# *Multi-factor collaborative analysis of global environmental changes*

Speaker | Heng Chang

*School of Computing and Communications Lancaster University*

Lancaster University

# OUTLINE

# Background and significance

## Background to the selection of the topic

Because in the early stage, I read some articles about the environment and found that many studies focused on environmental issues with a single factor, this led to the idea of using multi-factor collaborative analysis to conduct more comprehensive and specific research on the environment. And due to global warming and melting glaciers some time ago, along with the latest image segmentation technology Semantic-Sam. Therefore, we chose to use the image segmentation project to automatically segment the glacier snow and obtain the changes in the snow area over the years, so as to better study the environmental factors.

## The significance of the choice of topic

Through multi-factor collaborative analysis, we can understand and evaluate environmental conditions more comprehensively. The use of Semantic-Sam in this study can improve the accuracy of information extraction and save time and effort in segmenting the required map areas. Combining the above two points, we can analyze changes in the environment more accurately.

# Introduction

1.Data collection and organization.

2. Use Semantic-Sam to segment glacier and snow images.

3.Convert the image to gray value through image processing in Python, and then count the number of pixels with different gray values to obtain the area occupied by different positions.

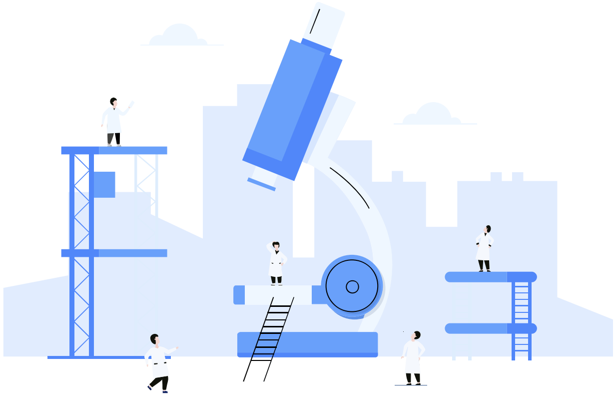
4. Calculate the Pearson correlation coefficient of the four groups of factors.

5. Compare the Pearson correlation coefficient standard to obtain the mutual synergy relationship between the four groups.

# Related Work

## Multi-factor Collaborative Analysis

After studying some environmental articles, I found that most studies are based on one or two influencing factors. To help us more fully understand the relationships between data and reveal complex patterns that cannot be discovered by single-variable analysis. In this paper, I decided to use multi-factor collaborative analysis to improve the accuracy and reliability of predictions and reduce errors caused by outliers of a single variable. By analysing the relationship between multiple variables, we can better understand the trade-offs between various factors, allowing us to make more informed decisions.



# Related Work

## Understanding of image segmentation technology

In the early days, the analysis of satellite images mainly relied on manual annotation, rule-driven methods and traditional image processing techniques. These methods often require a lot of manual intervention, are time-consuming and unsatisfactory in efficiency. More importantly, with the explosive growth of image data, the limitations of these traditional methods gradually become apparent: they are unable to process and analyse large amounts of data quickly and accurately. Through research on some papers and understanding of some technical aspects, I feel that using image segmentation technology will make our analysis of satellite images more accurate and efficient in the future, and will provide a powerful tool for in-depth understanding and prediction of environmental changes on the earth.

# Related Work

## Threshold segmentation

Threshold segmentation is a simple image segmentation method that divides the pixel values in the image into two or more categories based on thresholds. The advantage of this method is that it is simple and easy to use, but there are many shortcomings in practical application. First, it is very sensitive to noise and illumination changes, and is prone to mis-segmentation. Secondly, it cannot handle complex image structures and textures, and it is difficult to segment fine object boundaries. Finally, it requires manual threshold adjustment and is difficult to apply to large-scale image segmentation tasks.

## U-Net

U-Net is an image segmentation method based on deep learning and has good performance and application prospects. Compared with Threshold segmentation, the improvement of U-Net is that it can automatically learn the feature representation of the image without manually adjusting the threshold. In addition, U-Net adopts an encoder-decoder structure, which can extract deeper image features to more accurately segment the boundaries of objects. However, U-Net also has some shortcomings. First of all, it requires a large amount of annotated data for training, and it is prone to over-fitting problems for small sample data sets. Secondly, it still needs improvement in handling details and textures in images.

## Semantic-SAM

Semantic-SAM is a new image semantic segmentation method released in July this year, which has good application prospects. Compared with U-Net, the improvement of Semantic-SAM is that it can better handle details and textures in satellite images. In addition, Semantic-SAM can transfer zero-shot to new image distributions and tasks. Its performance in multiple tasks has been experimentally evaluated, and the results show that its zero-shot performance is impressive and generally competes with or even exceeds previous fully supervised results. In addition, Semantic-SAM's granularity richness and semantic perception are further improved compared to other image segmentation technologies.

# Related Work

## Research on Environmental Issues

In recent years, people have increasingly realized that under the severe challenges brought by global warming, various environmental problems do not exist in isolation, but are a complex, diverse, and interrelated system. Other environmental problems, such as melting glaciers, shrinking forests, and the complex relationship between population growth and carbon dioxide emissions, are all made worse by global warming. Below I will introduce the reasons for selecting these four factors.

# Related Work

## glacier snow area

Changes in glacier snow cover are one of the important indicators of climate change. As the global climate warms, the area of glaciers and snow cover decreases year by year, which not only affects global sea level rise, but may also trigger natural disasters such as floods and droughts. Therefore, selecting glacier snow cover area as one of the environmental factors will help understand the impact of climate change on the natural environment and human society.

## population

Population distribution and changes are important factors affecting the natural environment and social and economic development. Densely populated areas are often accompanied by increased urbanization, industrialization and resource consumption, which may lead to problems such as environmental pollution, ecological damage and climate change. Therefore, selecting population as one of the environmental factors helps to understand the degree and mechanism of impact of human activities on the environment.

## forest

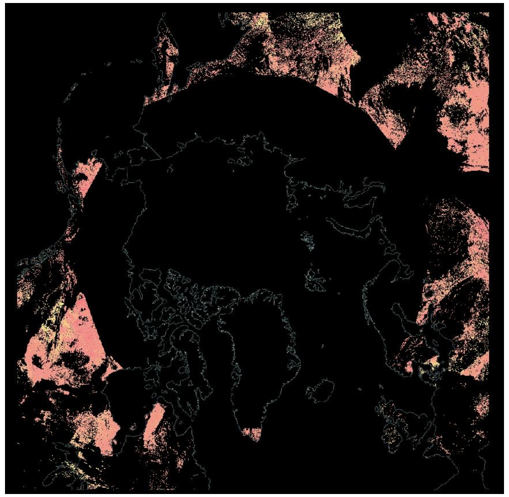
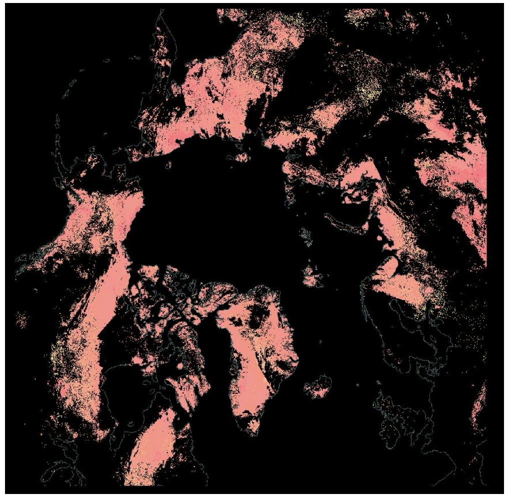
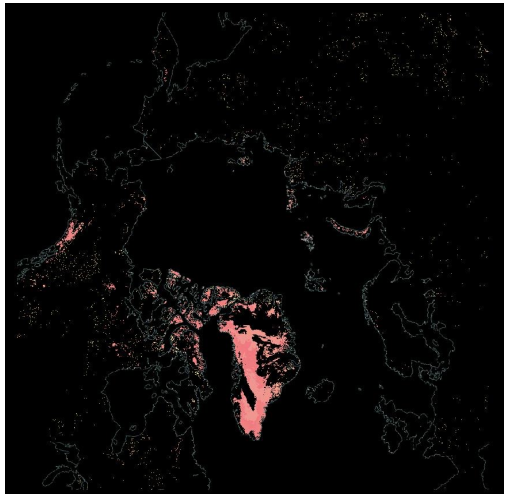
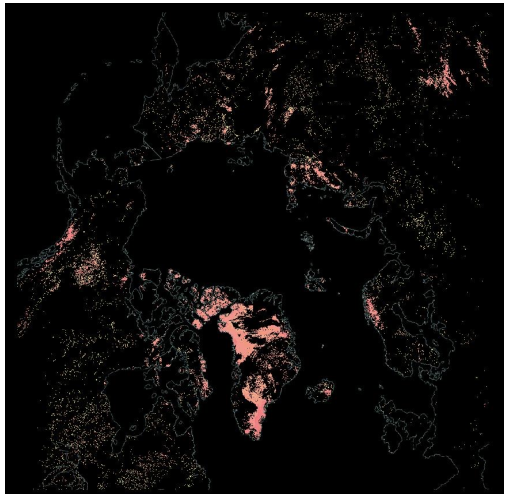
Forests are one of the most important ecosystems on earth and are of great significance to maintaining the earth's ecological balance and mitigating climate change. Forests can not only absorb large amounts of carbon dioxide, but also provide a variety of ecological services such as biodiversity protection, soil conservation, and water resource regulation. However, due to the impact of human activities and climate change, the global forest area is decreasing year by year, threatening the health and stability of the earth's ecosystem. Therefore, selecting forest as one of the environmental factors will help to understand the changing patterns and protection strategies of forest ecosystems.

## carbon dioxide

Carbon dioxide is one of the major greenhouse gases and has an important impact on global climate change. Human activities such as the burning of fossil fuels, industrial production and land use changes may lead to an increase in carbon dioxide emissions, thereby exacerbating the trend of global climate warming. Therefore, selecting carbon dioxide as one of the environmental factors helps to understand the contribution of human activities to climate change and the effectiveness of emission reduction strategies.

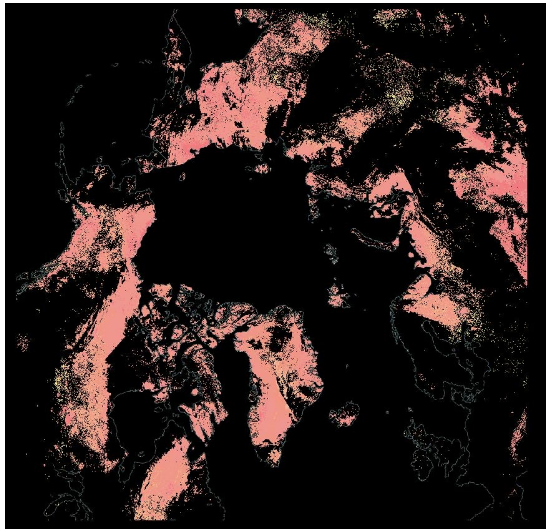
# Creating the Data Set

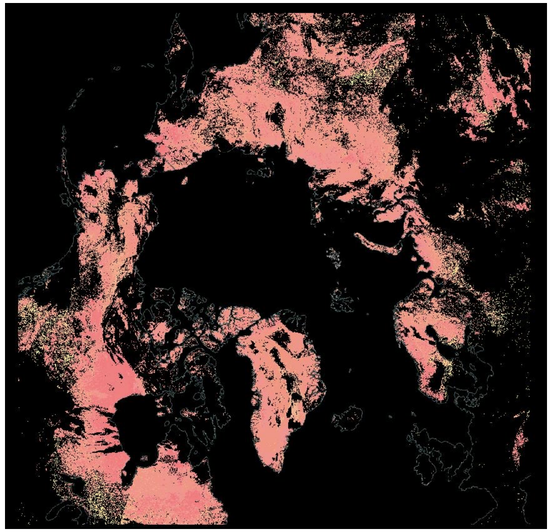
In this paper, in order to study the change of Arctic snow cover as a factor in the multi-factor collaborative analysis method, the data used are from NASA (National Aeronautics and Space Administration) and the coordinates of the Arctic region are selected as (87.5462°N, - 144.57589°E), and then intercepted images of this area with the same map scale and the same pixels to ensure the consistency of the data, and selected January, April, July and October from 2003 to 2022 as data collection time point. These four months were selected to reflect changes in snow cover during different seasons. Socioeconomic indicators related to global warming were also collected from the World Bank Data Catalog. These include population density, forest area and carbon dioxide emissions between 1990 and 2022. The following four pictures show satellite images of Arctic glacier and snow cover changes in January, April, July and October.

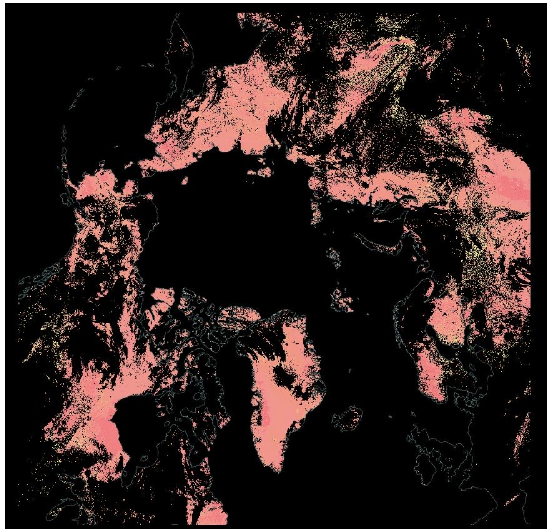





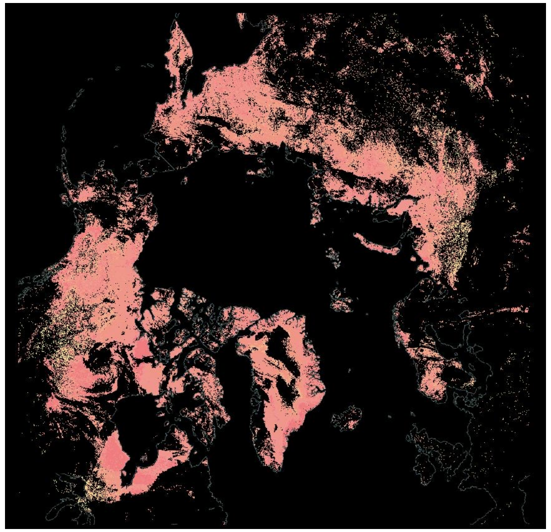
## Data Selection

After preliminary sorting of the pictures, the quality of each month's pictures is judged by sorting them in time series. Through research on the data, it was found that the image data is the most complete and clearest in April every year. In addition, the performance of ice cover and snow cover area in April is more accurate. Therefore, in order to facilitate subsequent analysis, this article selects pictures in April of each year as research data. In the data provided by World Bank Data Catalog, the data needs to be cleaned. This includes deduplicating data, removing erroneous data, and filling in missing data, among other things. The following four pictures show glacier snow cover in April in different years.



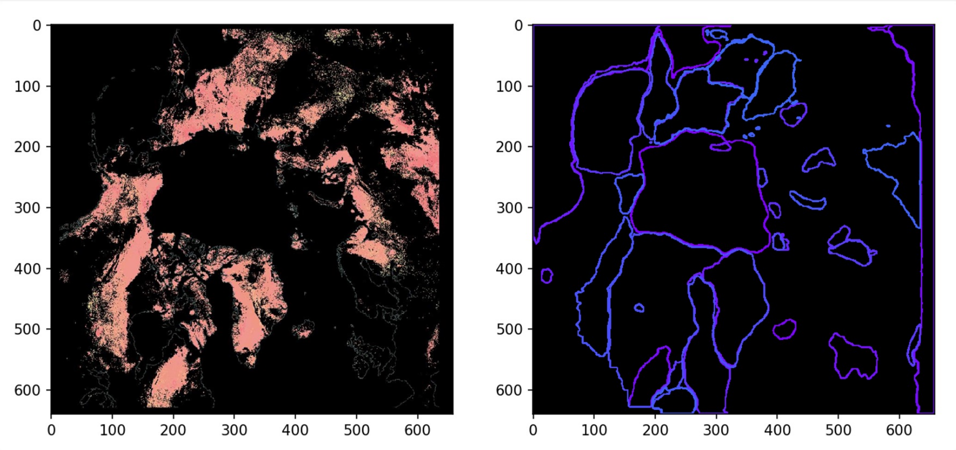


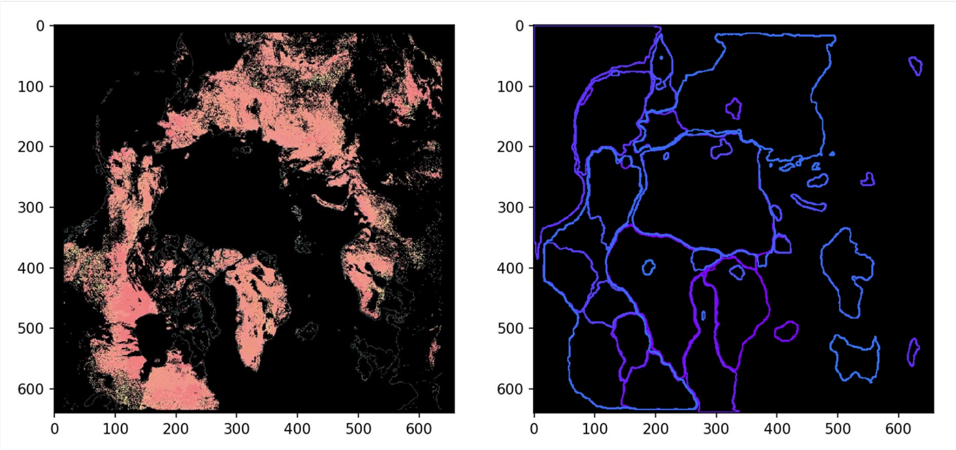




# Image segmentation of Arctic glacier snow images using the Semantic-Sam model（Extract data method）

I chose to use the Semantic-SAM model as the image segmentation tool for this study. This model was trained on more than 11 million images and 1 billion masks, which makes it highly accurate and reliable on image segmentation tasks. This is particularly important as we deal with the complexity and diversity of ice and snow images in the Arctic. The high scalability of this model also determines its excellent robustness. It can effectively handle various types of high-quality images and reduce the interference of noise and other interference factors on the image. The SAM model has the capability of zero-shot learning, which means that the model can effectively process new tasks when there is no specific training data. This feature gives us more flexibility and makes it easier to use the model in various Arctic glacier and snow images. The figure below lists the effects of image segmentation in two years.

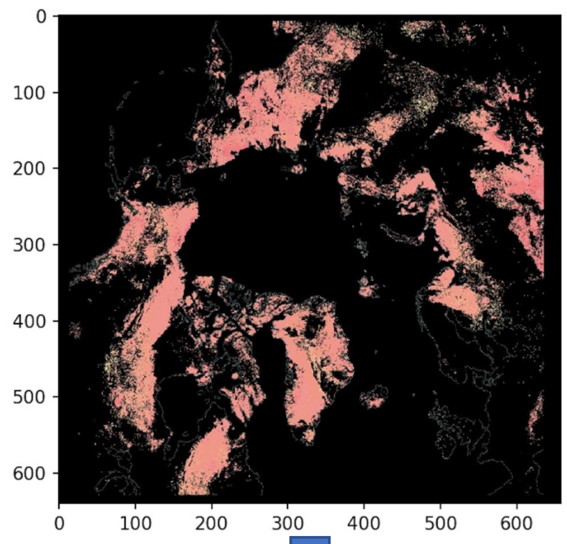


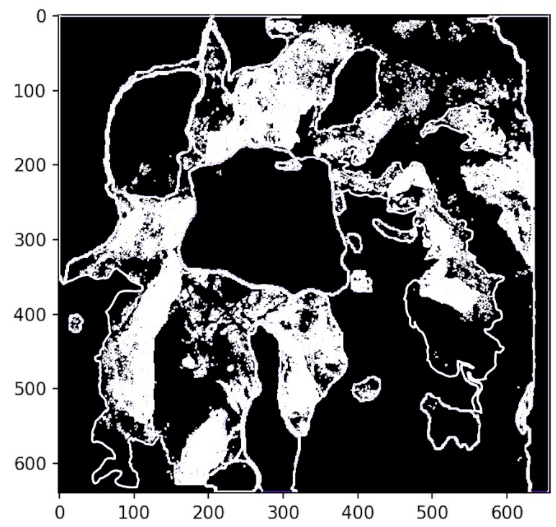



# Method

## Use pixels to calculate the snow ratio of the image ( Statistical snow cover data )

In order to obtain the snow-covered area in the image, this study uses the Python language for data processing. After using the Semantic-Sam model to segment the glacier snow-covered image, the snow-covered area is manually filled and filled with a single color for better accuracy. Good distinction between snow and background areas. We use Python language to preprocess the filled snow images. Mainly relies on the PIL (Pillow) library for image reading and manipulation. By using the image function to open the preprocessed image, we convert it into a grayscale image. By calculating the grayscale distribution of the image, we find that the pixel values are concentrated around two peaks. After analysis, it was confirmed that one of the peaks clearly corresponds to the snow-covered area, while the other peak corresponds to the non-snow-covered area. In order to determine an appropriate threshold, several representative images were selected for experimental reclassification. After trying several different thresholds, it was found that a pixel value of 50 can effectively distinguish snow-covered and non-snow-covered areas. Pixels with a grayscale value greater than or equal to 50 are then labeled as snow (value 1), while pixels less than 50 are labeled background (value 0). Use the SUM function to count the number of pixels with a value of 1 (snow) and a value of 0 (background). Then the value of the proportion of glacier snow can be obtained. The picture on the right shows what the segmented image looks like after manual filling and grayscale value conversion.



# Method

## Pearson correlation coefficient of four factors ( Main research methods )

In order to better reflect the correlation of these four factors, Obtain the correlation coefficients between the four factors by using spss software we use the Pearson coefficient to judge their correlation.

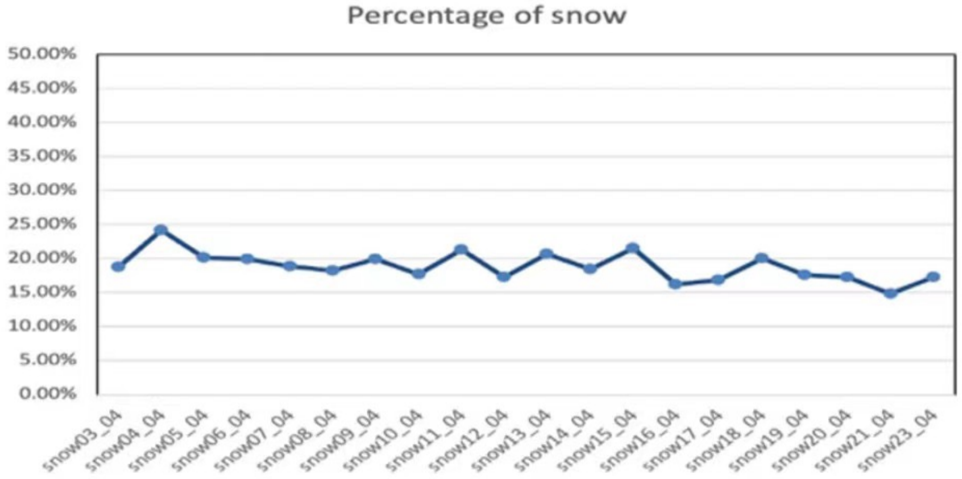
**TABLE 1 Pearson Correlation Coefficients**

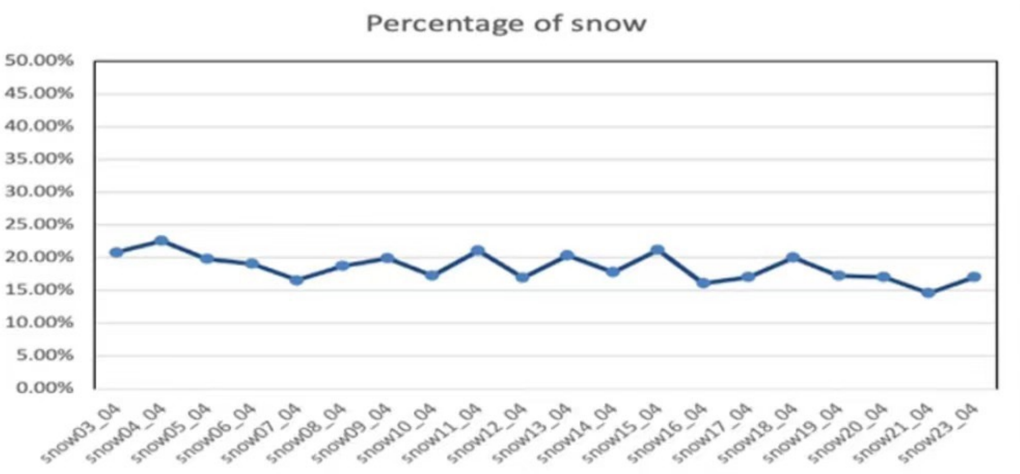
| | CO2 Emission | Forest Cover | Population | Snow Cover |
|---|---|---|---|---|
| CO2 Emission | 1 | | | |
| Forest Cover | -.939** | 1 | | |
| Population | .978** | -.980** | 1 | |
| Snow Cover | -.599** | .718** | -.641** | 1 |

* p<0.05 ** p<0.01

# Results

## The original image and the segmented image are converted into grayscale values to obtain the proportion of glacier snow for comparison. ( Verify accuracy )




| DATE | snow03-04 | snow04-04 | snow05-04 | snow06-04 | snow07-04 | snow08-04 | snow09-04 | snow10-04 | snow11-04 | snow12-04 |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of snow rasters | 125926 | 162674 | 135533 | 134068 | 127031 | 122759 | 133928 | 119193 | 143355 | 115830 |
| Number of black rasters | 544835 | 509726 | 536867 | 538332 | 545369 | 549641 | 538472 | 553207 | 529045 | 556570 |
| Percentage of snow | 18.77% | 24. 19% | 20. 16% | 19. 94% | 18. 89% | 18. 26% | 19. 92% | 17. 73% | 21. 32% | 17. 23% |
| DATE | snow13-04 | snow14-04 | snow15-04 | snow16-04 | snow17-04 | snow18-04 | snow19-04 | snow20-04 | snow21-04 | snow23-04 |
| Number of snow rasters | 139005 | 123950 | 144476 | 109012 | 113416 | 135075 | 118154 | 115993 | 99657 | 115882 |
| Number of black rasters | 533395 | 548450 | 527924 | 563388 | 558984 | 537325 | 554246 | 553947 | 571104 | 556518 |
| Percentage of snow | 20.67% | 18. 43% | 21. 49% | 16. 21% | 558984 | 20. 09% | 17. 57% | 17. 31% | 14. 86% | 17. 23% |

| DATE | snow03-04 | snow04-04 | snow05-04 | snow06-04 | snow07-04 | snow08-04 | snow09-04 | snow10-04 | snow11-04 | snow12-04 |
|---|---|---|---|---|---|---|---|---|---|---|
| Number of snow rasters | 133282 | 156753 | 133435 | 132087 | 110631 | 127867 | 133928 | 115322 | 142212 | 114326 |
| Number of black rasters | 506335 | 536797 | 538907 | 558432 | 556369 | 553212 | 538472 | 553207 | 530241 | 558563 |
| Percentage of snow | 20.83% | 22. 60% | 19. 84% | 19. 12% | 16. 58% | 18. 77% | 19. 91% | 17. 25% | 21. 14% | 16. 99% |
| DATE | snow13-04 | snow14-04 | snow15-04 | snow16-04 | snow17-04 | snow18-04 | snow19-04 | snow20-04 | snow21-04 | snow23-04 |
| Number of snow rasters | 136743 | 122354 | 142675 | 103324 | 112432 | 133065 | 115435 | 113778 | 97633 | 114648 |
| Number of black rasters | 534362 | 556430 | 536455 | 536455 | 546344 | 530461 | 554126 | 554043 | 571432 | 556518 |
| Percentage of snow | 20.37% | 17. 81% | 21. 25% | 16. 14% | 17. 06% | 20. 05% | 17. 24% | 17. 03% | 14. 59% | 17. 08% |

The image on the left represents the proportion of glacier snow area in the segmented image. The image on the right represents the proportion of glacier snow in the original image. It can be seen that the percentage difference between the two sets of data is very small. From this, it can be concluded that Semantic-Sam can be used to Make satellite image analysis more efficient.

Lancaster University

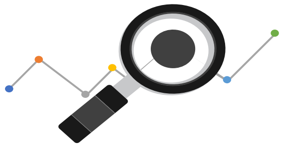
The University of Lancaster

# Results

## Verify the accuracy and precision of the Semantic-Sam model

In order to verify that the image segmentation effect of semantic-Sam is indeed better than that of general image segmentation, I conducted a comparative experiment using the U-Net model. I used the same data set to first train the two sets of models, using (accuracy and precision) to evaluate two segmentation models.

| MODEL | Semantic- Sam | U-Net |
|---|---|---|
| Accuracy | 85.2 | 72.3± 0.2 |
| Precision | 90.3 | 77.2±1.5 |



As can be seen from the above figure, the Semantic-Sam model has a better segmentation effect than the U-Net model. By comparing the accuracy and precision of the two sets of models, we can clearly observe that the Semantic-Sam model has a better segmentation effect than U-Net. The segmentation effect is better, partly benefiting from its powerful zero-sample learning ability. Also benefit from its higher performance and more flexible basic framework. This may also explain why using Semantic-Sam to segment satellite images will indeed bring convenience to future researchers. It also provides a more powerful tool for future environmental research.

# Results

## correlation analysis

TABLE 1 Pearson Correlation Coefficients

| | CO2 Emission | Forest Cover | Population | Snow Cover |
|---|---|---|---|---|
| CO2 Emission | 1 | | | |
| Forest Cover | -.939** | 1 | | |
| Population | .978** | -.980** | 1 | |
| Snow Cover | -.599** | .718** | -.641** | 1 |

* p<0.05 ** p<0.01

| Correlation Coefficient | Evaluation |
|---|---|
| 1 | Perfect correlation, positive linear relationship |
| 0.80-0.99 | Strong correlation, positive linear relationship |
| 0.50-0.79 | Moderate correlation |
| 0.30-0.49 | Weak correlation |
| 0.10-0.29 | Very weak correlation or no linear relationship |
| 0 | Zero correlation, no correlation |

The Pearson correlation coefficient was obtained by using spss software, and then comparing the obtained Pearson correlation coefficient with the Pearson correlation coefficient table, it can be concluded that there is a strong negative correlation between forest coverage area and carbon dioxide emissions. There is a strong positive correlation between population and carbon dioxide emissions. Carbon dioxide emissions show a moderate negative correlation with snow cover area. There is a strong correlation between population and forest cover. There is a clear positive correlation between forest cover area and snow cover area. There is a moderate negative correlation between population size and snow cover. Generally speaking, there is a strong correlation between the four factors, which further validates the theme of this study. For environmental research, we need to use multi-factor collaborative analysis, which can make our research or prediction more accurate.

# Discussion

## Data on four factors

The sample size of the data from the glacier and snow region in this study is very limited, which will have an impact on how image segmentation is used, even though most of the data can still provide important information for models and algorithms. The Semantic-Sam model will also have an impact on the calculation of the correlation coefficient. The official NASA website has limited content as it does not contain photos from the same time and place in the past. In future research, I will adopt more systematic and rigorous data collection methods to ensure that the data is more balanced. There are some missing and outliers in the processing data section, and the data is organized. This part will also have a certain impact on the accuracy of the research results. Due to time constraints, the data in this article mainly comes from NASA and World Bank Data CatLog, which will also limit the results to a certain extent.

## Image segmentation and Pearson correlation

In the intercepted satellite image data, some images contain noise and blur, which will lead to a decrease in the accuracy of model segmentation. In future research, the preprocessing of data needs to be further optimized to improve the model's robustness to images. Of course, there are indeed some errors in accuracy when segmenting images using semantic-sam. In future research, we need to find some methods to use more accurate models for image segmentation. More complex image processing models such as Semantic-Sam lack intuitive explanations, and more models that can be automatically interpreted can be developed in the future. Measuring these four factors through correlation coefficients may lead to over-interpretation or misinterpretation of the results, and the suitability of the selected measures needs to be further investigated and validated in future studies.

# Conclusions

## Conclusions

In this study, we employ an advanced Semantic-Sam model to accomplish the task of segmenting Arctic glacier snow cover from satellite images. Semantic-Sam demonstrates its suitability for future research involving satellite image analysis through its superior segmentation capabilities. We confirmed the accuracy of its segmentation by using pixel values to calculate the proportion of snow, and then compared with U-Net, a traditional image segmentation model, Semantic-Sam does provide faster and more accurate results on satellite maps. Accurate segmentation results. This discovery provides a powerful tool for future studies of satellite imagery, allowing future studies to be conducted more accurately and more easily.

In addition, this paper demonstrates the synergistic relationship between multiple factors, including the proportion of snow area in satellite images, as well as population density, forest area and carbon dioxide emissions. The interrelationships between these four factors were verified using the Pearson correlation coefficient. A relatively strong correlation between them was revealed, confirming our initial hypothesis that environmental change is influenced by a range of factors rather than just one.

By verifying the impact of such multiple factors on environmental changes, more comprehensive information can be provided for future research environments. Specifically, understanding how these factors are interrelated can provide some useful information for protecting the environment and formulating future policies to protect the environment.

Lancaster University

# Thanks for your attention