643 \quad (3.29)$$

$\text{slope}_{RED/SWIR2}$  và  $\text{intercept}_{RED/SWIR2}$  là slope và intercept trong hàm hồi quy tuyến tính của phản xạ bề mặt giữa các band RED và SWIR2 cũng như slope và intercept trong hàm hồi quy tuyến tính của

phản xạ bề mặt giữa các band BLUE và RED, còn NDVI là chỉ số thực vật,  $\Theta$  là góc tán xạ.

### **3.4.3. Đánh giá AOD ước tính từ thuật toán**

#### **3.4.3.1. AOD Landsat 8**

Trong luận án này, tôi áp dụng thuật toán để tạo bản đồ AOD ở khu vực Hà Nội, phần nằm trong scene Landsat 8 có Path/Row là 127/045 tương ứng với các ngày 18/12/2013 và 01/7/2015.

#### **3.4.3.2. Đánh giá AOD của Landsat 8 với trạm AERONET**

Sản phẩm AOD được ước tính từ thuật toán đề xuất được so sánh với ba trạm AERONET tại Việt Nam, Sơn La, Nghĩa Đô và Bạc Liêu từ năm 2014 đến năm 2016 cho kết quả tương quan rất cao. Tại trạm Sơn La, hệ số tương quan  $R^2$  là 0,99 và RMSE là 0,20. Tại trạm Nghĩa Đô, hệ số tương quan  $R^2$  là 0,97 với RMSE là 0,33.

#### **3.4.3.3. Đánh giá AOD của Landsat 8 với AOD của MODIS**

Kết quả cho thấy rằng hầu hết các trường hợp đều xác định mối tương quan giữa AOD Landsat 8 với AOD MODIS khá cao. Cụ thể, các scene có Path/Row 127/045, khi so sánh với MODIS/Terra AOD, cho thấy sự tương quan tốt, vào các ngày 01/07/2015, 07/06/2018, tương ứng hệ số tương quan  $R^2$  là 0,7127, 0,8610. Mặt khác, scene Landsat 8 với Path/Row là 128/045, vào ngày 8 tháng 7 năm 2015 và ngày 30 tháng 6 năm 2018 cho  $R^2$  với MODIS/Terra AOD lần lượt là 0,2431 và 0,6579. Ngoài ra, các chỉ số thống kê khác, sai số trung bình RMSE của kết quả không quá lớn với giá trị từ 0,06 đến 0,26, đa số dao động trong khoảng 0,1 đến 0,2. Độ lệch của dữ liệu sai lệch khá nhỏ, cao nhất là 0,19 và hầu hết đều dưới 0,1.

## CHƯƠNG 4: DỰ ĐOÁN NỒNG ĐỘ BỤI PM CỦA KHU VỰC ĐÔ THỊ

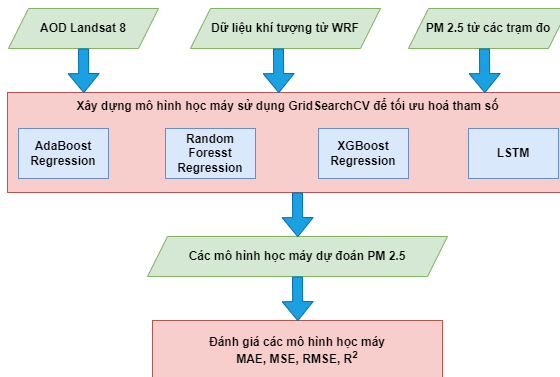
### 4.1. Giới thiệu bài toán dự đoán nồng độ bụi PM<sub>2.5</sub>

### 4.2. Dữ liệu

Dữ liệu nồng độ PM<sub>2.5</sub> thập tại các trạm quan trắc cố định liên tục trên phạm vi tỉnh Bắc Ninh trong khoảng thời gian 2020-2021. Dữ liệu Mô hình nghiên cứu và dự báo thời tiết (Weather Research and Forecasting - WRF).

### 4.3. Phương pháp

Trong nghiên cứu này, tôi xây dựng các mô hình học máy dự đoán nồng độ PM<sub>2.5</sub> theo các phương pháp khác nhau. Trên cơ sở đánh giá kết quả thu được đề xuất việc sử dụng mô hình thích hợp cho việc dự đoán nồng độ PM<sub>2.5</sub> tại vùng nghiên cứu và mở rộng ra trên toàn lãnh thổ Việt Nam.



Hình 4.3. Xây dựng mô hình học máy dự đoán nồng độ PM<sub>2.5</sub> từ dữ liệu AOD của Landsat 8 và dữ liệu khí tượng từ mô hình WRF



### **4.3.1. Mô hình học máy sử dụng trong dự đoán nồng độ $PM_{2.5}$**

### **4.3.2. Sử dụng tham số trong đánh giá mô hình học máy**

## **4.4. Kết quả**

Xây dựng bản đồ nhiệt và phân phối của nồng độ  $PM_{2.5}$ . Dựa trên kết quả có thể nhận thấy  $PM_{2.5}$  có tương quan cao với AOD và áp suất. Kết quả đánh giá các mô hình dự đoán được áp dụng cho thấy  $R^2$  của bốn mô hình đạt giá trị từ 0,86 đến 0,99, có nghĩa là tất cả các mô hình này đều có hiệu suất tốt và rất tốt. Giá trị  $R^2$  cao nhất xác định được 0,99 của mô hình sử dụng phương pháp học sâu LSTM. Các mô hình sau đó được xem xét cho RMSE, tốt hơn với RMSE thấp hơn. Trong các mô hình cây RMSE cao hơn so với mạng nơ-ron, thấp nhất RMSE = 13,46 đối với mô hình AdaBoost, có nghĩa là nó cho hiệu suất tốt hơn các mô hình cây khác là Random Forest và XGBoost. Trong khi đó RMSE của mô hình mạng nơ-ron học sâu LMST là 2.85 thấp hơn rất nhiều so với các mô hình cây. Ngoài RMSE, MAE cũng là một số liệu phổ biến để đánh giá độ chính xác của các biến liên tục. Tương tự khi so sánh RMSE, đối với các mô hình cây MAE cũng cao hơn nhiều so với mô hình LSTM, có nghĩa là mô hình LSTM tốt hơn các mô hình còn lại.

## KẾT LUẬN

Trên cơ sở những nghiên cứu trong phạm vi đề tài luận án, tôi tóm tắt lại các kết quả đóng góp như sau:

Luận án đã đề xuất được phương pháp xác định nguồn phát thải ô nhiễm không khí dựa trên nguồn dữ liệu vệ tinh CALIPSO và Sentinel-5P và dữ liệu thực địa.

Luận án đã phát triển được một thuật toán ước tính độ sâu quang học sol khí từ ảnh vệ tinh độ phân giải cao Landsat 8 OLI có sự tương quan cao khi đánh giá sản phẩm AOD thu được từ thuật toán với dữ liệu từ các trạm AERONET và dữ liệu AOD của MODIS.

Luận án đã xây dựng các mô hình dự đoán  $PM_{2.5}$  cho khu vực đô thị lẫn khu công nghiệp đặc trưng là tỉnh Bắc Ninh. Dựa trên kết quả của các mô hình, luận án đã tiến hành đánh giá và đề xuất mô hình phù hợp.

### **Công việc tiếp theo**

Áp dụng phương pháp giám sát nguồn phát thải chất gây ô nhiễm không khí bằng cách kết hợp nhiều nguồn dữ liệu vệ tinh khác nhau để có thể tạo ra độ phủ về mặt thời gian và không gian.

Nghiên cứu áp dụng thuật toán đã xây dựng cho dữ liệu của các vệ tinh khác để có sự so sánh, đánh giá chính xác hơn, cũng như ứng dụng trong việc xây dựng các mô hình dự báo ô nhiễm không khí tại các khu vực khác nhau.

## **DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN**

- [**TTV-1**] Kristofer Lasko, Krishna P Vadrevu, Vinh T Tran, Evan Ellicott, Thanh T N Nguyen, Hung Q Bui and Christopher Justice, *Satellites may underestimate rice residue and associated burning emissions in Vietnam*, 2017. Environmental Research Letters, Volume 12, Number 8. DOI: <https://doi.org/10.1088/1748-9326/aa751d> (Q1 theo Scimago)
- [**TTV-2**] Vinh T Tran , Ha V Pham, Thanh TN Nguyen, Thanh X Pham, Hung Q Bui, Anh X Nguyen, Thuy T Nguyen, *Satellite Aerosol Optical Depth over Vietnam, an analysis from VIIRS and CALIOP aerosol products*, Land-Atmospheric Research Applications in South and Southeast Asia, Book Series: Springer Remote Sensing/Photogrammetry, 2018. ISBN 978-3-319-67474-2. DOI: 10.1007/978-3-319-67474-2 (Chương sách)
- [**TTV-3**] Truong X. Ngo, Ngoc T.N. Do, Hieu D.T. Phan, Vinh T. Tran, Tra T.M. Mac, Anh H. Le, Nguyet V. Do, Hung Q. Bui & Thanh T.N. Nguyen (2021) Air pollution in Vietnam during the COVID-19 social isolation, evidence of reduction in human activities, International Journal of Remote Sensing, 42:16, 6128-6154, DOI: 10.1080/01431161.2021.1934911 (Q1 theo Scimago)
- [**TTV-4**] Tran Tuan Vinh, Pham Van Ha, Nguyen Thanh Thuy and Nguyen Thi Nhat Thanh (2020), Analysis of CALIPSO satellite imagery for air pollution source identification in Hanoi, Vietnam, The 2020 12th International Conference on Knowledge and Systems Engineering (KSE), 73 – 78, ISBN (Xplore compliant): 978-1-7281-4510-5.

## **DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC CỦA TÁC GIẢ LIÊN QUAN ĐẾN LUẬN ÁN CHỜ CÔNG BỐ**

- [**TTV-5**] Vinh Tuan Tran, Truong Xuan Ngo, Thuy Thanh Nguyen, Thanh Thi Nhat Nguyen, An Algorithm for Retrieving Aerosol Optical Depth from Landsat 8 Operational Land Imager in Vietnam, Geocarto International. (Q2 theo Scimago), (Đã qua vòng 1, đang phân biện)