A "Warm" or "Cold" Early Mars: Evidence From Valley Networks

Xuezhi Cang
xuezhicang@gmail.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Part of the Geographic Information Sciences Commons, Other Earth Sciences Commons, and the Spatial Science Commons

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/6893
ABSTRACT

A “WARM” OR “COLD” EARLY MARS: EVIDENCE FROM VALLEY NETWORKS

Xuezhi Cang, Ph.D.
Department of Geographic and Atmospheric Sciences
Northern Illinois University, 2021
Wei Luo, Director

Although Mars is cold and dry today, many lines of evidence suggest that ancient Mars had large amount of water and surface fluvial processes. However, climate modelers have encountered difficulties in modeling such early warm and wet conditions with an above freezing temperature, mainly due to the Martian orbit being further away from the Sun and the faintness of the young Sun. The main purpose of this dissertation is to test the ancient Martian climate hypotheses by investigating the spatial pattern of valley network (VN) properties that can survive post-formational modification (i.e., the robust characteristics) and by analogizing the Earth streams and Martian VNs. In the dissertation, I developed a spatial analysis method to assess the associations between properties of streams/VNs and their corresponding environmental factors. I then utilized this tool to analyze and analogize one VNs’ robust characteristic (junction angle) and its corresponding environmental factors and estimated the early Mars climatic conditions. I also calculated Martian watersheds’ maturity by comparing their energy dissipation with those of Optimal Channel Network (OCN) to investigate whether the “warm” duration on Mars is long enough. Our results suggest that the ancient Mars was at least “episodically warm” and had surface runoff long
enough to carve the Martian surface in some regions and the climatic conditions of the early Mars was similar to those of the arid/sub-arid regions on Earth.
A “WARM” OR “COLD” EARLY MARS:
EVIDENCE FROM VALLEY NETWORKS

BY
XUEZHI CANG
© 2021 Xuezhi Cang

A DISSERTATION SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOGRAPHIC AND ATMOSPHERIC SCIENCES

Dissertation Director:
Wei Luo
ACKNOWLEDGEMENTS

First, I am deeply indebted to my wife, Jingzhi Zhang. Her patience and willingness to take care of me so that I could focus on this work has meant more than she will ever know. Her invaluable encouragement and support during the journey were important. Thanks for always being there for me.

I would like to express my deepest appreciation to my advisor Dr. Wei Luo. He helped me throughout the entire program and went above and beyond to ensure that I complete this dissertation. He is not only my advisor but also a mentor to me. His valuable advice and suggestions on the research design, procedures and publication always put me in the right direction. I am very thankful to him for being patient with me. Without his guidance and persistent help this dissertation would not have been possible.

I would like to thank my parents, who have been very supportive and eager to see me earning this degree all these years. It would not have been possible without their love and encouragement.

I am also thankful to my committee members Dr. Andrew Krmenec, Dr. Thomas Pingel and Dr. Paul Stoddard for their input and feedback, which have been very helpful in shaping this research. I am very grateful for their guidance and suggestions.

I am very thankful to Graduate Coordinators and Department Chairs. Their support and encouragement have always motivated me. I am very grateful for the assistantship and the opportunity to teach the courses in the department, which have added greatly to my professional experiences. I also acknowledge the Department of Geographic and Atmospheric Sciences and the Department of Computer Science, including the faculty, staff, and other students. Without their help, I could not have done this dissertation.
I greatly appreciate the comments and suggestions from anonymous reviewers which helped improve the quality of Chapters 2 and 3.

Finally, I would like to thank everyone who has helped me directly or indirectly to complete this research. Thank you!
DEDICATION

To the Moon, Mars and the sea of stars.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>ix</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
</tbody>
</table>

## Chapter

1 INTRODUCTION AND BACKGROUND                      | 1   |

1.1 Water on the Mars                             | 1   |

1.2 Climatic Scenario of VNs’ Formation           | 3   |

1.3 Martian Valley Networks Analog Analysis       | 5   |

1.4 Research Questions                            | 6   |

1.5 Summary                                       | 8   |

2 SPATIAL ASSOCIATION DETECTOR.                   | 10  |

2.1 Abstract                                       | 10  |

2.2 Introduction                                   | 11  |

2.2.1 Introduction of Geographical Detector       | 11  |

2.2.2 Analysis of Geo-detector                     | 12  |

2.3 SPatial Association DEtector (SPADE) Considering Spatial Variance, Multi-level Discretization, and Information Loss | 17  |

2.3.1 Spatial Variance                             | 17  |

2.3.2 Multi-level Discretization and Information Loss | 19  |

2.3.3 Test of Significance                         | 22  |

2.4 Simulation Test of PSMD                        | 24  |

2.4.1 Simulation Scheme                            | 24  |
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4.2 Estimated Spatial Data Association Comparison by Simulations</td>
<td>25</td>
</tr>
<tr>
<td>2.5 Application to U.S. Surface Dissection Density Data</td>
<td>26</td>
</tr>
<tr>
<td>2.6 Discussion</td>
<td>30</td>
</tr>
<tr>
<td>2.6.1 Selection of Model Parameters</td>
<td>30</td>
</tr>
<tr>
<td>2.6.2 Assumption of the Probability Distributions</td>
<td>32</td>
</tr>
<tr>
<td>2.6.3 Selection of Significant Test</td>
<td>38</td>
</tr>
<tr>
<td>2.7 Conclusion</td>
<td>39</td>
</tr>
<tr>
<td>3 NOACHIAN CLIMATIC CONDITIONS ON MARS INFERRED FROM VALLEY NETWORK</td>
<td>41</td>
</tr>
<tr>
<td>JUNCTION ANGLES</td>
<td></td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>41</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>42</td>
</tr>
<tr>
<td>3.3 Dataset</td>
<td>46</td>
</tr>
<tr>
<td>3.4 Method and Procedure</td>
<td>47</td>
</tr>
<tr>
<td>3.4.1 Junction Angle Extraction</td>
<td>47</td>
</tr>
<tr>
<td>3.4.2 Spatial Association Measures by SPADE</td>
<td>50</td>
</tr>
<tr>
<td>3.4.3 Comparative Analysis Procedure</td>
<td>50</td>
</tr>
<tr>
<td>3.4.3.1 Terrestrial Data Analysis</td>
<td>50</td>
</tr>
<tr>
<td>3.4.3.2 Terrestrial Data Modeling</td>
<td>51</td>
</tr>
<tr>
<td>3.4.3.3 Inferring Mars Climatic Conditions</td>
<td>52</td>
</tr>
<tr>
<td>3.4.3.4 Estimating the Duration of “Warm” Mars</td>
<td>53</td>
</tr>
<tr>
<td>3.5 Results</td>
<td>55</td>
</tr>
<tr>
<td>3.5.1 Terrestrial Junction Angle Analysis</td>
<td>55</td>
</tr>
<tr>
<td>3.5.2 Global and Local Mars VNs Analysis</td>
<td>57</td>
</tr>
<tr>
<td>3.5.3 Estimated Climatic Conditions and Duration of “Warm” Period</td>
<td>60</td>
</tr>
<tr>
<td>Chapter</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>3.6</td>
<td>Discussion and Conclusion</td>
</tr>
<tr>
<td>3.6.1</td>
<td>Terrestrial Model Selection</td>
</tr>
<tr>
<td>3.6.2</td>
<td>Estimated Mars Climatic Conditions</td>
</tr>
<tr>
<td>3.6.3</td>
<td>Estimated Duration of “Warm” Period</td>
</tr>
<tr>
<td>3.6.4</td>
<td>Additional Uncertainties from Channel Bed Conditions and Orbital Obliquity</td>
</tr>
<tr>
<td>3.6.5</td>
<td>Conclusion</td>
</tr>
<tr>
<td>4</td>
<td>THE MARS WATERSHEDS’ MATURITY MEASURED BY OPTIMAL CHANNEL NETWORKS SIMULATION</td>
</tr>
<tr>
<td>4.1</td>
<td>Abstract</td>
</tr>
<tr>
<td>4.2</td>
<td>Introduction</td>
</tr>
<tr>
<td>4.3</td>
<td>Dataset</td>
</tr>
<tr>
<td>4.4</td>
<td>Method and Procedure</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Hydrologic Analysis and Watershed Delineation</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Optimal Channel Network and Watershed Maturity</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Spatial Association Measured by SPADE Detector</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Comparative Analysis Procedure</td>
</tr>
<tr>
<td>4.4.4.1</td>
<td>Maturity Analysis of Terrestrial Watersheds</td>
</tr>
<tr>
<td>4.4.4.2</td>
<td>Martian Watersheds Maturity Analysis</td>
</tr>
<tr>
<td>4.5</td>
<td>Results</td>
</tr>
<tr>
<td>4.5.1</td>
<td>Terrestrial Watershed Maturity Analysis</td>
</tr>
<tr>
<td>4.5.2</td>
<td>Martian Watershed Maturity Analysis</td>
</tr>
<tr>
<td>4.6</td>
<td>Discussion and Conclusion</td>
</tr>
<tr>
<td>4.6.1</td>
<td>Earth Watersheds Maturity</td>
</tr>
<tr>
<td>4.6.2</td>
<td>Outliers of Martian VN’s Maturity</td>
</tr>
</tbody>
</table>
### Chapter 4.6.3 Martian VN's Maturity and Early Martian Climate

- Page 99

### Chapter 4.6.4 Conclusion

- Page 100

### Chapter 5 SUMMARY AND CONCLUSIONS

- Page 102
  
  5.1 Overview

- Page 102
  
  5.2 Future Work and Research Directions

- Page 103

### Chapter 6 RESPONSE TO COMMITTEE’S COMMENTS

- Page 106
  
  6.1 Overview

- Page 106
  
  6.2 The Uncertainties of Junction Angle Influenced by the Map Projection

- Page 106
  
  6.3 The Uncertainties of Models

- Page 108
  
  6.4 Philosophical Thinking about Analogizing in Planetary Research

- Page 109
  
  6.5 Immature Watershed or Outliers

- Page 110

### WORKS CITED

- Page 112
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Controlling factors.</td>
</tr>
<tr>
<td>2</td>
<td>PSMD results in Regions 3, 6, 7 and 8</td>
</tr>
<tr>
<td>3</td>
<td>PSMD of continuous variables under different discretization levels in Region 3</td>
</tr>
<tr>
<td>4</td>
<td>PSMDs under different distance decay methods</td>
</tr>
<tr>
<td>5</td>
<td>Basic statistical result of Mars junction angles</td>
</tr>
<tr>
<td>6</td>
<td>Associations using SPADE (larger value means higher association)</td>
</tr>
<tr>
<td>7</td>
<td>Model performance on terrestrial data.</td>
</tr>
<tr>
<td>8</td>
<td>Estimated duration of “warm” Mars (unit: million years)</td>
</tr>
<tr>
<td>9</td>
<td>Spatial analysis of landform characteristics and estimated $AI$.</td>
</tr>
<tr>
<td>10</td>
<td>Average and SD of terrestrial watershed maturity</td>
</tr>
<tr>
<td>11</td>
<td>Associations between watersheds’ maturity and environmental factors at the conterminous U.S. scale</td>
</tr>
<tr>
<td>12</td>
<td>Associations between watersheds’ maturity and environmental factors in Regions</td>
</tr>
<tr>
<td>13</td>
<td>Average, median, and SD of Martian watershed maturity</td>
</tr>
<tr>
<td>14</td>
<td>Average, median, and SD of Martian VNs maturity</td>
</tr>
<tr>
<td>15</td>
<td>Comparisons of junction angles under different map projections</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Method of the geographical detector</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>PD values of elevation under different discretization level (from 4 to 22)</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>Discretizing continuous independent variable using different discretization levels (zones)</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>Comparison of PD, compensated PD (CPD), PSD and compensated PSD (CPSD), under different shuffling rates</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>Comparison of estimated spatial association and critical value under different shuffling rates. X-axis is the discretization level and Y-axis is estimated spatial association between two variables</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>Dissection density in the US</td>
<td>30</td>
</tr>
<tr>
<td>7</td>
<td>Dissection density and its controlling factors in Region 3</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>Comparison spatial estimators form different distributed variables</td>
<td>34</td>
</tr>
<tr>
<td>9</td>
<td>QQ-plot of null hypothesis under varied distributions</td>
<td>35</td>
</tr>
<tr>
<td>10</td>
<td>VN junction angle calculation based on VN topology</td>
<td>48</td>
</tr>
<tr>
<td>11</td>
<td>Frequency distribution of junction angles</td>
<td>58</td>
</tr>
<tr>
<td>12</td>
<td>Spatial distribution of junction angle of Mars</td>
<td>59</td>
</tr>
<tr>
<td>13</td>
<td>Box and whisker chart of the estimated $AI$</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>OCN under different $\gamma$ values</td>
<td>84</td>
</tr>
<tr>
<td>15</td>
<td>Histograms of terrestrial watersheds’ maturity</td>
<td>90</td>
</tr>
<tr>
<td>16</td>
<td>Spatial distribution of terrestrial watersheds’ maturity</td>
<td>91</td>
</tr>
<tr>
<td>17</td>
<td>Spatial distribution of environmental factors</td>
<td>92</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>18</td>
<td>Histogram of Martian watersheds’ maturity.</td>
<td>95</td>
</tr>
<tr>
<td>19</td>
<td>Martian watersheds’ maturity.</td>
<td>96</td>
</tr>
<tr>
<td>20</td>
<td>Histogram of Martian VN’s maturity.</td>
<td>97</td>
</tr>
<tr>
<td>21</td>
<td>Spatial distribution of Martian VN’s maturity.</td>
<td>98</td>
</tr>
</tbody>
</table>
1.1 Water on the Mars

Mars is one of the nearest planets to Earth and has some similarities. The current mean obliquity is very close to Earth’s. The rotational period is also similar, the length of a day on Mars is 24.6 hours. However, unlike Earth, Mars is inhospitable to humans now. Mars is a dry and cold planet. The average temperature of Mars is much lower than the Earth’s average temperature. Moreover, the average Martian pressure, which is 6 to 7 millibars, is significantly smaller than that of Earth at sea level, which is about 1,013 millibars. Based on the water phase diagram, liquid water could not exist stably on its surface under such low temperature and low pressure conditions.

Although Mars is cold and dry today, many lines of evidence suggest that ancient Mars had large amount of water and surface fluvial processes. Based on the degree of crater degradation and crater counting, the main Mars geologic history has been divided into three stages (Carr and Head, 2010): Noachian period (4.1 to 3.7 Gya), Hesperian period (3.7 to 3.0 Gya) and Amazonian period (3.0 Gya to present) (Carr & Head III, 2010).

In the Noachian period, the rates of cratering, erosion, and surface weathering is much higher than the Hesperian period and the Amazonian period. A Mars ocean hypothesis stated that the Vastitas Borealis, a northern basin covering nearly a third of the surface of Mars, was filled by water early in the Noachian period. The ocean hypothesis (Brandenburg, 1987) is supported by many lines of evidence from different fields of research. The analogous
landform of tsunami (Rodriguez et al., 2016) showed that the northern basin of Mars once had an ocean in the Noachian period. The deltas (Hynek et al., 2010) are located at approximately the same geopotential level, consistent with sea level of an ocean in the Noachian period. The change of deuterium and hydrogen ratio (D/H) (Villanueva et al., 2015) suggests that Mars had abundant water three billion years ago. The coastline landform (Malin & Edgett, 2000) of Mars also supports the ocean hypothesis. Though this evidence is challenged by the difference of elevation of the coastline (Carr & Head III, 2003), the deformation caused by true polar wander associated with the rising of the Tharsis Montes (Perron et al., 2007) explained the difference between elevations of coastline. Also, the amount of crater degradation suggested that water activities, such as surface runoff erosion and weathering, occurred on the Martian surface several billion years ago.

In the Noachian period, most valley networks were formed by surface runoff erosion which was caused by at least episodic precipitation. It indicates that the early Mars not only had large amount of water, but its climatic conditions were “warm” and “wet” enough to allow liquid water to flow on the surface. The causes and the durations of the “warm” and “wet” climatic conditions are unclear. The possible explanations are global warming from impacts and volcanism (Halevy & Head III, 2014), global warming from greenhouse gases (Ramirez et al., 2014), or a combination effect (Wordsworth et al., 2021; Wordsworth, 2016).

After the Noachian period, because of the geological inactivity, the fluvial landforms formed in the Noachian period were preserved. The surface runoff erosion rate declined sharply in the Hesperian and Amazonian periods. In the Hesperian and Amazonian periods, most of the water and atmosphere, especially the hydrogen atoms, on Mars has evaporated and escaped into space (Carr & Wanke, 1992). Mars became a dry and cold planet like today. Although the large outflow events, as evidenced by large outflow channels, occurred in the Hesperian period, the duration of the outflow is short (Carr & Head III, 2010) and the source of outflow was the Martian aquifer and not precipitation (rainfall or melting snow).
For this reason, the outflow is not the evidence pointing to the “warm” and “wet” climatic conditions.

1.2 Climatic Scenario of VNs’ Formation

The existence of liquid water on early Mars in the Noachian period is supported by evidence in geochemistry and geomorphology. The wide distribution of aqueous alterations observed by satellite remote sensing and rovers indicate that large quantities of liquid water covered a lot of regions (Carter et al., 2013). The degradation of craters also showed that Mars had a warm period (Craddock & Howard, 2002; Hynek & Phillips, 2003). The valley networks (VNs), which were mainly formed in the Noachian period, are the important evidence that the early Mars had liquid water flowing on its surface. The remote sensing images acquired by Mariner 9 Mars orbiter first revealed the presence of dendritic valley networks (VNs) on the heavily cratered Noachian highlands (Masursky, 1973), suggesting the possibility of an early warm and wet climate with liquid water as erosional agent.

The widely distributed VNs on Mars, which were mapped based on digital elevation model (DEM) data generated from Mars Orbiter Laser Altimeter (MOLA) and the remote sensing images obtained from orbit (Di Achille & Hynek, 2010; Fassett & Head III, 2008; Luo & Stepinski, 2009), suggested widely distributed fluvial processes on early Mars. Although controversies still exist regarding the origin of the valley networks (e.g., by volcanic lava flow or by floodwater), most valley networks are considered to be related to the fluvial processes because their dendritic patterns are consistent with the basic trend of topography similar to what happens on Earth (Craddock & Howard, 2002).

Although the widely distributed VNs are the evidence for flowing liquid water on early Mars, the climate models support the “cold” early Mars theory (Wordsworth et al., 2015;...
Wordsworth, 2016). Climate models have difficulties in producing the early warm and wet conditions for the red planet, mainly because of the Martian orbit and the faintness of the young Sun (Wordsworth, 2016), thus they do not favor active global water cycling on early Mars.

However, large uncertainties exist in the initial conditions and evolution processes of climate models of early Mars. One of the most important initial conditions for early Mars climate model is its atmosphere composition (Wordsworth et al., 2015). A great number of hypotheses about the atmospheric composition have been proposed and rejected. For example, Wordsworth et al. (2015) think that the H$_2$-N$_2$ warming mechanism could not support a long “warm” period because H$_2$-N$_2$ warming mechanism needs a low H$_2$ escape rates and high combined H$_2$ and CO$_2$ outgassing rate. Several hypotheses are still under investigated (Ramirez et al., 2014; Urata & Toon, 2013; Wordsworth et al., 2015). Because of the cost of the Mars exploration missions and limitations of space technology, it is impossible to obtain all the data needed to address the questions. Thus, the uncertainty associated with the climate model will continue to exist for the foreseeable future.

Although the differences still exist, the consensus between supporters of “Warm” Mars and “Cold” Mars is that abundant liquid water flowed on Mars in the Noachian period (Ramirez & Craddock, 2018; Wordsworth et al., 2018). In this dissertation, I aim to address the warm and cold scenarios by investigating the duration, intensity, and evolution of fluvial processes on Mars. The VNs’ characteristics and their spatial distribution are the fingerprints of the past fluvial processes, so VNs could be utilized to test the early Mars climatic conditions.
1.3 Martian Valley Networks Analog Analysis

Comparing measurable spatial patterns of Mars with those on Earth, where we have better physical access and deeper understanding of the processes, can often offer significant insights into Martian geological processes.

Mars analog sites are places on Earth with assumed past environmental conditions of Mars. The ample data acquired from previous research projects on these Mars analog sites on Earth can help researchers explore and interpret the similar geographical pattern on Mars. The evidence of widely distributed VNs on Mars supports the fluvial processes on early Mars, but cannot accurately estimate the duration and intensity of the hydrological cycle of early Mars. To study the hydrological cycle of early Mars and the evolution of VNs, some analog sites on Earth were selected in previous research projects, such as the Great Basin area, NV (Matsubara et al., 2011; Matsubara et al., 2013) and the Usuktuk River, AK (Matsubara et al., 2015). The results showed that the ratio of the net lake evaporation rate to runoff depth on Mars is comparable to that of the Great Basin region in the western United States during the Last Glacial Maximum (LGM). This suggests that liquid water on early Mars existed under climatic conditions similar to those in the Great Basin region during the LGM. The uncertainty of this method is that whether the climate of small analog sites can represent the early Mars global climate.

Some globally distributed landform properties were also examined by previous research projects. The longitudinal profile of VNs and watersheds area were investigated (Galofre, Bahia, et al., 2020; Penido et al., 2013; Som et al., 2009). Som et al. (2009) and Penido et al. (2013) used the same properties, which are the relation between longitudinal profile and watersheds’ size, at different scales. Som et al. (2009) showed that the point source discharges were the water source for VNs. In contrast, Penido et al. (2013) revealed that
the precipitation is the water source. The possible explanation for the opposite results is that longitudinal profile of watershed may have been modified by the impact cratering and post-formation processes. For this reason, more robust and widely distributed landform properties that are not influenced by post-formatoinal processes should be selected in the research to obtain the reliable estimates in inferring the early Mars climatic conditions.

1.4 Research Questions

The discussions above showed that the early Mars environmental conditions for VN formation are still not clear. The main purpose of this research is to address this question. With the development of the computation techniques, more detailed terrain information over larger areas can be automatically extracted from Mars and Earth. The extracted terrain information of Earth and Mars can also be examined by taking advantage of recent advances in spatial analysis, which can provide new insights that the traditional geoscience methods cannot offer.

The research is presented in three parts. I will briefly introduce them here and more details can be found in each chapter. In the first manuscript (Chapter 2), a novel spatial variance-based analysis method, SPatial Association DEtector (SPADE Cang and Luo, 2018), is introduced. The SPADE can be used to assess the associations between two continuous variables and associations between one continuous variable and one discrete variable. The SPADE improved a popular geographical association estimator, Geographical Detector(or Geo-detector), by explicitly considering the effects of distance decay, which is effect that interaction between two objects decreases as the distance between them increases, and the level of discretization for continuous variables. In this manuscript, I demonstrated the algorithm and tested the performance of the SPADE.
The research questions posed by this chapter are:

- Does the comparison of spatial heterogeneity of variables represent their association?
- Does the distance decay influence the spatial variables’ association?
- How does the spatial heterogeneity analysis measure association between two continuous variables?

In the second manuscript (Chapter 3), the junction angle of the Earth channels and Mars VNs are extracted and analogized (Cang & Luo, 2019). This chapter aims to investigate the aridity index of the Martian VNs formation environment and estimate the duration of their formation. In several terrestrial watersheds, the correlation between the junction angles and their formation environment has been investigated; however, the quantitative analysis between the junction angles and their formation environment has not been done. The quantitative correlation also has not been utilized to investigate the early Mars environmental conditions.

The research questions posed by this chapter are:

- Can junction angle indicate past climatic conditions?
- Was the ancient Martian climate condition “dry” or “wet”?
- How long did the period of precipitation on Mars last?

In the third part (Chapter 4), the terrestrial and Martian watersheds’ maturity is investigated. In the mature watershed, the hierarchy relations of elevation grids dissipate the minimum energy because of the self-organized effect of fluvial process. This chapter aims to investigate the duration of surface runoff by measuring the maturity of watersheds. The association between environmental factors and maturity are accessed and the maturities of
terrestrial and Martian watersheds are compared. The following questions were addressed in this chapter:

- Is the terrestrial watersheds close to the optimal status?
- Which environmental factors influence the maturity significantly?
- What are the maturity of Martian watersheds?
- Does Mars have the watersheds that are close to the optimal status?

1.5 Summary

This research estimated the early Martian climatic conditions by investigating the properties of widely distributed dendritic VNs on a global scale and from a spatial perspective. This dissertation built a spatial analysis method to assess the association between spatial variables by comparing their heterogeneity. This method can be used to model the association between the properties of landform and the environmental factors. The models can be applied to the Martian landform properties to infer the early climatic conditions there. I then investigate two robust properties, VNs’ junction angle and grids’ hierarchy relation, to estimate the early Martian climatic conditions. I first built the model relating climate conditions and stream junction angles on Earth and applied it to Mars to reveal climatic conditions on Mars based on VN junction angles. The grids’ hierarchy relation is modified by the self-organized effect of fluvial process until the watersheds only dissipated minimum energy. The junction angle and grids’ hierarchy relation are influenced by the environmental factors statically and dynamically, thus the conclusions from the two robust properties can be confirmed mutually. After examining the junction angle and grids’ hierarchy relation,
I found that the two results are consistent, which suggest that Mars had wide distributed surface runoff in a long enough period to alter the Martian surface.

The dissertation is composed of three related manuscripts that stand alone as individual research articles. The chapters 2-4 represent the separate research papers, each intended to be published separately. Therefore, there is some repetitiveness throughout the dissertation. For example, the SPADE method has been discussed in varying extent of the chapters. This collective design was intentional because it allowed for the publication of the material throughout the dissertation development.
2.1 Abstract

The geographical detector model can be applied to either spatial or non-spatial data for discovering associations between a dependent variable and potential discrete controlling factors. It can also be applied to continuous factors after they are discretized. However, the power of determinant (PD), measuring data association based on the variance of the dependent variable within zones of a potential controlling factor, does not explicitly consider the spatial characteristics of the data and is also influenced by the number of levels into which each continuous factor is discretized. Here, we propose an improved spatial data association estimator (termed as SPatial Association DEtector, SPADE) to measure the spatial data association by the power of spatial and multilevel discretization determinant (PSMD), which explicitly considers the spatial variance by assigning the weight of the influence based on spatial distribution and also minimizes the influence of the number of levels on PD values by using the multilevel discretization and considering information loss due to discretization. We illustrate our new method by applying it to simulated data with known benchmark association and to dissection density data in the United States to assess its potential controlling
factors. Our results show that PSMD is a better measure of association between spatially distributed data than the original PD.

2.2 Introduction

2.2.1 Introduction of Geographical Detector

The geographical detector (Geo-detector) is a relatively new spatial analysis method (Wang et al., 2010) that explores the association between variables. It was first developed in medical geography to estimate the associations between a health outcome, such as mortality rate, and risk factors, such as water pollution and social economic factors, based on their spatial distribution. The premise is that if the dependent variable is controlled by an independent variable (or potential factor), their spatial distribution will be identical or very similar and the similarity can be measured in terms of the variance of the dependent variable within zones of the independent variable (Wang et al., 2010). Specifically, if dependent variable Y is controlled by factor X, then the spatial distribution of the two should be identical or very similar, and the similarity in spatial distribution can be measured by dividing the dependent variable into zones (levels) of each independent variable and by comparing the local zonal variance and the global variance. For example, let’s assume the dependent variable (such as a health outcome or a quantitative geographical phenomenon) is sampled by a grid system \( Y = \{ Y_i | i = 1, 2, 3, ..., n \} \) (see Figure 1). A potential controlling factor (independent variable) is represented by a GIS data layer X; it is either already a categorical variable or can be discretized into a finite number of zones or levels (e.g., in Figure 1, X has 3 zones, \( X_h | h = 1, 2, 3 \)). The Power of Determinant (PD), which measures the similarity
in spatial distribution between the dependent variable Y and a potential factor X, is defined as follows,

\[ PD = q = 1 - \frac{SSW}{SST} = \frac{SSB}{SST} \]  
(2.1)

\[ SSW = \sum_{h=1}^{L} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y}_h)^2 = N_h \sigma_h^2 \]  
(2.2)

\[ SSB = \sum_{h=1}^{L} N_h (\bar{Y}_h - \bar{Y})^2 \]  
(2.3)

\[ SST = SSW + SSB = \sum_{i=1}^{N} (Y_i - \bar{Y})^2 = N \sigma^2 \]  
(2.4)

where \( SSW \) is the sum of squares within zones; \( SSB \) is the sum of squares between zones; \( SST \) is the total sum of squares; \( L \) is the total number of zones of X; \( N_h \) is the count of samples of the \( h \)th zone; \( Y_{hi} \) is the \( i \)th sample of dependent variable within \( h \)th zone of X; \( \bar{Y}_h \) is the mean of \( Y \) within the \( h \)th zone of X; \( \bar{Y} \) is the global mean of Y (i.e., mean of Y over the entire study area); \( Y_i \) is the \( i \)th sample of the entire study area and \( N \) is the total count of samples.

\subsection*{2.2.2 Analysis of Geo-detector}

The main advantage of the Geo-detector is that it makes fewer assumptions than other methods such as the regression (Wang & Xu, 2017). Geo-detector has been applied to many different fields, including physical geography (Du et al., 2016; Luo et al., 2016) and urban geography (Ren et al., 2014; Zhu et al., 2015). A complete list of applications and software
can be found on the website of the geographical detector (http://geodetector.cn/). However, there are a number of drawbacks to the original Geo-detector method. First, PD neglects the characteristics of spatial data by utilizing the sum of squares (within zones or between zones) to describe the similarity of a data set. The sum of squares, although it is used widely in statistics, cannot describe the similarity of spatial data properly, because it cannot represent the most important characteristics of spatial data, i.e., spatial dependence. The only thing that is “spatial” or “geographic” about the Geo-detector is the implied overlay of GIS data layers in the data preparation, but the method itself never explicitly considers spatial distribution and can be applied to either spatial or non-spatial data; thus, ‘Geo-detector’ is really a misnomer. Spatial dependence plays two types of roles in spatial data association: within a GIS layer or between GIS layers. Within a GIS layer, spatial dependence refers to the spatial autocorrelation. Spatial autocorrelation, also commonly called spatial association in the literature, describes the association between nearby observations and can be measured
by Moran’s I, Geary’s C or Local Indicators of Spatial Association (LISA) (Anselin, 1995). Between GIS layers, spatial dependence is described by geographically weighted regression (GWR) (Brunsdon et al., 1996). The purpose of the Geo-detector and our method is to measure the association between GIS layers. To distinguish the concept of spatial autocorrelation, we will refer to the association between GIS layers as spatial data association or simply spatial association. Spatial data association can be measured by GWR. GWR is suitable for assessing the association between continuous variables or between continuous variable and dummy variable transformed from discrete variable. For the association between variables, GWR takes advantage of the linear model in the assessing and assumes the correlations between variables are linear relationship (Brunsdon et al., 1996). However, the association measured by the Geo-detector is not limited by the linear relationship, because the Geo-detector measures the association by stratified heterogeneity (Wang & Xu, 2017; Wang et al., 2016). The stratified heterogeneity is measured by comparing the difference between strata (zones) of a data layer. Therefore, when estimating association between two continuous variables, one of them needs to be discretized into zones first. The scope of application of the Geo-detector is wider than GWR; however, the Geo-detector neglects the spatial dependence as mentioned above. In a spatial data set, the associations between locations vary with distance. Normally, based on the first law of geography, attribute values at closer locations are more similar and have stronger relations. The drawback of the sum of squares is that it treats every value equally in the same data set. To address the drawback of the sum of squares, we present a solution to this problem by using spatial variance. Besides the problem of similarity measure, the second drawback of the Geo-detector is that the selection of the number of discretization zones of continuous variables is arbitrary and may cause underestimation of the association between two continuous variables. As outlined above, the Geo-detector method works with categorical (or discrete) variables. Continuous variables must first be discretized into a finite number of zones (or levels). How a continuous
variable is discretized and how many zones it is discretized into can influence the resultant PD values. Cao et al. (2013) discretized continuous variables into two to eight zones using a number of different discretization methods (e.g., Equal Interval, Natural Breaks, Quantile, Geometrical Interval, and Standard Deviation) and compared the resultant PD values. Their results suggest that the quantile method generally produced the highest PD value and is thus better than other methods. However, they only examined a small number of zones (two to eight) and did not examine in detail the impact of number of zones on PD value. In traditional cartography, the mapping of quantitative variables as different levels of gray is usually limited to less than eight (e.g. Brewer and Pickle, 2002, Cauvin et al., 2010), because human eyes are only capable of seeing 30 shades of grays (Kreit et al., 2013) and a small number of classes is easier for humans to distinguish and understand than a large number of classes. So the rule of thumb in cartography is to use three to seven classes for grayscale representation (Harvey, 2015). With the development of geocomputation, advanced spatial analysis methods can take advantage of more classes than the limited number of grayscales designed for easy distinguishing and interpretation by humans (Kwan, 2004). Conceptually, if a continuous variable is discretized into a large number of zones, each zone will be small and the variance within each zone will also likely be smaller, leading to a larger PD value. This appears to be the case based on existing publications using the Geodetector method. For example, Wang et al. (2010) used both continuous factors (distance to fault, distance to buffer, elevation) and categorical or nominal factors (soil type, watershed and lithozone). The continuous factors were discretized to five categories. All the categorical variables have more than five categories: soil type has nine categories, watershed has nine categories and lithozone has seven categories. The Geo-detector showed that the top three variables are the three nominal variables. Cao et al. (2014) presented a spatial data discretization method which utilized local spatial autocorrelation indices to discretize continuous variables and suggested that the variables need to be discretized into more than 40 classes. They also found
that the PD value increases with the increasing number of categories. Luo et al. (2016) applied the Geo-detector method to evaluate the factors controlling the surface dissection density in the conterminous United States by physiographic regions. In that study, all the continuous factors are discretized to 6 categories and the categorical variable lithology has 21 categories, much more than 6. The results showed that for four of eight physiographic regions, the dominant factor (with maximum PD value) is lithology. To further confirm this, we discretize the continuous factor elevation into a range of categories (4—22) and calculate their corresponding PD values in Region 3 (Interior Highlands) of the United States. The result is shown in Figure 2. It is clear that there is a general trend of the PD value increasing as the number of discretization zones increases. We will refer to the number of discretization zones as the discretization level hereafter.

Figure 2: PD values of elevation under different discretization level (from 4 to 22).

Conceptually, one can imagine two extreme cases: if the discretization level is one (i.e., the whole study area is one zone), the zonal variance and global variance are the same, then the PD would be 0; if the discretization level is the same as the total number of samples of
a continuous variable (i.e., each sample is its own zone), the zonal variance will be 0 and the PD value will be 1 (see Equation 2.1). Thus, the effect of the discretization level on the PD value in the Geo-detector method must be controlled in order for PD values to be more comparable, meaningful, and interpretable across different situations. Here we present a solution to this problem using multi-level discretization and considering information loss due to discretization.

2.3 SPatial Association DEtector (SPADE) Considering Spatial Variance, Multi-level Discretization, and Information Loss

2.3.1 Spatial Variance

Spatial dependence, as the most important characteristics of spatial data, is ignored by the Geo-detector (Wang et al., 2010). The spatial dependence can be represented as spatial autocorrelation measured by Moran’s I, Geary’s C, Semivariogram, spatial interaction models, etc. The common point of these models is that they all take advantage of the spatial weighted cross-product statistic (Getis, 1991). The general form of the mean of spatial weighted cross-product is shown as below:

$$\Gamma = \frac{\sum_i \sum_{i \neq j} W_{ij} C_{ij}}{\sum_i \sum_{i \neq j} W_{ij}}$$ (2.5)

where $w_{ij}$ is the weight between $i$th location and $j$th location. Here, we set the inverse of distance as the weight (see more discussion in Discussion section). $c_{ij}$ measures the attribute similarity, such as semi-squared difference $\frac{(y_i - y_j)^2}{2}$ or absolute difference $|y_i - y_j|$. Here we will
use semi-squared difference (because doing so will make the original Geo-detector a special case of the new method, as will be shown next):

\[ c_{ij} = \frac{(y_i - y_j)^2}{2} \]  

(2.6)

In the extreme case that the weight matrix is a matrix of ones, which means that all weights between locations are equal to 1, the mean spatial weighted cross-product becomes the variance equation (Equation 2.7) (Bachmaier & Backes, 2008):

\[
\Gamma = \frac{\sum_i \sum_{i \neq j} W_{ij} (y_i - y_j)^2}{N(N-1)}
\]

\[
= \frac{1}{2} \frac{1}{N(N-1)} \sum_{i \neq j} (y_i - y_j)^2
\]

\[
= \frac{\sum_{i=1}^{N} (y_i - \overline{y})^2}{N-1}
\]

(2.7)

For \( N \) samples, there are \( N(N-1) \) pairs of cross-product. As the Equation ?? showed, the Equation 2.2 is a special case of equation 2.5 when the weights are equal to 1. Based on the equation transformation (Wang et al., 2016), the sum of squares can be represented by the product of counts of variable and variance (Equations 2.2 and 2.4). Here, the spatial sum of squares is represented by the product of variable counts and spatial variance. We define the ratio of local spatial sum of squares and the global counterpart as power of spatial determinant (PSD) or \( q_s \):

\[ q_s = 1 - \frac{\sum_{h=1}^{L} N_h \Gamma}{N \Gamma} \]

(2.8)

where \( N_h \) is the total count of samples in \( h \)th category; \( \Gamma_h \) is the spatial variance within level \( h \); \( L \) is the total number of levels; \( N \) is the total count of samples; \( \Gamma \) is the total spatial
variance. To give a specific example, let’s assume $X$ and $Y$ are two spatial variables which cover the same area. $Y$ is a continuous dependent variable. $X$ is an independent categorical variable and has $h$ levels, so we do not need to consider the discretization process in this example. The spatial data association ($q_s$) or PSD between $Y$ and $X$ is:

$$q_s = 1 - \frac{\sum_{h=1}^{L} N_h \Gamma_h}{NT_t}$$

$$= 1 - \frac{\sum_{h=1}^{L} N_h \sum_{i=1}^{N_h} \sum_{j \neq i}^{N_h} w_{hi,hj} \frac{(y_{hi} - y_{hj})^2}{2}}{\sum_{i=1}^{N} \sum_{j \neq i}^{N} w_{i,j} \frac{(y_i - y_j)^2}{2}}$$

(2.9)

where $w_{(i,j)}$ is the weight between the $i$th location and the $j$th location, which is taken as the inverse of distance in our calculation; subscript $hi$ and $hj$ are the $i$th and $j$th samples in the $h$th category.

### 2.3.2 Multi-level Discretization and Information Loss

Discretization level corresponds to the minimum perception of measurement. The discretization makes a variable easier to be understood. In essence, discretization keeps the main information of the variable and ignores the finer details of information. The smaller the discretization level (i.e., fewer categories), the more the information loss. Too few levels cannot represent variation within the original variable well, but too many levels may result in redundant information and could thus hinder the analysis and interpretation. However, it is always difficult to find the best discretization level, because the best discretization level is related to the characteristics of the data and the specific research questions (e.g., Liao et al., 2010). Because the study fields are different, the best scales are not even the same for similar research works (Cao et al., 2014; Wang et al., 2010). Thus it is desirable to
conduct the analysis at multiple discretization levels and compensate information loss due to discretization. As shown in Figure 3, we discretize the continuous independent variable into different levels and apply spatial variance to each discretization level.

Figure 3: Discretizing continuous independent variable using different discretization levels (zones)

In this article, we select the quantile discretization for three reasons. First, a previous comparative research suggested that the Quantile method is better than others discretization methods (e.g., Equal Interval, Natural Breaks, Geometrical Interval, and Standard Deviation) (Cao et al., 2013). Second, the example of the current paper is from a previous research project (Luo et al., 2016). In that paper, Luo et al. (2016) utilized the quantile discretization. Selecting the same discretization method will greatly facilitate the comparison. Third, the quantile method can minimize the information loss as measured by information entropy, which is defined as Equation 2.10 (Baez et al., 2011; Quinlan, 1986):

\[
F = - \sum_{i=1}^{N} p(i) \log_2 p(i) - (\sum_{h=1}^{L} p(h) \log_2 p(h)) \tag{2.10}
\]

where \(N\) is the total count of samples; \(p(i)\) is the probability of finding sample \(i\); \(h\) is the \(h\)th level, \(p(h)\) is the probability of a data point or sample belonging to the \(h\)th level. The
The first term of Equation 2.10 is the information contained in the original un-discretized data; the second term is the information contained after discretization; $F$ is the information lost due to discretization. To minimize Equation 2.10, the second term should be maximized and this is achieved when all the $p(h)$s are equal (MacKay, 2003, p43-44), i.e., Quantile discretization. For example, if we have 100 samples (of a continuous variable) and discretize them to 4 levels ($L = 4$), $F$ will be minimized (second term maximized) if $p(h) = 0.25$. There are two extreme examples, which will not be used in practice but can help understand the concept. One is that if $L = N$ (i.e., there is no discretization), then $F = 0$ which means that there is no information loss; the other is that if $L = 1$ (i.e., all samples are grouped in one level), then $F$ reaches the maximum (first item of this equation), which means that all the information is lost. As described in Figure 3, the SPADE uses levels of the independent variable to stratify (or divide) the dependent variable and to compare the zonal spatial variance and global spatial variance as a way to measure the spatial association between them. It only measures the relationship between the continuous dependent variable and the discretized independent variable, but omits the information loss as a result of the discretization. The information remain after discretization can be measured by examining the spatial data association between the original continuous independent variable and its discretized counterpart. If all information is kept (i.e., loss = 0), the spatial data association between the original continuous independent variable and its discretized counterpart should be 1. Following this line of reasoning, we can calculate the spatial data association between the original continuous independent variable and its discretized counterpart as a measure of the information still remain after discretization (We will call this $q_{s,info\_kept}$ to distinguish it from the $q_s$ as defined in Equation 2.8). The information kept can be expressed as:

$$q_{s,info\_kept} = 1 - \frac{\sum_{h=1}^{L} N_h \Gamma_{h,ind}}{N \Gamma_{total,ind}} \quad (2.11)$$
where the subscript $h_{\text{ind}}$ represents the $h$th level of independent variable; $\Gamma$ represents the spatial variance.

We define the new compensated power of spatial discretization determinant ($Q_s$) as the ratio of the above two quantities (i.e., relative to the information kept, thus compensating for the information loss due to discretization):

$$Q_s = \frac{q_s}{q_{s,\text{info kep}}} = \frac{1 - \frac{\sum_{h=1}^{L} N_h \Gamma_{k,\text{dep}}}{N_{\text{total,dep}}}}{1 - \frac{\sum_{h=1}^{L} N_h \Gamma_{k,\text{ind}}}{N_{\text{total,ind}}}}$$  \quad (2.12)

For reasons that will become clear later, we define the power of spatial and multi-level discretization determinant as follows:

$$PSMD\_Q_s = MEAN(Q_s)$$  \quad (2.13)

where $MEAN$ represents the mean of $Q_s$ at all discretization levels.

### 2.3.3 Test of Significance

After some transformations (Wang et al., 2016), the probability density function (PDF) of PDs is a noncentral F-distribution with the first degree of freedom ($d.f.)L - 1$, the second $d.f.N - L$, and noncentrality $\lambda$. The null hypothesis (two variables have no association) can be tested by critical value test. The PDF of PSMD is unknowable because it is influenced by the spatial pattern of locations which varies in each case. If all the weights are equal to 1, the PD is a noncentral F-distribution, just as the Geo-detector paper showed (Wang et al. 2016). Although the PDF of PSMD cannot be known, the null hypothesis (two variables have no association) can still be tested. Following the idea of current spatial analysis software
(such as ArcGIS, PySAL and GeoDa) (Anselin et al., 2006; Rey & Anselin, 2010), the null hypothesis can be tested by a Z-test, which assumes that the PDF of PSMD under null hypothesis is a normal distribution; or it can be tested by the pseudo p-value approach, which has a broader constraint of distribution than traditional significance test. For an observed PSMD, if the z-score (see Equation 2.14) is greater than 1.96 or if pseudo p-value (see Equation 2.15) is smaller than 0.05, the confidence level is above 95%. We select the pseudo p-value approach. The reason will be explained in the Discussion Section.

\[
Z = \frac{PSMD_{obs} - MEAN(PSMD_{random})}{SE(PSMD_{random})} \tag{2.14}
\]

where \(PSMD_{obs}\) is the observed PSMD; \(MEAN(PSMD_{random})\) is the mean of \(PSMD_{random}\); \(SE(PSMD_{random})\) is the standard error of the mean of \(PSMD_{random}\); \(PSMD_{random}\) is an array including \(M(99,999,\text{etc.})\) PSMDs which are calculated from randomization null hypothesis.

\[
pseudo_p = \frac{R + 1}{M + 1} \tag{2.15}
\]

where \(R\) is the number of times a computed statistic from the random data sets is equal to or more extreme than the observed PSMD; \(M\) is the number of permutations (99, 999, etc.). For a given spatial data set, the critical value can be calculated by the following steps. We calculate the PSMD under the situation that all the dependent and independent values are rearranged randomly. When we select 0.05 as the critical \(p\)-value, the \(R\)-th (\(R = 0.05 \times (M + 1) - 1\)) highest PSMD from null hypothesis is the critical PSMD value. If the observed PSMD value is less than the critical PSMD, the association between dependent and independent variable is not significant.
2.4 Simulation Test of PSMD

2.4.1 Simulation Scheme

To compare the result of our new method with Wang et al. (2010), We design simulations to test the performances of our method. The simulation includes two parts: spatial data generation and spatial data association assessment. The basic idea is to create simulated variables $Y$ and $X$ with perfect spatial association (=1); we then randomly shuffle the variable $X$ with known shuffling rate and use PSMD to measure the spatial association and compare that with the benchmark association, which is 1 minus the shuffling rate. To generate the variables with perfect spatial association, we first create an area with $30 \times 30$ small lattices. Then, we select 50 positions within this area (the locations of points are random, so probably are not at the centers of lattices) and assign 50 random numbers from a random distribution (e.g. Gaussian $\sim N(0,10)$) to these points. To obtain a spatial surface data, we select the radial basis function to interpolate the values at the centers of all lattices using the 50 random distributed points. The set of center point values and their locations is the spatial dependent variable ($Y$). After we created the spatial dependent variable, we produce the independent variable ($X$) as follows. We generate 900 ($30 \times 30$) randomly distributed (e.g. Gaussian $\sim N(0,10)$) random values, rank them and assign them to the center points of lattices based on the ranking of the dependent variable. After this step, the rankings of dependent and independent variable are matched perfectly. In this case, the ranking of dependent values and their corresponding independent values are the same. The association between two data reaches maximum which is 100% or 1. After we created the maximum association spatial data set, we use the controlled parameter (shuffling rate) to decrease the maximum association to generate benchmark association for testing the measures of spatial
association. We switch some independent values with each other randomly (i.e. shuffling) to wane the association between dependent and independent variables. The shuffling rate is the ratio between the counts of rearranged independent values and the total counts of independent values (so the range of shuffling rate is from 0 to 1). If the shuffling rate is 0, all the independent values are not rearranged; the association between spatial dependent and independent variables is maximum, in other words, it is 1. Conversely, if the shuffling rate is 1, all the independent values are rearranged randomly; the association between the spatial dependent and independent variable is 0 (random).

2.4.2 Estimated Spatial Data Association Comparison by Simulations

To evaluate our SPADE, we compare the PD (from Wang et al., 2010) with PSMD (proposed by this article) under the controlled shuffling rates. First, under some fixed shuffling rates, we compare the PD, compensated PD (PD compensated by the information loss, or CPD), PSD (PD using spatial variance, Equation 2.9) and compensated PSD (PSD compensated by the information loss, or CPSD Equation 2.12). We also calculate the critical value by the pseudo p-value method to test whether the estimated associations are significant. The aims of comparison are as follows: (1) to illustrate which estimator is stable across different discretization levels and (2) to test if the estimated values follow the general trend of different benchmark associations. The results under different shuffling rates (0.0, 0.2, 0.4, 0.6, 0.8, 1.0) are shown in Figure 4. Each point represents the mean of 100 simulations. Figure 5 shows the comparisons between the estimated associations and critical values. Following pseudo p-value, we repeated 99 times of the calculation of the spatial association under the full shuffling rate (1.0) and obtained 4th largest value of simulation results \( p = 0.05 \) as the critical value \( R = 0.05 \times (M + 1) - 1 = 0.05 \times (99 + 1) - 1 = 4 \).
The results showed that (1) the compensated estimators (CPD and CPSD) are generally stable throughout the discretization levels; (2) the stabilities of CPD and CPSD only have subtle difference; (3) when the shuffling rate is 100%, all the estimators are not significant (shown in Figure 5); (4) when the shuffling rate is 0%, the compensated estimators (CPD and CPSD) reach 1 (shown in Figure 4) and (5) overall, our CPSD more closely follow the trend of different benchmark associations.

2.5 Application to U.S. Surface Dissection Density Data

We applied the new approach to derive PSMD as described above to better capture spatial associations between dissection density and environmental factors and compared them with the results from Luo et al. (2016). Dissection density describes the degree of land surface dissection by erosional processes and is defined as the total length of valleys per unit area ((Luo et al., 2016)). On a continental scale, dissection density (a geomorphology concept not requiring identification of channels) is highly correlated with drainage density (a hydrology concept requiring identification of channels). Previous research projects have shown that factors controlling dissection density include climate (Melton, 1957; Montgomery & Dietrich, 1989; Tucker & Bras, 1998), slope and relief (Oguchi, 1997; Schumm, 1956; Strahler, n.d.), lithology (Gardiner, 1995; Tucker & Slingerland, 1996; Xiong et al., 2014) and soil properties (Dietrich et al., 1992; Montgomery & Dietrich, 1989). The understanding of which factors play dominant roles in controlling dissection density is an important theme in geomorphology and hydrology because of its scientific and practical values. The latter relates to assessing the risk of soil loss and to designing measures to reduce such loss. Unlike most previous studies, which were at local scales and lacked an analytical framework designed especially for comparing controlling factors over a regional or continental scale, Luo et al. (2016) utilized the
Geo-detector as a general framework to assess the associations between dissection density and environmental factors in each physiographic region (Fenneman 1928, Figure 6) and tested the hypothesis that the dominant controlling factor, or the interactions between factors, vary from region to region due to differences in each region’s local characteristics and geologic history. The dissection density data were derived using geomorphons method (Jasiewicz & Stepinski, 2013) and aggregated to the basic units of watersheds based on the 12-digit hydrologic unit boundaries (Federal standards and procedures for the national Watershed Boundary Dataset). The 13 controlling factors are shown in Table 5, which include 3 main groups: geology, climate and terrain, and are aggregated to the same basic units of watersheds. The terrain factors are derived from Digital Elevation Model (DEM) data. The geology and climate factors are downloaded from open source database: the precipitation factor is from the website of PRISM climate Group (http://www.prism.oregonstate.edu/), the glaciation and the lithology factors are from United States Geological Survey (USGS), the permeability factor is from Gleeson et al. (2014) (http://crustalpermeability.weebly.com/glhymps.html) and the porosity factor is from the STATSGO2 database.

To illustrate the PSMD, we use Region 3 (Interior Highlands), Region 6 (Laurentian Upland), Region 7 (Pacific Mountain System) and Region 8 (Rocky Mountain System) as examples. In the calculation, the discretization levels in the multilevel discretization are from 4 to 20. Table 6 shows the PSMD values and their rankings in those regions in comparison with those of the original PD. The spatial association for most continuous variables increased. The ranking order of the spatial associations between dependent and independent variables also changed slightly. The ranking of discrete variables, such as the Litho, decreased. However, the dominant factors (high ranked factors in PD) still retain their statuses. To illustrate the physical meaning of the PSMD, we map the three controlling factors, which are high, medium and low ranking, to compare the spatial distribution of dissection
Figure 4: Comparison of PD, compensated PD (CPD), PSD and compensated PSD (CPSD), under different shuffling rates

The benchmark association equals 1 minus the shuffling rate. The suffix _mean represents the mean of simulation results. X-axis is the discretization level which is from 5 to 30 and Y-axis is estimated spatial association between two variables.
Figure 5: Comparison of estimated spatial association and critical value under different shuffling rates. X-axis is the discretization level and Y-axis is estimated spatial association between two variables.

density (Figure 7). It is clear that the controlling factors with higher PSMD have a more similar spatial distribution with the dependent variable.

The results (Table 6 and Figure 7) show that (1) the associations between dependent and independent continuous variables are underestimated by the original PD in most cases, because the original Geo-detector omits the spatial dependence resulting from dynamical geographical processes and the information loss due to discretization; (2) The information loss from discretizing different continuous independent variables is different, because their distributions are different. Thus, the ranking of the continuous independent variables also changed.
2.6 Discussion

2.6.1 Selection of Model Parameters

The major parameters of SPADE include the range of discretization level and the weighting method. The range of discretization level is a robust parameter because the information loss is compensated. From Figure 4, the compensated PSDs are stable across different discretization levels; most of the differences of the compensated PSDs between different discretization levels are smaller than 0.1. Here, we apply the different discretization level ranges

Figure 6: Dissection density in the US
<table>
<thead>
<tr>
<th>Category</th>
<th>Factor</th>
<th>Factor code</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geology or soil property</td>
<td>Glaciation</td>
<td>Glaci</td>
<td>Resampled shapefile data to 4 km</td>
</tr>
<tr>
<td></td>
<td>Lithology</td>
<td>Litho</td>
<td>1 km resolution, resampled to 4 km</td>
</tr>
<tr>
<td></td>
<td>Permeability</td>
<td>logK</td>
<td>Resampled shapefile data to 4 km</td>
</tr>
<tr>
<td></td>
<td>Porosity</td>
<td>poro</td>
<td>Resampled shapefile data to 4 km</td>
</tr>
<tr>
<td>climate</td>
<td>Precipitation</td>
<td>precip</td>
<td>4 km resolution</td>
</tr>
<tr>
<td></td>
<td>Elevation</td>
<td>elev</td>
<td>ETOPO1 DEM resampled to 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>Aspect</td>
<td>asp</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td>Topography</td>
<td>Slope</td>
<td>slp</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>difference in elevation (relief)</td>
<td>difelev</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>distance to erosional base</td>
<td>distb</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>Elevation to erosional base</td>
<td>elevb</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>Planar Curvature</td>
<td>planc</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
<tr>
<td></td>
<td>Tangential Curvature</td>
<td>tanc</td>
<td>Derived from DEM, 4 km resolution</td>
</tr>
</tbody>
</table>

on Region 3 as an example to illustrate the effect of discretization level ranges. The results are shown in Table 7. We apply three groups of discretization level ranges (5–10, 10–15 and 15–20) to calculate PSMDs. The results showed that the rankings of variables are the same and that most of differences between different range are smaller than 0.05. The weighting method includes contiguity-based weights or distance-based weights. The selection of method depends on the nature of the research. In this research, the distance-based weights are more suitable, because the scale of dynamic geography processes (erosion, sedimentation and tectonic processes) is much larger than the statistics unit (watershed), which means that noncontiguous watersheds can be influenced by the same process and the intensity of influ-
ence decays through the distance (Taylor & Openshaw, 1975). The distance decay can be measured by different functions (summarized by Martinez and Viegas 2013) quantitatively. In this article, we choose the power law because the gravity model, a form of power law, has been used in different geography research fields, both in human geography and in physical geography. The general form of weight based on the power law of distance decay is shown in Equation 2.16:

\[ w = \frac{1}{d^\beta} \]  

(2.16)

where \( w \) is the weight between two locations; \( d \) is the Euclidean distance between two locations and \( \beta \) is the exponent of distance, representing the decay rate. The \( \beta \), with a typical range from 0 to 3 (Chen, 2015), is the intensity of influence between neighbors. A larger \( \beta \) in Equation 2.16 means that the closer values have higher weights in the calculation. The selection of \( \beta \) is usually research dependent and often determined empirically; or can be estimated by the product of the Zipf’s exponent of size distributions and the fractal dimension of spatial distributions (Chen & Huang, 2018). In producing Table 6, we used inverse distance decay, that is, \( \beta = 1 \). Here, we test the robustness of \( \beta \) by applying two \( \beta \) values (1, 2) on the same region (Table 8). The dominant factors of Regions 3 and 6 changed subtly, and most of ranking only changed one or two positions. The results showed that the ranking of variables may change with different weighting methods, but they are not very sensitive to different weighting methods.

2.6.2 Assumption of the Probability Distributions

Normally, the assumption of association estimator is that the PDF of variables are normal distribution; however, in the geography research field, the symmetrical distribution and
Figure 7: Dissection density and its controlling factors in Region 3.
(a) Dissection density (dependent variable); (b) logk (highest PSMD); (c) elevation (7th controlling factor); (d) tanc (11th controlling factor).
Figure 8: Comparison spatial estimators form different distributed variables
X-axis is the shuffling rate and Y-axis is estimated spatial association between two variables. The
suffix _q represents the estimator without compensation and the suffix _Q represents the compensated estimator.
Figure 9: QQ-plot of null hypothesis under varied distributions
The red rectangle shows the critical values (fourth point).
Table 2: PSMD results in Regions 3, 6, 7 and 8

<table>
<thead>
<tr>
<th>Ranking</th>
<th>Region 3 (Interior Highlands)</th>
<th>Region 6 (Laurentian Upland)</th>
<th>Region 7 (Pacific Mountain System)</th>
<th>Region 8 (Rocky Mountain System)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Variable</td>
<td>PSMD</td>
<td>Variable</td>
<td>PD</td>
</tr>
<tr>
<td>1</td>
<td>logk</td>
<td>0.389</td>
<td>logk</td>
<td>0.309</td>
</tr>
<tr>
<td>2</td>
<td>poro</td>
<td>0.363</td>
<td>precip</td>
<td>0.262</td>
</tr>
<tr>
<td>3</td>
<td>precip</td>
<td>0.348</td>
<td>poro</td>
<td>0.219</td>
</tr>
<tr>
<td>4</td>
<td>distb</td>
<td>0.336</td>
<td>litho</td>
<td>0.185</td>
</tr>
<tr>
<td>5</td>
<td>difelev</td>
<td>0.280</td>
<td>elevb</td>
<td>0.175</td>
</tr>
<tr>
<td>6</td>
<td>elevb</td>
<td>0.258</td>
<td>distb</td>
<td>0.163</td>
</tr>
<tr>
<td>7</td>
<td>elev</td>
<td>0.235</td>
<td>slp</td>
<td>0.154</td>
</tr>
<tr>
<td>8</td>
<td>litho</td>
<td>0.235</td>
<td>difelev</td>
<td>0.127</td>
</tr>
<tr>
<td>9</td>
<td>slp</td>
<td>0.166</td>
<td>elev</td>
<td>0.122</td>
</tr>
<tr>
<td>10</td>
<td>planc</td>
<td>0.123</td>
<td>planc</td>
<td>0.107</td>
</tr>
<tr>
<td>11</td>
<td>tanc</td>
<td>0.118</td>
<td>tanc</td>
<td>0.094</td>
</tr>
<tr>
<td>12</td>
<td>asp</td>
<td>0.042</td>
<td>glaci</td>
<td>0.057</td>
</tr>
<tr>
<td>13</td>
<td>glaci</td>
<td>0.001</td>
<td>asp</td>
<td>0.049</td>
</tr>
</tbody>
</table>

N is the number of watersheds in the region. All the PSMD estimated associations are significant, except that the values with the symbol ^, representing that the values are not significant. The PSMD is the mean of PD using spatial variance under different discretization levels. The original PDs are from Luo et al. (2016).
Table 3: PSMD of continuous variables under different discretization levels in Region 3

<table>
<thead>
<tr>
<th></th>
<th>5_to_10</th>
<th></th>
<th>10_to_15</th>
<th></th>
<th>15_to_20</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$logk$</td>
<td>0.396</td>
<td>$logk$</td>
<td>0.354</td>
<td>$logk$</td>
<td>0.422</td>
<td></td>
</tr>
<tr>
<td>$poro$</td>
<td>0.367</td>
<td>$poro$</td>
<td>0.335</td>
<td>$poro$</td>
<td>0.389</td>
<td></td>
</tr>
<tr>
<td>$precip$</td>
<td>0.362</td>
<td>$precip$</td>
<td>0.324</td>
<td>$precip$</td>
<td>0.364</td>
<td></td>
</tr>
<tr>
<td>$distb$</td>
<td>0.341</td>
<td>$distb$</td>
<td>0.319</td>
<td>$distb$</td>
<td>0.350</td>
<td></td>
</tr>
<tr>
<td>$difelev$</td>
<td>0.288</td>
<td>$difelev$</td>
<td>0.262</td>
<td>$difelev$</td>
<td>0.295</td>
<td></td>
</tr>
<tr>
<td>$elevb$</td>
<td>0.259</td>
<td>$elevb$</td>
<td>0.251</td>
<td>$elevb$</td>
<td>0.265</td>
<td></td>
</tr>
<tr>
<td>$elev$</td>
<td>0.245</td>
<td>$elev$</td>
<td>0.206</td>
<td>$elev$</td>
<td>0.258</td>
<td></td>
</tr>
<tr>
<td>$slp$</td>
<td>0.168</td>
<td>$slp$</td>
<td>0.158</td>
<td>$slp$</td>
<td>0.174</td>
<td></td>
</tr>
<tr>
<td>$planc$</td>
<td>0.125</td>
<td>$planc$</td>
<td>0.118</td>
<td>$planc$</td>
<td>0.126</td>
<td></td>
</tr>
<tr>
<td>$tanc$</td>
<td>0.121</td>
<td>$tanc$</td>
<td>0.109</td>
<td>$tanc$</td>
<td>0.125</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: PSMDs under different distance decay methods

<table>
<thead>
<tr>
<th></th>
<th>Region 3</th>
<th></th>
<th>Region 6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\beta=2$</td>
<td>$\beta=1$</td>
<td>$\beta=2$</td>
<td>$\beta=1$</td>
</tr>
<tr>
<td>$distb$</td>
<td>0.406</td>
<td>0.389</td>
<td>0.406</td>
<td>0.214</td>
</tr>
<tr>
<td>$logk$</td>
<td>0.402</td>
<td>0.363</td>
<td>0.402</td>
<td>0.202</td>
</tr>
<tr>
<td>$difelev$</td>
<td>0.393</td>
<td>0.348</td>
<td>0.393</td>
<td>0.167</td>
</tr>
<tr>
<td>$poro$</td>
<td>0.391</td>
<td>0.336</td>
<td>0.391</td>
<td>0.164</td>
</tr>
<tr>
<td>$elevb$</td>
<td>0.355</td>
<td>0.280</td>
<td>0.355</td>
<td>0.148</td>
</tr>
<tr>
<td>$precip$</td>
<td>0.338</td>
<td>0.258</td>
<td>0.338</td>
<td>0.138</td>
</tr>
<tr>
<td>$elev$</td>
<td>0.322</td>
<td>0.235</td>
<td>0.322</td>
<td>0.136</td>
</tr>
<tr>
<td>$slp$</td>
<td>0.225</td>
<td>0.235</td>
<td>0.225</td>
<td>0.131</td>
</tr>
<tr>
<td>$tanc$</td>
<td>0.187</td>
<td>0.166</td>
<td>0.187</td>
<td>0.125</td>
</tr>
<tr>
<td>$planc$</td>
<td>0.147</td>
<td>0.123</td>
<td>0.147</td>
<td>0.104</td>
</tr>
<tr>
<td>$litho$</td>
<td>0.133</td>
<td>0.118</td>
<td>0.150</td>
<td>0.100</td>
</tr>
<tr>
<td>$asp$</td>
<td>0.106</td>
<td>0.042</td>
<td>0.147</td>
<td>0.099</td>
</tr>
<tr>
<td>$glaci$</td>
<td>0.053</td>
<td>0.001</td>
<td>0.128</td>
<td>0.017</td>
</tr>
</tbody>
</table>

asymmetrical distribution are both very common. For this reason, we create random variables which follow different PDF to test the influence of different PDFs on SPADE. First, we investigate the PSMD under different distributions. We select the normal distribution to represent the symmetrical distribution and select Pareto distribution to represent the asymmetrical distribution. We create four groups of data sets whose dependent and independent
variables are normal distribution \((N(0, 10))\) or Pareto distribution \((P(3))\). Then, we analyze the influence of variables’ distribution on the association estimation. Based on simulation test described in Section 3, we average the compensated PSDs from 5 to 30 discretization levels as the result of multilevel associations to compare the association to the shuffling rates. The results are compared with the spatial association benchmark line \((y = -x, x \in [0, 1])\) (see Figure 8). The benchmark values of spatial association are from 0 to 1, whose interval is 0.2. The results showed that the compensated estimators (PMD\(_Q\) and PSMD\(_Q\)) can cover the full range of benchmarks and are less sensitive to the changes of variables’ distributions than PMD\(_q\) and PSMD\(_q\).

### 2.6.3 Selection of Significant Test

As mentioned above, the PDF of PSMD is unknowable. The current spatial analysis software (such as ArcGIS, PySAL and GeoDa) (Anselin et al., 2006; Rey & Anselin, 2010) test the significance of spatial estimator by Z-test or pseudo p-value approach. We use the simulations to discuss the difference between the Z-test and pseudo p-value approach and explain why we choose p-value approach. We repeat 99 times of the calculation of PSMDs under the 100% shuffling rate. In each simulation, first, we create four pairs of possible combinations from the two PDFs (normal distribution \((N(0, 10))\) or Pareto distribution \((P(3))\)) in one time of simulation; next, shuffle one variable of every pair with 100% shuffling rate; then, calculate the PSMDs. Based on the four groups of PSMDs, the QQ-plot (Figure 9) shows the similarity between null hypothesis distribution and normal distribution. At the right end of each line, the actual values are a little bit greater than the expected value. The shape represents that the tail of distribution from null hypothesis is heavier than the normal distribution. In this situation, the pseudo p-value approach has a higher critical value than
the normal distribution significance test for a given spatial estimator. For this reason, we select pseudo p-value approach because it is more conservative.

2.7 Conclusion

The purpose of this article is to improve the measure of spatial association between dependent variable and potential controlling factors within the Geo-detector framework by explicitly utilizing the spatial information and minimizing the influence of discretization levels. We solved the problem that the original Geo-detector lacks measure of spatial dependence by utilizing spatial variance, which is derived from general spatial weighted cross-product, to replace the traditional variance. We also addressed the problem that the Geo-detector measured association (PD) was influenced by the number of levels into which continuous variables are discretized by compensating the information loss due to discretization. The information kept was measured by PSD value between the continuous variable and its discretized counterpart. Using simulated data with known benchmark association, we demonstrated that (1) the compensated association can cover the whole range of benchmark association; (2) the compensated association is stable across different discretization levels and (3) the significance of null hypothesis (association between variables not significant) can be tested by pseudo p-value approach, whose result is more conservative than Z-test which assumes that the distribution from null hypothesis is a Gaussian distribution. When applying the new method to measure the spatial association between dissection density and controlling factors in United States, the ranking of PSMD values of some variables changed, but most dominant factors still remain the same. So the general conclusion of Luo et al. (2016) that the dominant factor for each physiographical region reflects that region’s geological history and character still holds. The exception happened in Regions 6 and 8, both with low original PD value.
The dominant factors of the two regions were litho, which means that a higher category number can cause a higher estimated spatial association and that the previous discretization level in Luo et al. (2016) underestimated the associations between dependent and independent variables. Through our simulated data with known benchmark association and case study of dissection density in United States, we have demonstrated that the PSMD value is a stable and more accurate measure than original PD because PSMD explicitly considers spatial variation and minimizes the influence of discretization levels. Thus, in practice, we have more confidence in using PSMD to measure the association between spatial data and do not need to discretize the continuous variables into a large number of levels.
CHAPTER 3

NOACHIAN CLIMATIC CONDITIONS ON MARS INFERRED
FROM VALLEY NETWORK JUNCTION ANGLES

3.1 Abstract

Valley networks (VNs) on Mars offer convincing evidence for its past water activities. Previous research empirically and theoretically suggested that the geometry properties, such as junction angles, of streams on Earth formed under different climatic conditions are different. Thus, the geometry properties of VNs can be used to infer the early Mars climatic conditions under which VNs were formed. The frequency distribution of junction angles is less influenced by the post-formational modification processes than valley density and cross-section and can be accurately extracted from low resolution data. We first analyzed the association between the junction angles and environmental factors on Earth. The results suggested that the climatic factors are stronger than or on par with the geologic factors in controlling junction angles and that climatic parameters Aridity Index ($AI$) and Mean Annual Precipitation ($MAP$) can be estimated from junction angles. We then applied the associations between terrestrial junction angle and climatic conditions to estimate the $AI$ and $MAP$ of Mars, considering the different solar radiation of the two planets. The spatial analysis of inferred climatic conditions showed that the Noachian Mars had an active global

---

This chapter was previously published as “Cang, X., & Luo, W. (2019). Noachian climatic conditions on Mars inferred from valley network junction angles, 526.” It is reprinted by permission of Elsevier.
hydrological cycling and Mars was “warm” during the VNs’ formation period. The duration of “warm” Mars is estimated by the ratio between the required water volume to form VNs and the runoff discharge derived from MAP. The range of “warm” period is from \( \sim 4.4 \) to \( \sim 77 \) million years. The results support the hypothesis that Mars was “episodically warm”.

### 3.2 Introduction

Mars is currently a cold and dry planet, where liquid water could not exist stably on its surface. However, observations of valley networks (VNs) (Craddock & Howard, 2002), deltas (Di Achille & Hynek, 2010) and tsunami deposits (Rodriguez et al., 2016), among other evidences, suggest that Mars once had abundant liquid water during the Noachian period and the early Mars climate was “warm” enough to support liquid water (e.g., Craddock and Howard, 2002). Despite the many lines of geomorphic, geochemical, and geologic evidence that suggest an early warm Mars with abundant liquid water and an active hydrologic cycle (Luo et al., 2017), climate modelers have encountered difficulties in modeling such early warm and wet conditions with an above freezing temperature, mainly due to the Martian orbit and the faintness of the young Sun (e.g., Wordsworth et al., 2015). Since Mars is in an orbit farther away from the Sun, it receives only about 43% of the solar energy that Earth does and the Sun’s luminosity 3.8 billion years ago (when most of the VNs were believed to be formed) was only \( \sim 75\% \) of its present value (Gough, 1981). Wordsworth et al. (2015) compared the spatial distribution of modeled precipitation/ice accumulation with that of the VN density and found no correlation between them. In particular, the model predicted high precipitation in Arabia Terra, where low VN density was observed there. However, Davis et al. (2016), through detailed regional analysis and mapping of the orbiter images over Arabia Terra, showed evidence for extensive networks of inverted sinuous ridges, suggesting
that previous VNs were eliminated by the resurfacing process; thus the evidence supports that the Mars had a more widespread precipitation and runoff than the modeled results from the climate model (Wordsworth et al., 2015). To reconcile the different views on early Mars climate, a number of alternative scenarios have been proposed. For example, the early climate of Mars could be warm and wet episodically because of brief and strong volcanic activities and associated outgassing of greenhouse gasses and aerosols (Halevy & Head III, 2014); VNs could be formed by groundwater sapping associated with magma intrusion and hydrothermal activities, thus not requiring continuous warm and wet conditions (Gulick, 1998); VNs could be created during short-lived episodes of top-down melting of thick cold-based ice on the equatorial highlands (Rosenberg et al., 2019; Wordsworth et al., 2015).

The formation mechanism of Mars VNs is important in understanding the early Mars climate. The VNs' characteristics, such as cross-section shape, stream order, scaling relation, have been investigated quantitatively to derive the climate condition of the Noachian period (Hynek & Phillips, 2003; Penido et al., 2013; Williams & Phillips, 2001). Their results support that Mars was “warm” enough to have liquid water flowing on the surface in the Noachian Mars; however, they did not answer the question about duration of “warm” period or the formation timescale of VNs, which directly relates to the long term climatic conditions. This question is answered by some research which applied the sediment transport models and the characteristic of VNs (depth, width, volume, etc.) to estimate the cumulative volume of water required to carve the VNs (Luo et al., 2017; Rosenberg et al., 2019) and the formation timescale of VNs (Hoke et al., 2011; Orofino et al., 2018). Due to the large uncertainties of derived VN characteristics data and the difference between models, the conclusions from these studies are different or even mutually exclusive (Penido et al., 2013; Som et al., 2009). Hitherto, scientists are still debating on whether episodic warmth or nonprecipitation dominated erosion can create the observed valley networks and trying to reconcile the geochemical and physical understanding of Mars with the geologic evidence of
its watery past (Baker et al., 2015; Ehlmann et al., 2011; Wordsworth, 2016). Although the controversies still exist, the consensus between supporters of “Warm” Mars and “Cold” Mars is that abundant liquid water flowed on Mars in the Noachian period (Ramirez & Craddock, 2018; Wordsworth et al., 2018). Since characteristics used by earlier studies were likely influenced by the impact cratering and post-formation processes, it is necessary to choose a new characteristic of VNs, which is less influenced by the impact cratering and post-formation process, to estimate the climatic condition of the Noachian Mars.

The junction angle between two joining streams is one of the main characteristics of stream networks. Perron et al. (2012) revealed that the junction angle is not an outcome of random topology, but “an organized signature of erosional mechanics.” To predict the junction angles, Horton (1932;1945) investigated the association between junction angles and other characteristics of basin, such as the slope/area ratio. Seybold et al. (2017) showed that the influence of aridity index to junction angles is stronger than the influence of elevation, slope and stream network concavity and suggested that the strength of coupling between groundwater and surface water processes influenced the size of junction angle.

Unlike other characteristics of streams, junction angle is less influenced by post-formational modification and can be accurately extracted even from low resolution data. For this reason, the junction angles of VNs on Mars were investigated in early research (Pieri, 1980). There is renewed interest in junction angles in recent literature that suggest that terrestrial stream junction angles and their frequency distribution reflect the relative dominance in the underlying channel forming processes and the climatic conditions under which the streams were formed (Hooshyar et al., 2017; Seybold et al., 2017). Seybold et al. (2018) extended the research area of Seybold et al. (2017) from the conterminous U.S. to the global Earth surface, and analogized the junction angles of Earth and Mars to examine the paleoclimate in the Noachian period.
Seybold et al. (2017) analyzed millions of stream junctions throughout the continental U.S. and showed that mean junction angles by hydrologic units vary systematically with $AI$ (aridity index = precipitation / potential evapotranspiration): $\sim 45^\circ$ in the driest places and $\sim 72^\circ$ in the humid areas. They interpreted the observed trend as due to the different geomorphic processes controlled by climate. Seybold et al. (2017) concluded that the junction angles are strongly associated with the climatic condition and can be utilized as an indicator of arid or humid climatic condition.

Based on the results of Seybold et al. (2017), Seybold et al. (2018) calculated the junction angles on the global Earth surface and found that the mean junction angles in each watershed increase with the aridity index. Then, Seybold et al. (2018) compared frequency of junction angles of Mars with that of Mars Desert Research Station (MDRS), which is an analog site located in the Upper Colorado-Dirty Devil basin in the arid southwestern United States, and found similarity between them, which confirms that the climatic environment of Noachian Mars is similar to that of the MDRS. Seybold et al. (2018) also compared the frequency of junction angles of Mars with the frequency of junction angles of the permafrost and nonpermafrost areas in Alaska. The frequency of junction angles of Mars is more similar to that in the nonpermafrost area of Alaska, which suggests that the VNs on Mars were formed by overland flow erosion in nonpermafrost area.

The findings from Seybold et al. (2017 and 2018) suggest that the frequency of junction angle is less influenced by the post-formational modification processes and can be an indicator of the climatic factors of VNs’ formation. In this paper, we compared the associations between the mean junction angles and different environmental factors (climatic and geological factors) on Earth and modeled the associations between the frequency of junction angles and climatic factors ($AI$ and $MAP$). We then applied associations between junction angle and $AI$ and $MAP$ established on Earth to estimate the $AI$ of Mars and scaled the estimated $MAP$ to account for the different radiation each planet receives from the Sun. Next, we estimated
the duration of “warm” Mars by utilizing the minimum cumulative volume of water required to form the VNs (Luo et al., 2017) and the discharge derived from MAP (assuming that 1/3 to 1/2 of MAP runs off and spreads evenly over the watershed or grid). Our results suggest that the climatic condition of ancient Mars was similar to the arid or semiarid area on Earth and that the estimated duration of “warm” Mars supports an “episodically warm” Mars.

### 3.3 Dataset

For the global Mars VN lines, data from two independent sources were considered. The VN lines of Luo and Stepinski (2009) were automatically extracted utilizing a morphology-based method (surface curvature) coupled with flow simulation based on Mars Obiter Laser Altimeter (MOLA) digital elevation model (DEM) data (~463 m resolution, Smith et al., 2001). The VN lines of Hynek et al. (2010) were manually drawn based primarily on interpretation of Thermal Emission Imaging System (THEMIS) mosaic images (~230 m resolution, Christensen et al., 2004) and with reference to MOLA data. The VNs from Hynek et al. (2010) contain more details in high valley density areas because the VNs were drawn from higher resolution images. In low density areas, the VNs from Luo and Stepinski (2009) performed better because automatic algorithm was able to extract valleys based on topography that were not obvious on images due to resurfacing processes.

The terrestrial stream data and hydrological units were obtained from the NHDPlusV2 Dataset (McKay et al., 2012). The streams were derived from the several national topographical maps, with spatial scales comparable with the streams extracted from DEM at a resolution of approximately 30 m. For our basic units of analysis or statistical units, we used the HUC-6 (Hydrologic Unit Code-6) watersheds, small grids (158 km by 158 km), and large grids (500 km by 500 km) (see explanation in section 3.4.3.1).
We used Mean Annual Precipitation \((MAP)\) and Aridity Index \((AI)\) to represent the long term terrestrial climate. The \(AI\) is the index of the generalized climate classification which was defined by the UNEP(1997) and is the ratio of MAP and Mean Annual Evaporation \((MAE)\). All these three data were based on monthly average data from 1950 to 2000 provided by CGIAR-CSI (Zomer et al., 2007; Zomer et al., 2008).

The geological conditions are represented by the surficial lithology and faults. The surficial lithology (Soller et al., 2009), which was based on texture, internal structure, thickness, and environment of deposition is selected to represent the geological factor. Also, some streams, such as those with rectangular drainage pattern, are controlled by the linear geological structure, so we used the faults data to represent the linear structure. The faults data come from the Database of the Geologic Map of North America (Soller et al., 2009). The database includes drainage, geologic units, faults, etc. The database is developed based on the geological map at 1:5,000,000 scale. In the database, there are more than three thousand faults in the conterminous U.S. area.

### 3.4 Method and Procedure

#### 3.4.1 Junction Angle Extraction

The junction angle is the branching angle between two joining stream lines (Figure 10). We extracted the junction angles by utilizing the topological information of the VNs (or streams) stored in the GIS database. We developed a Python program to extract the junction angle automatically in ArcGIS. The Python code has the following steps. (1) It simplifies the VNs lines (Douglas & Peucker, 1973) to obtain the overall direction of VNs near the intersection points and thus removes the bias by local topography (otherwise, we would only
have 45° or 90°). (2) The program searches all the potential junction points and extracts the simplified VN line segments within the circles (buffers’ radius = 10 m) centered on the potential junction points (Figure 10). The numbers of the starting vertices and ending vertices in each circle are counted; if there are two ending vertices and one starting vertex at the same location, which means that two upstream VNs merge to a down-stream VN, the center point is a real junction point. (3) The program then calculates the junction angle (α) based the area of the slice formed by the two upstream VNs and the circle (S) and the area of the whole circle (A): \[ \alpha = \frac{S \times 360^\circ}{A}. \]

![Diagram illustrating VN junction angle calculation based on VN topology.](image)

Figure 10: VN junction angle calculation based on VN topology
Diagram illustrating VN junction angle calculation based on VN topology information. (Arrows represent the overall directions of VNs after simplification to remove bias by local topography.)

Because the method described above relies on topological information to extract junction angles, accurate topological information in the VN/stream data is critical. Since the VNs from Luo and Stepinski (2009) and streams from NHDPlus v2 are automatically extracted from DEM, their direction follows topography and the topological information is accurate. The VNs from Hynek et al. (2010) are manually drawn from the remotely sensed images, so some VNs’ direction may not always follow the topography. To assess the data quality of
juncture angles, we compared the junction angles obtained from the two datasets in Table 5.

The inconsistency in mean junction angles between the two results is due to the inaccurate topology of some VNs in the manually derived Hynek et al. (2010) data. This is confirmed by the abnormal high frequency of large junction angles and by manual check. The frequency of large angles (greater than 120°) from Hynek et al. (2010) is abnormally high (see Table 5). The range of frequency of junction angles which are greater than 120° is from 6.18% to 9.75% on Earth. Our manual checks also confirmed that some VNs’ flow directions of Hynek et al. (2010) are not correct since they are manually digitalized. For this reason, we will only report the result based on the VNs data from Luo and Stepinski (2009) for the rest of the study.

Table 5: Basic statistical result of Mars junction angles

<table>
<thead>
<tr>
<th></th>
<th>Mean of junction angle</th>
<th>No. of junction points</th>
<th>No. of junction angles &gt;120°</th>
<th>Mean of junction angles &lt;120°</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luo and Stepinski</td>
<td>58.39°</td>
<td>28,620</td>
<td>848 (≈ 2.96%)</td>
<td>56.49°</td>
</tr>
<tr>
<td>Hynek et al.</td>
<td>68.15°</td>
<td>22,659</td>
<td>2,832 (≈ 12.63%)</td>
<td>57.38°</td>
</tr>
<tr>
<td>All terrestrial (U.S.) data</td>
<td>73.31°</td>
<td>867,804</td>
<td>77,322 (≈ 8.91%)</td>
<td>66.71°</td>
</tr>
<tr>
<td>Terrestrial humid area</td>
<td>76.13°</td>
<td>157,666</td>
<td>5,356 (≈ 9.74%)</td>
<td>69.23°</td>
</tr>
<tr>
<td>Terrestrial arid area</td>
<td>62.91°</td>
<td>178,143</td>
<td>11,005 (≈ 6.18%)</td>
<td>58.81°</td>
</tr>
</tbody>
</table>

The averaged AI by watersheds are classified to five classed by the quantile classification method.
The humid area is the watersheds whose AI are greater than 1.034.
The arid areas’ AI are smaller than 0.382.
3.4.2 Spatial Association Measures by SPADE

To compare the associations between junction angle and environmental conditions, we applied the SPatial Association DEtector (SPADE) (Cang & Luo, 2018) to assess the associations. The SPADE is an improvement of the Geo-detector (Wang et al., 2016), which assesses the association by analyzing the variance. The SPADE solved the two main drawbacks of the Geo-detector, which does not consider distance decay and is influenced by the number of discretization zones for continuous variables. After we improve the Geo-detector, the SPADE is a spatial method considering the distance and minimize the influence of discretization so that it is applicable to both continuous and discrete variables. The association is represented by an association parameter usually ranging between 0 and 1, with 0 representing no association and 1 perfect association. The details of SPADE can be found in Cang and Luo (2018).

3.4.3 Comparative Analysis Procedure

3.4.3.1 Terrestrial Data Analysis

For terrestrial analysis, we selected streams in the conterminous U.S. as our terrestrial data because the conterminous U.S. include varied climates types and have detailed stream information and the geological and climatic database. To compare the influence of climatic and geological factors, we set the mean junction angle as the dependent variable and use the environmental factors (climatic factors: climate types, $AI$, and $MAP$ and geologic factors: lithology and density of faults) as the independent variables respectively, then used the SPADE to investigate their associations and ranking.
The dependent variable, mean junction angles, is the averaged junction angles by statistical units, which are HUC-6 (Hydrologic Unit Code-6) watershed, small grids (158 km by 158 km) or large grids (500 km by 500 km). To model the associations between junction angles and climatic factors, we averaged the AI and MAP by the same statistical units as the dependent variables. To assess the associations between junction angles and geological factors, we used zonal majority of surficial lithology by the zones of same statistical units and calculated density (m/km$^2$) of faults (representing the power of linear geological structure) by the same statistical units as the dependent variables. Since the control of geologic structure on streams is often reflected in the rectangular drainage pattern with mostly right junction angles, we also used the ratio of right junction angle in statistical units to represent the junction angle.

For the statistical unit, we used not only the HUC-6 watersheds, but also square grids (with each grid area equaling to the mean area of the HUC-6 or equaling to the grids used on Mars data) to process our data. The watershed boundaries on Mars are often poorly defined due to the influence of impact cratering and the low density of VNs on Mars, so it is desirable to use a simple unit such as square grid. To obtain enough junction angles in each grid, the size of statistic unit on Mars is assigned as 500 km by 500 km. To compare the difference between the different sizes and shapes, we chose all the three statistical units for the measures of terrestrial association and estimates of Mars climatic conditions.

3.4.3.2 Terrestrial Data Modeling

For the purpose of applying the associations between climatic factors and junction angles to Mars, we used the Ordinary Least Squares (OLS) regression and $k$-NN regression ($k$-
nearest neighbor regression) to model to the associations between the climatic factors and the frequency of junction angles.

We selected the OLS regression because the previous research Seybold et al. (2017, 2018) indicated that the association between the logarithmic AI and junction angles may be close to linear, but Seybold et al. (2017, 2018) did not prove that. To consider the possibility that the association between junction angles and climatic factors is nonlinear, we applied \( k \)-NN regression to model the association between the frequency of junction angle and environmental factors. The OLS regression and the \( k \)-NN regression are implemented by Pedregosa et al. (2011).

The dependent variables in the regression, or the training data in the \( k \)-NN regression, are the logarithmic climatic factors (AI and MAP) and the independent variables are the frequencies of junction angle in bins (bin=10°). The reason for selecting frequency distribution of junction angle, as opposed to the mean junction angles, is that the frequency distribution contains more information than the simple mean. The modeled associations were utilized in the section 3.4.3.3 to estimate the AI and MAP of the Noachian Mars.

### 3.4.3.3 Inferring Mars Climatic Conditions

We first compared the junction angles on Mars and Earth. The comparisons are made between the frequency distribution of all junction angles considered on Mars surface (hereafter simply referred to as global Mars) and frequency distributions of junction angles in the conterminous U.S. We used quantile classification method to group the HUC-6 watersheds to five groups based on their AIs. Then we calculated the frequency distribution of junction angles in each group and compared them with frequency distribution of Mars junction angle.
We also applied the model results from the OLS regression and $k$-NN regression obtained on Earth to the global frequency distribution of Mars junction angles to estimate the global $AI$ and $MAP$ of Mars. Note, because the radiation each planet received from the Sun is different, the estimated $MAP$ from the model does not represent the MAP on the Noachian Mars. To obtain the $MAP$ on the Noachian Mars, we need to consider the difference in solar radiation each planet receives from the Sun, which can be calculated as the ratio of MAE between the Mars and Earth. This ratio is then used to scale the $MAP$ estimated from the model. The scale processing is described in section 3.4.3.4.

To consider the spatial distribution of climate on Noachian Mars, we divided the entire Mars surface to grids, each with size of 500 km by 500 km, which can be considered a climatically homogeneous area. We selected the grids within denser junction angle belt for analysis. After we grouped the junction angles to each grid, we extracted the frequencies of junction angles in each grid and applied the OLS regression and $k$-NN regression models derived from terrestrial data, to estimate the $AI$ and $MAP$ of each grid on Mars. Similar to global frequency analysis, we also used the ratio of MAE between the Mars and Earth to scale the estimated $MAP$ of each grid on the Noachian Mars.

3.4.3.4 Estimating the Duration of “Warm” Mars

To consider the difference in solar radiation Earth and Mars received from the Sun, we use the Potential Evaporation ($PE$) to represent the MAE. The replacement is also used in the Earth climatic data production (Zomer et al., 2007; Zomer et al., 2008). The PE on Earth can be related to radiation by Hargreaves evapotranspiration equation (Hargreaves & Allen, 2003) as Equation 3.1.
\[ PE = 0.0023 \times RA(T_{\text{mean}} + 17.8) \times TD^{0.5} \text{ mm/month} \]  

(3.1)

where \( T_{\text{mean}} \) is mean temperature; \( TD \) represents daily temperature range; and \( RA \) is radiation received from the Sun.

Due to the faint young Sun and the distance between Mars and the Sun, Mars only receives 1/3 of solar radiation that Earth receives. The mean temperature and daily range have more uncertainty. At the lower end, we assume \( T_{\text{mean}}=0^\circ \text{C} \) and \( TD=22.2^\circ \text{C} \) (same as terrestrial value), which will make Mars PE about 1/5 of terrestrial value. At the upper end, we assume \( T_{\text{mean}}=9.75^\circ \text{C} \) (Wordsworth et al., 2015) and \( TD=60^\circ \text{C} \) (based on Viking record (Lewis et al., 1999)), which would give Martian PE about 1/2 of terrestrial value. We assume these same ratios in PE estimates also apply to MAP estimates when AIs are the same.

Next, we utilized the frequency of junction angles of global VNs to estimate the duration of “warm” Mars. We divided volume of water required to form the VNs (Luo & Stepinski, 2009) by the discharge (assuming that 1/3 to 1/2 of MAP runs off and spreads evenly over grid) to obtain the VNs’ formation timescale. To estimate the duration of the “warm” Mars, we adopted a 1% intermittence (which is the average time between floods) because our estimated the AI of Noachian Mars falls in the zone of semi-arid or arid, for which the intermittence is 1% (Alemanno et al., 2018).

To infer local/regional climate at grid level, we only selected the upstream grids (shown as hatched areas in Figure 12) to estimate the duration of “warm” Mars to avoid the complication that water flowed in the downstream grids were not only from the precipitation, but also from upstream areas.
3.5 Results

3.5.1 Terrestrial Junction Angle Analysis

There are a total of 867,804 junction angles extracted from the conterminous U.S. The mean and standard deviation of the junction angle are 73.31°, 32.98° respectively. The associations between the junction angles and the environmental factors using SPADE are shown in Table 6(a). The associations between the linear geological structure and the junction angles using SPADE are in Table 6(b).

Table 6: Associations using SPADE (larger value means higher association)

(a) Associations between mean junction angles and environmental factors on Earth

<table>
<thead>
<tr>
<th></th>
<th>Lithology</th>
<th>Density of faults</th>
<th>Climate types</th>
<th>Aridity index</th>
<th>Mean annual precipitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>HUC-6</td>
<td>0.33</td>
<td>0.14</td>
<td>0.38</td>
<td>0.46</td>
<td>0.39</td>
</tr>
<tr>
<td>158km × 158 km</td>
<td>0.41</td>
<td>0.15</td>
<td>0.39</td>
<td>0.53</td>
<td>0.47</td>
</tr>
<tr>
<td>Mean junction angle</td>
<td>0.27</td>
<td>0.06</td>
<td>0.26</td>
<td>0.63</td>
<td>0.38</td>
</tr>
<tr>
<td>500 km × 500 km</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Associations between density of faults and junction angle and between density of faults and ratio of right angles

<table>
<thead>
<tr>
<th>Density of faults</th>
<th>Mean junction angle</th>
<th>Ratio of right junction angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>HUC-6</td>
<td>0.08</td>
<td>0.14</td>
</tr>
<tr>
<td>158 km × 158 km</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>500 km × 500 km</td>
<td>-0.03</td>
<td>0.02</td>
</tr>
</tbody>
</table>

In Table 6(a), the influences of climatic types and lithology vary under different size and shape of statistical units. Although we cannot prove which one is more important, we find
that the lithology and climatic types both have significant influence on the junction angles, and influence of climatic types is stronger than or at least on a par with the influence of lithology; so the climatic influence cannot be ignored. The results in Table 6(b) showed that the associations between the linear geologic structures (faults) and the junction angles are weak at the conterminous U.S. scale. Although the linear structures can influence the junction angles at the fine scales and form the rectangular drainage pattern, the influence from the surficial lithology is a majority factor at the conterminous U.S. scale, so we do not consider the linear structure as an influence factor in this research. Based on this result, we confirmed that the junction angles are influenced by the surficial lithology and climatic factors. Since this research focuses on the influence of water to the junction angles, we investigated the association between the junction angles and $AI$ and $MAP$. Their associations are stronger than the lithology and climatic types, which means that $AI$ and $MAP$ are more predictable by the junction angles.

Table 7 shows the performance scores of OLS regression and $k$-NN regression models (higher score representing better performance). To ensure the robustness of the modeling results, we assigned $k =1$, 10, and 50 for HUC-6 watershed and grids with size of 158 km by 158 km and assigned $k =1$, 3, and 5 for grids with size of 500 km by 500 km because the total number of grids in this large size is 47, which is much fewer than the number of small grids (more than 300). Most scores are greater than 0.6. In the $k$-NN regression, the model scores, when $k$ is small, are greater than the OLS regression model, although it may have some overfitting risks. The model scores of $k$-NN regression decrease with the increase of $k$ values. When the $k$ is large enough (such as $k =50$ for the small statistical units or $k =5$ for the large statistical units), the scores of $k$-NN regression are smaller than the scores of OLS regressions.

The scores from the grids with size of 500 km by 500 km are larger than the scores from other statistical units, because the large grids may group more details together. However,
Table 7: Model performance on terrestrial data.

(a) Model based on the HUC-6

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>$k$-NN($k=1$)</th>
<th>$k$-NN($k=10$)</th>
<th>$k$-NN($k=50$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>logAI</td>
<td>0.59</td>
<td>1.00</td>
<td>0.67</td>
<td>0.57</td>
</tr>
<tr>
<td>logMAP</td>
<td>0.49</td>
<td>1.00</td>
<td>0.61</td>
<td>0.48</td>
</tr>
</tbody>
</table>

(b) Model based on the grids with size of 158km × 158km

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>$k$-NN($k=1$)</th>
<th>$k$-NN($k=10$)</th>
<th>$k$-NN($k=50$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>logAI</td>
<td>0.65</td>
<td>1.00</td>
<td>0.69</td>
<td>0.62</td>
</tr>
<tr>
<td>logMAP</td>
<td>0.55</td>
<td>1.00</td>
<td>0.62</td>
<td>0.53</td>
</tr>
</tbody>
</table>

(c) Model based on the grids with size of 500 km × 500 km

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>$k$-NN($k=1$)</th>
<th>$k$-NN($k=3$)</th>
<th>$k$-NN($k=5$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>logAI</td>
<td>0.59</td>
<td>1.00</td>
<td>0.86</td>
<td>0.80</td>
</tr>
<tr>
<td>logMAP</td>
<td>0.49</td>
<td>1.00</td>
<td>0.75</td>
<td>0.66</td>
</tr>
</tbody>
</table>

it does not mean that the large grids are much better than other smaller statistical units due to the complexity of spatial self-organization and spatial scale. For example, if the all junction angles are grouped to one group, the model score is 1 (high score, but meaningless). We did not compare which model is the best in detail. We chose the three statistical units and took advantage of the OLS regression model and the $k$-NN regression models excluding the large $k$ to estimate the climatic condition of Mars.

### 3.5.2 Global and Local Mars VNs Analysis

Based on the terrestrial junction angle analysis, the frequency distribution of junction angles acts as an indicator of the longterm climate. We extracted the global junction angles of Mars and plotted the frequency distribution of global Mars junction angles and terrestrial junction angles (statistical unit: HUC-6) together. Figure 11 shows that the frequency
Figure 11: Frequency distribution of junction angles
Frequency distribution of junction angles on the global and in different terrestrial zones of Earth. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

The frequency distribution of global Mars junction angles is generally between the arid \((AI < 0.38)\) and semi-arid \((0.38 < AI < 0.68)\) areas in the conterminous U.S. It indicates that climate of ancient Mars was wetter than the driest areas of U.S.

To analyze the spatial distribution of estimated climatic conditions, we divided the entire Mars surface to square grids of 500 km by 500 km and then calculated the frequency of junction angels within each grid and estimated \(AI\) of each grid. To minimize the influence of post-formation process, we only select the grids located in the dense VN's belt (Figure 12).

The spatial distribution of mean junction angles of each grid is shown in Figure 12. The general trend is that the junction angles in the north are larger than those in the south, which indicates that the northern areas were wetter than the southern areas (Seybold et al., 2017).
We also analyzed the spatial distribution of estimated $AI$ by calculating the association between $AI$ and coordinates (latitude and longitude) of grid centers. To avoid the influence of non-linear association between variables, we chose Spearman’s rank correlation coefficient (Spearman’s rho) which assumes that a perfect monotone function reaches the highest correlation (+1 or —1) if there are no repeated data values. The Spearman’s rho between the junction angles and the latitude (to make interpretation easier, the latitude of north pole is defined as $0^\circ$ and south pole $180^\circ$) is $-0.37$ (see also Table 9(a)), and the associations between the junction angles and the longitude is not significant.
Estimated Climatic Conditions and Duration of “Warm” Period

We applied the association between junction angle and AI learnt from the OLS regression and $k$-NN regression models on Earth to estimate the AI of the ancient Mars. Using the global frequency of junction angles of Mars, the range of estimated AI from different models is from 0.18 to 0.47, the median and mean are 0.35 and 0.32, respectively. We also applied the associations to the junction angles grouped by grids, and the statistics of the results based on grids are shown in Figure 13.

Figure 13: Box and whisker chart of the estimated AI Box and whisker chart of the estimated AI of Noachian Mars by models. The bottom and top of box represent first quartile and third quartile. The median is represented by line in the box. The mean is denoted by the “X” marker. Followed the Tukey industry standard, the maximum length form whisker to its nearest box is 1.5 times the length of the box. The box and whisker of large grids (OLS) includes too many outliers, so a part of the box is shown.
After we estimated the MAP of the global Mars by utilizing the terrestrial models and scaled MAP based on the estimated ratio of MAE between Earth and Mars, we calculated the formation timescale of VNs by dividing the minimum volume of water required (Luo et al., 2017) by discharge derived from MAP (assuming that 1/2 to 1/3 of scaled MAP runs off and spreads evenly over the area of the grids analyzed, in m$^3$/yr). Then, we divided the formation timescale by 1% (intermittence) to obtain the “warm” period duration. The results are shown in Table 8. We excluded the results from OLS regression (large grids). The reason is described in the section 3.6.1. The range of “warm” period duration is from 7.8(4) to 77.(05) million years.

Table 8: Estimated duration of “warm” Mars (unit: million years)

<table>
<thead>
<tr>
<th>HUC-6</th>
<th>158km × 158km</th>
<th>500km × 500km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>KNN_1</td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>7.8(4)</td>
<td>13.84</td>
</tr>
<tr>
<td>UL</td>
<td>32.68</td>
<td>57.67</td>
</tr>
<tr>
<td>Upstream</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LL</td>
<td>4.4(3)</td>
<td>7.4(1)</td>
</tr>
<tr>
<td>UL</td>
<td>18.45</td>
<td>30.87</td>
</tr>
</tbody>
</table>

We also estimated and scaled the local MAP by utilizing the frequency of junction angles and the terrestrial models. To avoid the complication that downstream areas receive water from precipitation and contribution from upstream and the fact that VNs at lower elevation are more likely influenced by nonfluvial processes (e.g., groundwater sapping), we only considered upstream grids (shown as hatched areas in Figure 12). Then, we calculated the mean “warm” period duration as the results, which are shown in Table 13. The duration of “warm” period is from 4.4(3) ∼ 34.81 million years.
Table 9: Spatial analysis of landform characteristics and estimated AI.

(a) Spearman’s rho matrix of landform characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean junction angle</th>
<th>Mean elevation</th>
<th>Latitude (0 N pole ~ 180°S pole)</th>
<th>Minimum water volume required to form VNs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean junction angle</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean elevation</td>
<td>-0.24**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Latitude (0 N pole ~ 180°S pole)</td>
<td>-0.37***</td>
<td>0.32***</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Minimum water volume required to form VNs</td>
<td>0.29***</td>
<td>-0.34***</td>
<td>-0.39***</td>
<td>-</td>
</tr>
</tbody>
</table>

(b) Correlation between AI estimated by different models and landform characteristics

<table>
<thead>
<tr>
<th></th>
<th>HUC-6 158 km</th>
<th>HUC-6 500 km</th>
<th>158 km</th>
<th>500 km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude (0°N pole ~ 180°S pole)</td>
<td>-0.25***</td>
<td>-0.30***</td>
<td>-0.40***</td>
<td>-0.31***</td>
</tr>
<tr>
<td>Mean elevation</td>
<td>-0.12(ns)</td>
<td>-0.22**</td>
<td>-0.25***</td>
<td>-0.07(ns)</td>
</tr>
<tr>
<td>Minimum water volume required to form VNs</td>
<td>0.20**</td>
<td>0.16*</td>
<td>0.24**</td>
<td>0.20**</td>
</tr>
</tbody>
</table>

ns: P>0.05; *: P ≤0.05; **: P≤0.01; ***: P≤0.001.

3.6 Discussion and Conclusion

3.6.1 Terrestrial Model Selection

We excluded the results from OLS model (large grids), because the estimated Mars climatic conditions from this model are abnormal and with many outliers. The estimated
AI and the duration of “warm” Mars are much greater than the results from other models. Some grids’ (12 of 39 upstream grids) have AI values greater than 1.5. Recall that AI is the ratio between precipitation and potential evapotranspiration, a value of 1.5 means that precipitation is 1.5 times of PE, a very humid condition, which were unlikely for Mars. Meanwhile, the results have the highest AI and the long duration at the same time, which is not self-consistent and are likely caused by the outliers. Usually, the higher the AI, the higher the MAP; and the higher MAP will result in the shorter formation timescale. In the other two statistical units, the estimated AIs from OLS regression are larger than the AIs from k-NN regression and the estimated formation timescales from OLS regression are shorter. This shows that the size relationship between the estimated AI and timescale is self-consistent.

### 3.6.2 Estimated Mars Climatic Conditions

The estimates of the Mars AI are consistent with previous research results by Matsubara et al. (2011 and 2013). Matsubara et al. (2011 and 2013) used the X-ratio to express climatic conditions. Landform simulation modeling has established that the range of X-ratio is from 1 to 7 (Matsubara et al., 2011; Matsubara et al., 2013). The X-ratio is defined as Equation 3.2.

\[
X = \frac{E - P}{PR_B}
\]  

(3.2)

where \(E\) is evaporation; \(P\) is precipitation; \(R_B\) is the fraction of precipitation that contributes to runoff. Higher X-ratio represents drier climate.

Based on the definition of X-ratio and AI, the two can be related to AI by the following equation 3.3.
\[ AI = \frac{1}{1 + XR_B} \]  

The range of \( R_B \) is from 0 to 1, so the range of \( AI \) is from 0.125 to 1 (considering X-ratio ranges from 1 to 7). In our research, we assume the range of \( R_B \) is from 0.33 to 0.50 (Trenberth et al., 2007). Following our assumption, the range of equivalent \( AI \) is from 0.22 to 0.75. The estimated global \( AIs \), which is calculated from the frequency distribution of global junction angle, is within this range.

Seybold et al. (2018) revealed that the climatic condition of Noachian Mars was similar to the MDRS because they have similar drainage patterns, suggesting that the basin and its surroundings are similar to Martian landscape and that’s why it is selected as an analog location to train astronauts. The MDRS is located at the Upper Colorado-Dirty Devil basin (HUC 1407), which is at (-111.092530°W, 37.222290°N). The mean \( AI \) of HUC 1407 watershed is 0.20 (Zomer et al., 2007; Zomer et al., 2008). The minimum, median and maximum \( AI \) estimated by the global frequency of junction angles of Mars is 0.18, 0.32 and 0.46, respectively. The \( AI \) of HUC 1407 watershed is within the range of our estimated \( AI \). Although our estimated \( AIs \) (0.18-0.47) are a little bit greater than the \( AI \) of HUC 1407 watershed, the HUC 1407 watershed was wetter during the late Pleistocene (Matsubara & Howard, 2009; Matsubara et al., 2011; Matsubara et al., 2013) when the landform of HUC 1407 watershed formed. The similar \( AIs \) showed that our estimated results are reasonable.

The spatial pattern of estimated AIs also supports that our estimated conditions are reasonable. No matter whether ancient Mars was “warm” or “cold”, the liquid water flowed on the surface of Mars (Wordsworth et al., 2018) in the Noachian period and accumulated in the northern basin. To test the hypothesis of Wordsworth et al. (2018), we calculated the Spearman’s Rho matrix of landform characteristics (mean junction angle, mean elevation, latitude, minimum water volume required to form VNs) (Table 9(a)). The results suggest
that the lower the elevation and the farther north one goes, the larger the junction angle and the more the water required to carve the VNs, which is consistent with the scenario of an integrated hydrologic cycle with water flowing from south to north. In this case, the northern area should have more water due to the accumulative effect. If this hypothesis is correct, the estimated $AIs$ should have positive correlation with the volume of required water to form VNs and have negative correlation with the distance of grid to north pole. To test whether the spatial distribution of our estimated $AIs$ is consistent with the spatial distribution of volume of required water to form VNs, we calculate the Spearman’s Rho between the estimated $AI$ and center latitude, mean elevation, and volume of required water to form VNs by grid, respectively. The results are showed in Table 9(b).

In Table 9(b), the estimated $AIs$ have significant positive correlation with the volume of water required to form VNs (Luo et al., 2017) across all models, consistent with wetter area supplying more water for erosion processes. Also, most estimated $AIs$ have significant negative correlation with the latitudes (the values of latitude from north pole to south pole is defined as $0^\circ \sim 180^\circ$). This shows that the northern areas are wetter than the southern areas.

Since more fluvial erosion needs more water, our estimated $AIs$, derived from junction angle frequencies, are consistent with the volume of water required to form VNs, independently derived from DEM (Luo et al., 2017). Both the values of estimated $AIs$ and the volume of water required to form VNs decrease from northern basin to southern highland. It indicates that the water flowed from southern highland to northern basin and accumulated more water downstream. The trend supports the hypothesis that the Mars had a global hydrological cycle and the surface water flowed from southern highland to the northern basin. This consistency indicates that the spatial trend of our estimated $AI$ is reasonable.
3.6.3 Estimated Duration of “Warm” Period

Our estimated duration of “warm” period using all VNs (with an upper limit of around 77 million years) is longer than the duration using upstream areas only (4.4(3) to 34.(81) million years). Because VNs at lower elevation are more likely to be influenced by nonfluvial processes (e.g., groundwater sapping), we prefer the estimate of duration using upstream grids (4.4(3) to 34.(81) million years). It is important to note that our estimated length of “warm” period was obtained by integrating the climatic model and the geomorphological evidence. The estimated duration using all VNs and estimated duration using only upstream VNs have the same order of magnitude, but the former is usually larger than the latter by a factor of two.

The intermittence values depend on the aridity of Mars. If the Mars climatic condition was hyperarid, the intermittence value is 0.1% (Orofino et al., 2018); if the Mars was semi-arid/arid, the intermittence value is 1% (Orofino et al., 2018); and if Mars climatic condition was humid/sub-humid, the intermittence is 5% (Hoke et al., 2011). We took the intermittence value for semi-arid/arid for Mars, because the range of our estimated AIs falls between the zones of arid and sub-arid. Moreover, previous research relating the climate condition of Noachian Mars (Matsubara et al., 2011; Matsubara et al., 2013; Seybold et al., 2018) have the same or similar conclusion.

The duration of “warm” Mars is also estimated by other researchers. Hoke et al. (2011) applied Darcy–Weisbach equation and sediment transport models to the landform characteristics of large VNs to estimate the formation timescale. In Hoke et al. (2011)’s research, the concept of intermittence formation timescale is similar to the duration of “warm” Mars in our paper. The typical range of intermittence formation timescale of Hoke et al. (2011)
is between 0.1 and 100 million years. The range of our estimated duration of “warm” Mars is consistent with their results, but we have a narrower range.

Orofino et al. (2018) took advantage of modified Manning’s equation and the VNs’ landform characteristics to estimate the discharge. They then divided the volume of water required by the estimated discharge to estimate the possible range of water flow duration, which is considered as the duration of “warm” Mars in our paper. The range of water flow duration is 0.01 to 200 million years (median is 0.5 million years). Our results are within this range and is better constrained to a narrower range.

3.6.4 Additional Uncertainties from Channel Bed Conditions and Orbital Obliquity

The volume of water required to form the VNs was the minimum cumulative volume globally (Luo et al., 2017). Thus the duration of the erosion we estimated in this study should also be considered the lower bound. In addition, if the VNs beds are floored by coarse gravels/boulders or armored by coarse sediment (Howard et al., 2016) or at least locally composed of bedrock, erosion rates could be dramatically slower Sklar and Dietrich, 2004 and the formation duration could be significantly longer than our results in Table 8.

In addition, since the climate of Mars was controlled by obliquity variation, the “warm” Mars might only occur a small fraction of the time during the obliquity cycles (Mischna et al., 2013). If this hypothesis is correct, the obliquity cycle gave an additional intermittency on top of the intermittency of runoff events. The total duration of VNs’ formation (from the first to the last runoff incisions of VNs) could have been dramatically longer than even the least conservative published estimates, but the effective incision duration would still be similar to our estimates.
3.6.5 Conclusion

Stream/VN junction angles offer an important means to uncover processes and climate that formed the streams/VNs, because the angles and their spatial and frequency distribution contain fingerprints left by these processes and climatic conditions. Furthermore, the junction angle is a robust property that is minimally influenced by post formational modification processes.

Our terrestrial analysis in the conterminous U.S. confirmed that surficial environmental factors (lithology and climate type) significantly correlated with the junction angles and that the $AI$ and $MAP$ can be predicted by frequency of junction angles. The frequency of global junctions on Mars allows us to estimate the climatic condition ($AI$ and $MAP$) of the Noachian Mars. The range of estimated $AI$s using all junction angles is from 0.18 to 0.47, which is consistent with previous research using other landform characteristics. The duration of “warm” Mars can be estimated by the discharge derived from scaled MAP and water volume required to form VNs. The range of estimated duration of “warm” Mars using all junction angles is from 7.8(4) to 77.(05) million years; the range of estimated duration of “warm” Mars using upstream junction angles is from 4.4(3) to 34.(81) million years. Although the estimates using all VNs are larger than the estimates using upstream VNs by about a factor of two, they are in the same order of magnitude.

With junction angle as a robust and reliable property that reflects past climate and processes as shown in our terrestrial analysis, we are confident with our interpretation of the junction angle frequency on Mars. The global junction angles frequency on Mars supports that the $AI$ of Noachian Mars is similar to those of the arid/sub-arid regions on Earth. The correlation with water volume required to form VN and latitude results suggest that our estimated climatic conditions are consistent with the scenario of an integrated hydrologic
cycle that water flows from south to north. It is important to note that our estimated length of “warm” period was obtained by integrating the climatic model and the geomorphological evidence, although it should be considered the lower bound due to uncertainties in channel bed conditions and orbital obliquity. Considering the Noachian period is 400 million years, both results support the hypothesis that Mars was “ episodically warm”.
CHAPTER 4

THE MARS WATERSHEDS’ MATURITY MEASURED BY
OPTIMAL CHANNEL NETWORKS SIMULATION

4.1 Abstract

The dendritic Valley Networks (VNs) are distributed widely on the Martian Noachian surface. They offer convincing evidence for the past surface fluvial processes on Mars and indicate that the early Mars was warm enough to have liquid water on its surface. The duration and intensity of early Martian surface fluvial processes can provide critical information in understanding the duration of the “warm period” of early Mars. Previous research empirically and theoretically suggests that through the landform evolution processes, the energy dissipated along an optimal channel network (OCN) is minimal. That theory provides an opportunity to investigate the VNs’ maturity, which is related to the length of “warm period” duration, by computing the ratio between the optimum energy dissipation along simulated OCN and the energy dissipation in Martian fluvial landform (maturity is a number between 0 and 1, 1 being most mature). I first analyzed the watersheds’ maturity on Earth. The results suggested that the watershed maturities in the most terrestrial areas are greater than 0.86. Also, environmental factors and the watershed maturity have a stronger association in the arid/semi-arid area than the associations between the watershed maturity and environmental factors in the humid area. The Martian watersheds’ maturity showed that most Martian fluvial landform areas are immature compared to the fluvial landform areas on Earth. After grouping the watersheds’ maturity by using dendritic VNs, I found
that the maturity of many large VNs are similar to the watersheds’ maturity in the arid or semi-arid arid area on Earth. The results suggest that ancient Mars was at least “episodically warm” and the surface runoff and fluvial erosion processes operated long enough to carve the Martian surface in some regions.

4.2 Introduction

Mars is dry and cold now, and cannot have liquid water on its surface. However, the thousands of VNs, which were carved into the Martian southern highlands, indicate that environmental conditions were “warm” enough to support liquid water flowing on its surface. Also, the deltas (Di Achille & Hynek, 2010) and tsunami deposits (Rodriguez et al., 2016) suggest that liquid water flowed on the early Martian surface. Although many lines of geomorphic and geologic evidence suggest an early “warm” Mars with an active hydrologic cycle (Luo et al., 2017), Martian climate modelers have had difficulties in modeling such warm conditions, but could model a “cold” Mars, whose annual mean temperature was below 273 K. The main reasons are the Martian orbit and the faint young Sun (Wordsworth et al., 2015). Since Mars is in an orbit farther away from the Sun than Earth, it only receives about 43% of the solar energy that Earth does and the luminosity of Sun was only 75% of its present value (Gough, 1981). Under the “cold” Mars scenario, the VNs were formed in the episodic warm intervals when peak reducing gas release rates and background CO₂ levels are high enough (Wordsworth et al., 2021).

Understanding the formation of Mars VNs is important in understanding the early Mars climate. The characteristics of Martian fluvial landform have been investigated quantitatively to estimate the environmental conditions of the Noachian period (Hynek & Phillips, 2003; Penido et al., 2013; Williams & Phillips, 2001). The results of previous studies are con-
sistent with the “warm” Mars scenario and support that liquid water flowed on the Noachian Martian surface; however, these research projects did not answer the questions of how long the “warm” Mars lasted and at what timescale the VNs were formed, so their results cannot confirm or exclude the early Martian “warm” scenario.

The duration of “warm” Mars is estimated by applying the sediment transport models and the characteristic of Martian VNs (junction angles, depth, width, volume, etc.) (Cang & Luo, 2019; Hoke et al., 2011; Orofino et al., 2018). Since the large uncertainties existed in the extraction of the VN characteristics and the models are different, the conclusions from these studies are different or even mutually exclusive. Galofre, Bahia, et al. (2020) investigated the maturity of VNs by modeling and analogizing the Earth’s and Mars’ longitudinal profiles and detected several mature VNs, which had surface runoff for long enough time to alter the landscape and reach the steady state. Their research did not estimate the absolute “warm” period length, but provided the relative maturity ranking, which can be used as the geologic constraints for investigating the early Martian climate. Additionally, since the ancient Martian fluvial landforms are modified by the post-formation processes, such as eolian processes, impact cratering, and true polar wander, some Martian longitudinal profiles are irregular and cannot be modeled by the method in Galofre, Bahia, et al. (2020). Because the number of measures provided by Galofre, Bahia, et al. (2020) is limited, it cannot provide a full picture of early Mars climatic conditions.

Hitherto, scientists are still debating on whether episodic warmth or non-precipitation dominated erosion can create the observed valley networks. Although the debates are still going on, the consensus between supporters of “Warm” Mars and “Cold” Mars is that abundant liquid water flowed on Mars in the late Noachian period (Ramirez & Craddock, 2018; Wordsworth et al., 2018).

Since characteristics of VNs used by previous research projects were likely influenced by the impact cratering and post-formation processes, it is necessary to choose a robust
characteristic that is not or less influenced by the impact cratering and the post-formation processes to estimate the climatic condition of the Noachian Mars. Cang and Luo (2019) estimated the aridity of the early Mars and duration of “warm” Mars by utilizing the frequency distribution of junction angles. Although the junction angle is a robust characteristic, the spatial resolution for the aridity estimation is very coarse (500 km * 500 km). The reason is that the estimating function needed enough samples for the frequency of junction angles in the area and the Martian VNs are in general much sparser than the Earth’s streams.

In the fluvial landform, the DEM cells’ hierarchy relation is the basic characteristic of the fluvial landform. According to the view of fluvial geomorphology, the hierarchy relation is the topological relation for describing the aggregating process of flowing water. Normally, the hierarchy relation can be represented by the tree-like network whose root is the outlet and leaves are the boundary of watershed. Most of the watershed analysis methods, such as Hack’s law (the power law relating stream length and contributing area) (Rigon et al., 1996), utilized the cells’ hierarchy relation to the exponent of the Hack’s law. Those methods have been widely used in the Earth hydrology research to understand the fluvial processes and their formation conditions (Singh et al., 2014).

On Mars, impact cratering created sinks on its surface, so the hierarchy relation extraction and the watershed boundary delineation become big challenges for the Martian watershed analysis. Although it is hard to determine the hierarchy relation and the watershed boundary, Mest et al. (2010) showed that the hierarchy relation and the watershed boundary can be corrected partly by the DEM filling function because most post-Noachian impact cratering did not change the general trend of the topography in the large area. For this reason, the filled DEM can generally represent the hierarchy relation in the watershed.

Based on the previous research (Rodriguez-Iturbe & Rinaldo, 2001; Rodriguez-Iturbe et al., 1992), the hierarchy relation in the watershed is modified by the self-organization process, which minimizes the energy dissipation by modifying the flow direction spontaneously.
The watershed with the minimum energy dissipation, which is calculated from its hierarchy relation, is called the OCN status. Based on the definition of OCN (Rodriguez-Iturbe et al., 1992), the OCN follows three principles: 1) the principle of minimum energy dissipation in any link of the network for the transportation of a given discharge; 2) the principle of equal energy dissipation per unit area of channel anywhere in the network; and 3) the principle of minimum energy dissipation in the network as a whole. Ijjasz-Vasquez et al. (1993) compared the energy dissipation between the OCN and real world watersheds. The results showed that the energy dissipation are similar between the real watershed and theoretical OCN model. It indicates that OCN channel from the model could represent mature fluvial landform altered by the self-organization process. By measuring the degree to which a fluvial network approaches OCN, I can measure the maturity of the fluvial landform, i.e., how long it has been altered by fluvial processes.

To understand the Martian VN s formation, I investigated the maturity of Martian VNs, which is the ratio between the energy dissipation of OCN and that of the real world watershed. In this paper, I also examined the associations between the maturity of Earth’s watersheds and the environmental factors and compared the maturity between the Earth’s and Mars’ watersheds. Then, I analyzed the spatial distribution of the VN maturity on Mars. Our results suggest that the maturities of many Martian VNs are similar to those of terrestrial watersheds and mature VNs distribution is wide, so Mars likely had long enough “warm” period to support surface runoff to alter the landform.

4.3 Dataset

The terrestrial DEM used in this research is the North America DEM from the HydroSHEDS database (Lehner et al., 2008). Since OCN analysis is computationally intensive,
I resampled the HydroSHEDS DEM to 1000 meters. The hydrological units were obtained from the NHDPlusV2 Dataset (McKay et al., 2012). Hydrologic Unit Code-10 (HUC10) watersheds was used as the basic statistical unit.

The Mars global DEM is based on data from the Mars Orbiter Laser Altimeter (Smith et al., 2001) and published by NASA. The spatial resolution on average is about 463 meters. We also resampled the DEM to 1000 meters to be consistent with our terrestrial analysis and to save computation time.

Although the fluvial landform is distributed widely on the Martian surface, some of them are covered by sand and dust or are eroded by eolian processes (Weitz et al., 2008). To identify the area preserved the fluvial landform on the Martian surface, I use the dendritic VN lines to select the qualified watersheds. The Martian VN s have been manually drawn by Hynek et al. (2010) and automatically extracted by Luo and Stepinski (2009). Alemanno et al. (2018) combined Luo and Stepinski (2009)’s and Hynek et al. (2010)’s results and interpreted the new Martian high resolution satellite images to extract the Martian VN s and classify them to six different classes: valley network, longitudinal valleys, valleys on volcanoes, valleys adjacent to canyons, single valleys and valleys segments, and small outflow channels. The dendritic pattern of VN s is a strong evidence indicating that the dendritic VN s are formed by the surface runoff, so the watersheds containing dendritic VN s are most likely eroded and modified by fluvial processes. The dendritic VN s (i.e., valley network from Alemanno et al. (2018)’s data) are selected to identify the qualified watersheds whose maturity would be measured.

The geological condition is represented by the lithology. The lithology (Theobald et al., 2015), which is based on texture, internal structure, thickness, and environment of deposition, is selected to represent the geological factor. The climatic conditions are represented by the aridity index. The aridity index (Zomer et al., 2006; Zomer et al., 2008) is the ratio between the average annual precipitation and modeled annual average evaporation. The
spatial variation is explored by dividing the conterminous U.S. to eight physiographic regions (Fenneman, 1928).

4.4 Method and Procedure

4.4.1 Hydrologic Analysis and Watershed Delineation

Hydrologic analysis extracts the hierarchy structure of DEM cells, digitalizes the stream lines and delineates the watershed boundary based on the DEM data. The general procedure for delineating watershed boundary is 1) filling the sinks; 2) determining the flow direction of each cell; 3) calculating the flow accumulation of each cell and extracting the stream cell; 4) labeling each stream; and 5) delineating the watershed boundary of each stream.

The sink region is the cells whose elevation is lower than their neighboring cells. In the filling sink step, the sinks on the DEM are filled by increasing the sink area’s elevation until the water in the sink area could flow out of the sink area. In the low resolution Earth DEM, most of the sinks are artifacts and errors. We filled all the sinks on the Earth DEM data by using the ArcGIS filling function. The sinks in the Mars MOLA DEM are the errors in the DEM generation or the small craters which formed after the period of fluvial processes. To eliminate the small craters, I used the 500 meters as the threshold to fill the sink, which is used in the Mest et al. (2010). After the filling process, all the sinks with a filling elevation of less than 500 meters are filled. The large craters, which were likely formed before/during the VNs’ formation period, are still preserved on the surface.

The flow direction builds the hierarchy structure of cells. The common algorithm of the flow direction is D8, multi-flow direction, and D-Infinity algorithm. The comparisons have been made by previous research projects (Wolock & McCabe Jr, 1995). Because I
only use the hierarchy structure to identify the upstream cells of the outlets, the impacts between adopting different flow direction algorithms are low. For this reason, I choose the D8 algorithm, because the spatial resolution in this project is low and the purpose of this project is to find the basic trend of flow. In the D8 algorithm, the flow direction of each cell is determined by the direction of the steepest descent from the cell to its Moore-Neighbor (3x3 window centered on target cell). If the descents are the same, the neighborhood is enlarged until the steepest descent is found.

After extracting the flow direction, the flow accumulation algorithm counts the upstream cells by using the cells’ hierarchy structure. Then, the stream cells are filtered by a selected threshold value, i.e., the cells having the flow accumulation above the threshold are the streams. For the Martian MOLA data, the threshold is 150 cells (i.e., 150 km$^2$), which is suggested by Mest et al. (2010). The higher threshold value would cause fewer stream junction points, fewer watershed polygons, and larger average area of watersheds. The task of this project is to calculate the watersheds’ OCN, so the OCN configuration would be searched within each watershed. The OCN searching would cost too much time if the watershed is too large because the OCN algorithm is an exponential time algorithm. With the threshold as 150, the average Martian watershed area is about 361 km$^2$.

In the labeling stream step, the program assigns unique values to sections of a raster linear network between intersections. After this step, each linear network between intersections has a unique ID. The end vertex is the outlet of the watershed. The outlet is the point on the surface at which water flows out of the area. It is the lowest point along the boundary of a watershed.

The watershed delineation step extracts boundary of the upstream area for each outlet. The extracted watersheds can be a large basin, which had no upstream area for the extract watershed. The extracted watersheds can also be a subbasin, which has water input from
its upstream subbasin. In this step, the input data is the outlet point and the flow direction data from previous steps. The output data is the watershed polygons.

### 4.4.2 Optimal Channel Network and Watershed Maturity

Based on the three principles of OCN, Rodriguez-Iturbe et al. (1992) derives the cost function of the OCN as Equation 4.16.

The derivation uses empirical hydrologic correlation and rely on some theoretical assumptions. Here, I re-examine Rodriguez-Iturbe et al. (1992)’s derivation to analyze the parameter in the cost function.

Given a channel without tributary (with width $w$, length $L$, slope $S$, flow depth $d$, and discharge $Q$), the flowing water in the channel gains the energy from the work of gravity. The gained energy is dissipated by the frictional force and by maintaining the channel’s shape. The direction of the gravity’s work is the down-slope direction. The down-slope component of gravity, $F_1$, is shown in Equation 4.1.

$$F_1 = \rho gdLwS \quad (4.1)$$

where $\rho$ is density of water and $g$ is gravitational constant.

In Equation 4.1, I used $S$ (i.e., $\tan \beta$) instead of $\sin \beta$ ($\beta$ is slope angle) because the slope angle is small in the most channels. I assume that the shape of the channel’s cross-section is rectangle.

The frictional force resists the movement and is the stress per unit area ($\tau$) times the wetted perimeter ($2d + w$) times length ($L$), so the frictional force is shown in Equation 4.2.

$$F_2 = \tau(2d + w)L \quad (4.2)$$
If the flow velocity does not change in the channel (i.e., $F_1 = F_2$), then I can have Equation 4.3.

$$\tau = \rho g SR$$

(4.3)

where $R$ is the hydraulic radius ($R = \frac{A_w}{P_w} = \frac{wd}{2d+w}$). $A_w$ is the area of channel’s cross-section. $P_w$ is the wetted perimeter of the channel’s cross-section.

In the turbulent incompressible flow, the boundary shear stress is proportionally related to the square of the average flow velocity ($v$) as Equation 4.4 showed.

$$\tau = C_f \rho v^2$$

(4.4)

where $C_f$ is a dimensionless resistance coefficient.

Equation 4.3 and Equation 4.4 can be combined since they both have $\tau$, then I can have Equation 4.5.

$$S = C_f \frac{v^2}{Rg}$$

(4.5)

Assuming the friction force is equal to the downslope component of gravity, the energy consumed by friction for the volume of water per unit time ($Q$) is shown as Equation 4.6.

$$P_{friction} = S \rho g QL$$

$$= \rho C_f \frac{v^2}{R} QL$$

(4.6)

The flow also consumes energy by carrying the sand and gravel to maintain the channel shape. The energy in this part is represented by $F(soil, flow)P_wL$, where $F(soil, flow)$ is a function related with the soil and flow properties. I assume that $F(soil, flow)$ is the bed load transport equation, so the $F(soil, flow)$ is $K\tau^m$, where $K$ and $m$ relates the soil and
fulid properties and are constant in the watershed, so the energy consumed by maintaining the channel shape is shown as Equation 4.7.

\[ P_{\text{maintain}} = K\tau^m P_w L \] (4.7)

In a channel given above, the energy dissipation is the sum of energy consumed by fiction (Equation 4.6) and energy consumed by maintaining channel shape (Equation 4.7) as shown in Equation 4.8.

\[ P = P_{\text{friction}} + P_{\text{maintain}} \]
\[ = \rho C_f \frac{v^2}{R} QL + K\tau^m P_w L \] (4.8)

The Equation 4.8 may be written as Equation 4.9 after I substitute Equation 4.4 and \( R = \frac{A_w}{P_w} = \frac{wd}{2d+w} \) in Equation 4.8.

\[ P = P_{\text{friction}} + P_{\text{maintain}} \]
\[ = \rho C_f \frac{v^2}{R} QL + K\tau^m P_w L \]
\[ = \rho C_f \frac{v^2}{A_w} QL + K\tau^m P_w L \]
\[ = \rho C_f P_w \frac{(wdv)^2 (wd)^2}{A_w (wd)^2} L + KC_f^m \rho^m v^{2m} P_w L \] (4.9)
\[ = \rho C_f P_w \frac{(wdv)^3}{(wd)^3} L + KC_f^m \rho^m v^{2m} P_w L \]
\[ = \rho C_f P_w \frac{Q^3}{A_w^3 L} + KC_f^m \rho^m v^{2m} P_w L \]

Based on the second principle of the OCN (i.e., principle of equal energy dispersion per unit area of channel anywhere in the network), unit dissipation \( P_1 \) in Equation 4.10 is constant values in the channel.
\[ P_1 = \frac{P}{P_w L} \]
\[ = \rho C_f \left( \frac{Q}{A_w} \right)^3 + K C_f^m \rho^m v^{2m} \]
\[ = \rho C_f \left( \frac{w d v}{w d} \right)^3 + K C_f^m \rho^m v^{2m} \]
\[ = \rho C_f v^3 + K C_f^m \rho^m v^{2m} \]  \hspace{1cm} (4.10)

In Equation 4.10, since the \( C_f, \rho \) and \( K \) are determined by the watershed’s conditions and are constant in the watershed, I can conclude that the flow velocity(\( v \)) is also the same in the OCN watershed.

Since \( Q = w d v \), I can derive that \( w = \frac{Q}{v d} \). Substituting \( w = \frac{Q}{v d} \) in Equation 4.9, I have Equation 4.11. In Equation 4.11, the last bracketed terms are constant through the network for a give flow condition.

\[ P_{new} = \rho C_f (2d + w) \frac{Q^3}{(w d)^3} L + K C_f^m \rho^m v^{2m} (2d + w) L \]
\[ = \rho C_f (2d) \left( \frac{w d v}{w d} \right)^3 L + \rho C_f w \frac{Q w d v^2}{(w d)^3} L \]
\[ + K C_f^m \rho^m v^{2m} 2d L + K C_f^m \rho^m v^{2m} w L \]
\[ = \rho C_f (2d) \left( \frac{w d v}{w d} \right)^3 L + \rho C_f w \frac{Q w d v^2}{(w d)^3} L \]
\[ + K C_f^m \rho^m v^{2m} 2d L + K C_f^m \rho^m v^{2m} \frac{Q}{v d} L \]  \hspace{1cm} (4.11)
\[ = 2L d \rho C_f v^3 + \frac{Q L}{d} \rho C_f v^2 \]
\[ + 2L d K C_f^m \rho^m v^{2m} + \frac{Q L}{d} (K (C_f)^m \rho^m v^{2m-1}) \]
\[ = \frac{Q L}{d} (C_f \rho v^2 + K (C_f)^m \rho^m v^{2m-1}) \]
\[ + L d (2C_f \rho v^3 + 2K C_f^m \rho^m v^{2m}) \]
Based on the first principle of OCN (the principle of minimum energy dissipation in any link of the network for the transportation of a given discharge), the partial derivative with respect to $d$ is 0. Then, I have equation 4.12.

$$\frac{\partial P_{\text{new}}}{\partial d} = 0$$

$$= -\frac{QL}{d^2}(C_f \rho v^2 + K C_f^m \rho^m v^{2m-1})$$

$$+ L(2C_f \rho v^3 + 2K C_f^m \rho^m v^{2m})$$

(4.12)

Since the bracketed terms in Equation 4.12 are constant in the watersheds, I get Equation 4.13.

$$Q = d^2\left(\frac{2C_f \rho v^3 + 2K C_f^m \rho^m v^{2m}}{C_f \rho v^2 + K C_f^m \rho^m v^{2m-1}}\right) = d^2 t$$

(4.13)

where $t$ represents the bracketed term, which is constant value in the watershed.

Based on Equation 4.13, I can also have that $d = t^{0.5}Q^{0.5}$. After substituting $d = t^{0.5}Q^{0.5}$ to Equation 4.11, I obtain the optimal energy dissipation at any node.

$$P = kQ^{0.5}L$$

(4.14)

where $k$ is the constant value in the watershed.

The total energy dissipation in the watershed is sum of energy dissipation of every node in the watershed. In the DEM data, the nodes are the DEM cells.

$$P = k \sum Q^{0.5}L_T$$

(4.15)

where $L_T$ is the distance between the nodes in the DEM data.

In the process of searching for the minimum energy dissipation, the $k$ and $L_T$ in the Equation 4.15 did not influence the results because they have the same values for each node.
Also, it is impossible to measure the discharge \((Q)\) at every node, so Rodriguez-Iturbe et al. (1992) assumed the relation between the flow contribution area \(A\) and \(Q\) is linear and use measurable flow contribution area \(A\) to replace the \(Q\). The cost function Equation 4.15 is converted to Equation 4.16. According to the third principle of the OCN (the principle of minimum energy dissipation in the network as a whole), the OCN searching process is to search cells’ hierarchy relation configuration having the minimum \(H\).

\[
H = \sum A_i^\gamma = \sum A^{0.5}
\] (4.16)

where \(A_i\) is the contribution area of \(i\) and \(\gamma\) is the exponent (to be more general).

To derive the cost function (Equation 4.16), Rodriguez-Iturbe et al. (1992) relied on many assumptions, e.g., the flow velocity in the channel are the same; the relation between the discharge and the contribution are linear. However, these assumptions may not be correct. Also, a lot of factors, such as heterogeneous surface property, vegetation, groundwater, also influence the cost function. The uncertainties may shift the value of \(\gamma\), which is the exponent of \(A\), or even make the value of \(\gamma\) varies within the watershed. Rodriguez-Iturbe and Rinaldo (2001) suggested that research project can use smaller \(\gamma\) value when the contribution area is smaller than the threshold and use larger \(\gamma\) value when the contribution is greater than the threshold. To illustrate the OCN under different \(\gamma\) values, I used 0.1, 0.5, and 0.9 as the \(\gamma\) value respectively to simulate the OCN in a 30 by 30 rectangle area. The results are shown in the Figure 14.

Considering the spatial heterogeneity and uncertainties of the \(\gamma\), I used a set of \(\gamma\) values to explore the OCN in different regions. The possible range of \(\gamma\) values was discussed by Rodriguez-Iturbe and Rinaldo (2001). Rodriguez-Iturbe and Rinaldo (2001) analyzed the patterns of stream networks under different \(\gamma\). Based on their results, the range of \(\gamma\) is between 0 and 1. If the \(\gamma\) is equal to or greater than 1, the aggregation network pattern is
Figure 14: OCN under different $\gamma$ values
The black lines are the river networks and are created by OCN algorithm. The blue bold line is the mainstream (the longest path) of the network. The figure demonstrates the structure of optimized network by using different $\gamma$. With the increase of $\gamma$, the flow path is straighter.

not produced. If the $\gamma$ is equal or less than 0, the stream tends to aggregate and forms a network with a spiral-like pattern. For this reason, I simulate the OCN of each watershed by using 3 different exponent values (i.e., $\gamma = 0.3, 0.5$ or 0.7) and compare the OCN’s energy dissipation and the real energy dissipation of each watershed to calculate the watersheds’ maturity.

To search for the OCN within the minimum energy dissipation of the watershed, the exhaustion method should be applied in the search algorithm. However, the complexity and the computation cost of the exhaustion method is too high, so solving the OCN by using exhaustion method is impossible (Rodriguez-Iturbe & Rinaldo, 2001). Previous research projects (Abed-Elmdoust et al., 2016; Abed-Elmdoust et al., 2017; Rodriguez-Iturbe & Rinaldo, 2001) showed that the suboptimal solution provided by the simulated annealing algorithm can be used as the approximate optimal solution of the OCN. Carraro et al. (2020) implement an R library for generating the OCN in the rectangle area. The OCN R library built by Carraro et al. (2020) cannot be applied on the non-rectangular watersheds, which are more common in the real world. I followed the annealing strategy in the Carraro
et al. (2020)’s simulated annealing algorithm and implement an algorithm to search for the OCN in the arbitrary shaped watershed based on the simulated annealing algorithm.

The simulated annealing algorithm is a probabilistic technique for obtaining the approximate global optimal result. The name is from the annealing in metallurgy, a technique relating heating and controlled cooling to reduce the crystals’ defects and increase crystals’ size. The concept of the slow cooling process in the simulated annealing simulation is represented as the slow decrease in the probability of accepting worse solutions. Specifically, the simulated annealing algorithm starts with an initialized random solution and assesses the solution’s quality. Then, it implements a for-loop to simulate the cooling, whose temperature progressively decreases from an initial positive value to 0. In each iteration, the algorithm randomly selects a solution close to the current one and assessed its quality. If the quality of the new solution is better than that of the old one, the new solution is accepted and will be compared with the next new solution. If the quality is worse than that of the old solution, the new solution would be accepted as well if the generated random number, which is within the range of 0 and 1, in this iteration is greater than the temperature-dependent probability. After the for-loop, the solution is the approximate optimal solution.

In the OCN searching, the algorithm has the following steps. (1) It initializes the flow direction of each cell and computes the initial energy dissipation. (2) It creates the temperature list. The size of the list is 40 * the number of cells in the watershed. (3) It executes a for-loop N times. In each loop, it changes one cell’s flow direction and re-calculate the energy dissipation. If the new energy dissipation is lower than the old one, the program keep the new configuration and record the new energy as the minimum energy dissipation. If the new energy is greater, the program generate a random number between 0 and 1. If the generated random number is smaller than the threshold, which is related with the temperature negatively, the program accept the new configuration and the new energy dissipation as the minimum energy dissipation. If the generated random number is greater than the threshold,
the program keeps the old flow direction configuration and old energy dissipation. After the loop section, the final flow direction and energy dissipation is the approximate OCN.

After the program searched the watershed’s OCN and its corresponding energy dissipation, I calculate the ratio between the energy dissipation based on the OCN and the energy dissipation based on the hierarchy relation calculated from the DEM data. The ratio (a value between 0 and 1) represents the watershed’s maturity. The high ratio represents that the watershed have been modified by the fluvial processes long enough and is close to the optimal status. The low ratio indicates that the fluvial process is not dominated or that landform in the watershed is not altered by the fluvial processes long enough.

4.4.3 Spatial Association Measured by SPADE Detector

To compare the associations between maturity of fluvial landform and environmental factors (aridity, physiographic regions, and lithology), I applied the SPatial Association DETector (SPADE) (Cang & Luo, 2018) to assess the associations. The SPADE assesses the association by comparing the spatial heterogeneity of factors. If the spatial heterogeneities of the two factors are similar, the association between two factors is high. Otherwise, the association is low. The SPADE is an improvement of the Geo-detector (Wang et al., 2010; Wang et al., 2016). The Geo-detector did not consider the distance decay and is influenced by the number of discretization zones for continuous variables. The association is represented by an association parameter usually ranging between 0 and 1, with 0 representing no association and 1 perfect association. The negative association represents no association and is caused by outliers or skewed distribution of dependent variation. The details of SPADE can be found in Cang and Luo (2018).
4.4.4 Comparative Analysis Procedure

4.4.4.1 Maturity Analysis of Terrestrial Watersheds

To analyze the maturity of fluvial landform on Earth, I selected the conterminous U.S. area as terrestrial data because the conterminous U.S. includes varied climates types and I have access to a large database of the environmental variables. I randomly selected 20% of the hydrologic unit 10 (HUC10) in the conterminous U.S. (3420 polygons are selected). Within each HUC10 polygon, I selected the largest watersheds extracted from the resampled HydroSHEDS DEM (spatial resolution is 1000 meters) because the HUC10 hydrologic unit polygon may contain multi-watersheds extracted from HydroSHEDS DEM and the OCN program can only process one watershed each time. After I extracted the largest watersheds from the HUC10 hydrologic units, I calculated the energy dissipation of every watershed based on the cells’ hierarchy relation and searched every watershed’s OCN to obtain the minimum energy dissipation. We then calculated the ratio between the energy dissipation of the simulated OCN of each watershed and the energy dissipation of the current watershed. Here, I use the ratio to measure maturity of the watershed. Theoretically, the measured energy dissipation is larger than the OCN’s energy dissipation, so the ratio value is smaller than 1 and greater than 0. The high ratio value means that the watershed area have been eroded by the fluvial process long enough and that the watersheds are self-organized significantly.

After I computed the maturity of the watersheds, I investigated the association between the watersheds’ maturity and environmental factors (aridity and lithology) by using the SPADE (Cang & Luo, 2018). I also investigate the spatial variation by assessing the relation between the maturity of watershed and physiographic regions. The dependent variable in the SPADE is the watershed maturity. The independent variables are the zonal average of
climatic factor (aridity), zonal majority of geologic factor (lithology), and zonal majority of physiographic region ID. After the association in the whole conterminous U.S. area is measured, I also measured the association between the watershed maturity and environmental factors in the eight physiographic divisions of the conterminous U.S. because relations of the fluvial landform and environmental factors have the spatial variation (Luo et al., 2016).

### 4.4.4.2 Martian Watersheds Maturity Analysis

To measure the maturity of the Martian fluvial landform, I need to identify the VNs’ area likely influenced by the fluvial processes before I measure the maturity of the watersheds. If the fluvial process did not occur in the extracted watersheds, the extracted watersheds are meaningless even though watershed extraction algorithm described in the previous section can produce the visually appealing results. I extracted the watersheds by using the procedure and the parameters described in the previous section firstly. Then, I selected the qualified watersheds that contain dense and dendritic VNs (Alemanno et al., 2018; Cang & Luo, 2019). Because the OCN algorithm is very time-consuming, I sampled 20% qualified watersheds to simulate the OCN and to compute the maturity. After I calculated the maturity of the selected qualified watersheds, I grouped the watershed by using each dendritic VN from Alemanno et al. (2018)’s VN data and used the median maturity in each group as the corresponding VNs’ maturity. I used the median value to exclude the influence of the outliers because the hierarchy relation formed in the Noachian period might be modified by the post-formation processes.
4.5 Results

4.5.1 Terrestrial Watershed Maturity Analysis

A total of 3420 watersheds are selected from the conterminous U.S. HUC10 polygons. The average and standard deviation of watershed size are 538.69 km$^2$ and 295.44 km$^2$, respectively. The average and the standard deviation of watersheds’ maturity are shown in Table 10. The histograms and spatial distribution of their maturity under different $\gamma$ values are shown in Figure 15 and Figure 16, respectively.

Table 10: Average and SD of terrestrial watershed maturity

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>Average</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.93</td>
<td>0.03</td>
</tr>
<tr>
<td>0.5</td>
<td>0.90</td>
<td>0.04</td>
</tr>
<tr>
<td>0.7</td>
<td>0.87</td>
<td>0.05</td>
</tr>
</tbody>
</table>

These results indicate that U.S. watersheds are generally very mature, with average maturity reaching around 0.9. This is consistent with long term fluvial erosion creating channels approaching OCN condition and minimizing energy dissipation. More discussion will be provided in Section 4.6.1

The association between the watersheds’ maturity and environmental factors at the conterminous U.S. scale are shown in Table 11. Figure 17 shows the spatial distribution of zonal average aridity index and zonal majority lithology, and the eight physiographic regions.

Table 11: Associations between watersheds’ maturity and environmental factors at the conterminous U.S. scale

<table>
<thead>
<tr>
<th>Factors</th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.5$</th>
<th>$\gamma = 0.7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aridity Index</td>
<td>0.223</td>
<td>0.125</td>
<td>0.067</td>
</tr>
<tr>
<td>Lithology</td>
<td>0.210</td>
<td>0.105</td>
<td>0.050</td>
</tr>
<tr>
<td>Physiographic Regions</td>
<td>0.084</td>
<td>0.043</td>
<td>0.026</td>
</tr>
</tbody>
</table>
Figure 15: Histograms of terrestrial watersheds’ maturity

After the watersheds are divided into the eight physiographic divisions, the associations between the maturity and environmental factors in each division are shown in the Table 12(a). The mean and SD of watersheds’ maturity by region are shown in Table 12(b).

These results suggest that the associations between maturity and environment factors are generally weak, with the strongest being between maturity and aridity index and weakest between maturity and physiographic regions. The association within each physiographic region does vary, with stronger one observed in arid to semi-arid regions. In general, the maturities in arid regions are lower than those in humid regions. More discussion will be provided in section 4.6.1
Figure 16: Spatial distribution of terrestrial watersheds’ maturity. For clarity, centroid of each watershed is shown as point.
Figure 17: Spatial distribution of environmental factors
Table 12: Associations between watersheds’ maturity and environmental factors in Regions

(a) Associations between watersheds’ maturity and environmental factors in Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>Aridity Index</th>
<th>Lithology</th>
<th>Aridity Index</th>
<th>Lithology</th>
<th>Aridity Index</th>
<th>Lithology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appalachian Highlands</td>
<td>0.048</td>
<td>0.016</td>
<td>-0.003</td>
<td>0.008</td>
<td>-0.022</td>
<td>0.019</td>
</tr>
<tr>
<td>Atlantic Plain</td>
<td>0.034</td>
<td>0.085</td>
<td>0.062</td>
<td>0.060</td>
<td>0.017</td>
<td>-0.002</td>
</tr>
<tr>
<td>Interior Highlands</td>
<td>0.046</td>
<td>0.199</td>
<td>0.073</td>
<td>0.081</td>
<td>0.062</td>
<td>0.075</td>
</tr>
<tr>
<td>Interior Plains</td>
<td>0.159</td>
<td>0.143</td>
<td>0.108</td>
<td>0.111</td>
<td>0.054</td>
<td>0.066</td>
</tr>
<tr>
<td>Intermontane Plateaus</td>
<td>0.230</td>
<td>0.229</td>
<td>0.161</td>
<td>0.057</td>
<td>0.100</td>
<td>0.027</td>
</tr>
<tr>
<td>Laurentian Upland</td>
<td>-0.094</td>
<td>0.152</td>
<td>-0.121</td>
<td>0.057</td>
<td>0.100</td>
<td>0.012</td>
</tr>
<tr>
<td>Pacific Mountain System</td>
<td>0.256</td>
<td>0.417</td>
<td>0.174</td>
<td>0.246</td>
<td>0.076</td>
<td>0.147</td>
</tr>
<tr>
<td>Rocky Mountain System</td>
<td>0.190</td>
<td>0.258</td>
<td>0.196</td>
<td>0.091</td>
<td>0.016</td>
<td>0.022</td>
</tr>
</tbody>
</table>

(b) Mean and SD of watersheds’ maturity in Regions

<table>
<thead>
<tr>
<th>Region</th>
<th>( \gamma = 0.3 )</th>
<th>( \gamma = 0.5 )</th>
<th>( \gamma = 0.7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Appalachian Highlands</td>
<td>0.930</td>
<td>0.025</td>
<td>0.905</td>
</tr>
<tr>
<td>Atlantic Plain</td>
<td>0.930</td>
<td>0.023</td>
<td>0.904</td>
</tr>
<tr>
<td>Interior Highlands</td>
<td>0.922</td>
<td>0.017</td>
<td>0.897</td>
</tr>
<tr>
<td>Interior Plains</td>
<td>0.906</td>
<td>0.030</td>
<td>0.901</td>
</tr>
<tr>
<td>Intermontane Plateaus</td>
<td>0.937</td>
<td>0.016</td>
<td>0.876</td>
</tr>
<tr>
<td>Laurentian Upland</td>
<td>0.930</td>
<td>0.023</td>
<td>0.888</td>
</tr>
<tr>
<td>Pacific Mountain System</td>
<td>0.937</td>
<td>0.016</td>
<td>0.899</td>
</tr>
<tr>
<td>Rocky Mountain System</td>
<td>0.938</td>
<td>0.016</td>
<td>0.900</td>
</tr>
</tbody>
</table>
4.5.2 Martian Watershed Maturity Analysis

A total of 3326 Martian watersheds are selected to calculate the maturity. I calculated the Martian watersheds’ maturity by choosing 0.3, 0.5 and 0.7 as the $\gamma$ value. The average, median, SD are shown in Table 13. The histogram and spatial distribution of the watersheds’ maturity are shown in Figure 18 and Figure 19. These results indicate that the Martian watersheds are less mature than their terrestrial counterparts and the Martian fluvial landform is comparatively immature.

Table 13: Average, median, and SD of Martian watershed maturity

<table>
<thead>
<tr>
<th></th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.5$</th>
<th>$\gamma = 0.7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.90</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Median</td>
<td>0.90</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>SD</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
</tr>
</tbody>
</table>

To investigate the local maturity and eliminate the outliers (out of $2\sigma$ range), I grouped the watersheds by the dendritic VNs formed in the Noachian period and selected the VNs intersected with more than 2 watersheds, and used the median maturity value in each group to represent the maturity of VNs. The histogram and spatial distribution are shown in Figure 20 and Figure 21. A total of 369 VNs are selected. The mean, median, and SD of VNs are shown in Table 14. The maturity values are close the arid or semi-arid area ($AI < 0.5$) in the conterminous U.S. area. The mean and SD of watersheds’ maturity are shown in Table 10 and Table 12(b).

Table 14: Average, median, and SD of Martian VNs maturity

<table>
<thead>
<tr>
<th></th>
<th>$\gamma = 0.3$</th>
<th>$\gamma = 0.5$</th>
<th>$\gamma = 0.7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>0.90</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>Median</td>
<td>0.90</td>
<td>0.88</td>
<td>0.87</td>
</tr>
<tr>
<td>SD</td>
<td>0.03</td>
<td>0.03</td>
<td>0.02</td>
</tr>
</tbody>
</table>
4.6 Discussion and Conclusion

4.6.1 Earth Watersheds Maturity

Table 10 indicated that most terrestrial watersheds’ maturity are greater than 0.86 even though under different $\gamma$ values. It means that a lot of watersheds in the conterminous U.S. area have been well modified by the fluvial processes. Figure 15 showed that distributions of maturity vary with different $\gamma$ values. The higher the $\gamma$ value, the higher the average maturity and the narrower the range of distribution. Considering the spatial variation, I cannot conclude which $\gamma$ value is more suitable. For this reason, I use the results based
Figure 19: Martian watersheds’ maturity.
For clarity, watersheds are represented by their centroids.
Figure 20: Histogram of Martian VN's maturity.

on all the $\gamma$ values I used in Table 10 to explore the associations between maturity and environmental factors.

Table 12 showed that environmental factors are more important in the Intermontane Plateaus region, the Pacific Mountain System region and the Rocky Mountain System region than other regions. The Intermontane Plateaus region, the Pacific Mountain System region and the Rocky Mountain System region are mainly arid or semi-arid areas ($AI < 0.5$, see also Figure 4 for their location). The possible explanation is that the environmental factors had minor influence on girds' hierarchy relations in the more humid regions where the duration and intensity of fluvial processes is long and strong enough.
Figure 21: Spatial distribution of Martian VN’s maturity.
4.6.2 Outliers of Martian VNs’ Maturity

In the results, some maturity values of Martian VNs are much lower than the maturity on Earth. I consider those values outliers. If I increase the threshold for selecting the VNs to six watersheds to measure the large VNs’ maturities, the minimal maturity of VNs increased from 0.69 to 0.82 when the $\gamma = 0.5$. The average maturity is similar with the arid or semi-arid area ($AI < 0.5$) in the conterminous U.S. area, which is consistent with previous research (Cang & Luo, 2019; Matsubara et al., 2013; Orofino et al., 2018). The potential causes to outliers is that the filling function can not perfectly re-construct the hierarchy relation in the DEM data.

4.6.3 Martian VNs Maturity and Early Martian Climate

The associations between the terrestrial watershed maturity and the environmental factors showed that the environmental factors and the watershed maturity have a stronger association in the arid/semi-arid area than the associations between the watershed maturity and environmental factors in the humid area. Also, based on the previous early Martian climate research, the aridity is not wetter than the arid/semi-arid areas on Earth (Cang & Luo, 2019; Matsubara et al., 2011; Matsubara et al., 2013). So, maturity of Martian watershed is a good indicator for the Martian environmental conditions. Since the physically-based landform model is not linked to the OCN model and initial landform condition is not clear, it is hard to model the association between the maturity and the environmental conditions by the quantitative model (e.g., regression model in Cang and Luo(2019)). However, the maturity is still an indicator for the duration of early Martian “warm” period, because the
watershed with a high maturity value had more flowed water than the watershed with a low maturity value, assuming that the underlying surface is homogeneous.

The histogram of maturity of Martian watersheds (Figure 20) showed that Martian VNs are immature comparing to the terrestrial watersheds. The median maturity is close to watershed maturity of the arid and semi-arid area in the conterminous U.S. The low maturity values of Martian watersheds suggest that Martian fluvial landform is immature comparing to the terrestrial landform, which is consistent with previous research (Galofre, Bahia, et al., 2020; Penido et al., 2013; Som et al., 2009). Figure 21 showed that large VNs, whose values are greater than 0.82 (the lower boundary of two-sigma range of terrestrial watershed maturity when $\gamma = 0.5$), are distributed on the Martian surface widely. It indicates that long-term fluvial activity on the Martian surface were distributed widely.

4.6.4 Conclusion

In this paper, I used the ratio of energy dissipation along the simulated OCN within a watershed to that along the real network as a measure of the maturity of the watershed. By analogizing with terrestrial analysis, I investigated the early Martian climatic conditions and important controlling factors.

The maturity measure by the OCN offers an important method to reveal whether the fluvial processes modified the landform long enough. Maturity is a robust characteristic because it is based on hierarchy relations within the watershed and they are not influenced by the post-formation processes easily.

In my terrestrial analysis, I confirmed that most Earth watersheds are close to the optimal status due to the active hydrologic cycling in the conterminous U.S. Also, I found that the maturity is easily influenced by the environmental factor in the semi-arid or arid area. Since
the ancient Mars was likely semi-arid or arid (Cang & Luo, 2019; Matsubara et al., 2011; Matsubara et al., 2013), it gives us an opportunity to use the maturity of watersheds to investigate the ancient Mars climatic conditions.

The lower average maturity of Martian watersheds comparing to that of the terrestrial counterparts suggests that Martian watersheds are immature, which is consistent with previous research (Galofre, Bahia, et al., 2020; Penido et al., 2013; Som et al., 2009). We also found that the maturity of the large VNs are within the 2-sigma range of terrestrial watershed maturity and are distributed widely on Mars. It supports that Mars had global long-term flowed liquid water.
5.1 Overview

The overall goal of this research is to test the ancient Martian climate hypotheses by investigating the spatial pattern of the Martian VNs’ robust characteristics and by analogizing the Earth streams and Martian VNs. This is accomplished through three separate but related articles: (1) developing a spatial analysis method to investigate the association between stream’s properties and environmental factors, (2) analyzing and analogizing a VNs’ robust characteristic (junction angle) and environmental factors, and (3) quantifying Martian watersheds’ maturity by comparing their energy dissipation with those of OCNs.

In the first manuscript, I built a spatial analysis method, SPADE, to measure the spatial associations between factors by comparing their spatial heterogeneity. The spatial heterogeneity refers to the landscapes’ uneven spatial distribution within an area. If the two spatial layers have a causal relation, then their spatial pattern correlates. The SPADE improved the Geographical Detector by explicitly considering the effects of distance decay and the level of discretization for continuous variables. The results in the first manuscript showed that the SPADE is a better measure of association between spatially distributed data than the original Geographical Detector.

In the second manuscript, I analyzed the association between the junction angles and environmental factors on Earth. The results suggested that the climatic factors are stronger than or on par with the geologic factors in controlling junction angles and the climatic
parameters (Aridity Index (AI) and Mean Annual Precipitation (MAP)) can be estimated from junction angles. I then applied the associations between terrestrial junction angle and climatic conditions to estimate the AI and MAP on Mars. The spatial analysis of inferred climatic conditions showed that the Noachian Mars had an active global hydrological cycle and that Mars was “warm” during the VNs’ formation period. The duration of “warm” Mars is estimated by the ratio between the required water volume (Luo et al., 2017) to form VNs and the runoff discharge derived from MAP. The range of “warm” period is from about 4.4 to about 77 million years.

In the third manuscript, I first analyzed the watersheds’ maturity in the conterminous U.S. area. The results suggested that the watershed maturity in most of conterminous U.S. areas are high. I next assessed the associations between the watersheds’ maturity and environmental factors (aridity and lithology). The results showed a stronger relationship between maturity and aridity than that between maturity and other environmental factors in the arid or semi-arid areas. To investigate the early Martian climatic conditions, I also measured the Martian watersheds’ maturity and found that Martian watersheds are less mature than terrestrial watersheds. The finding is consistent with previous research (Galofre, Bahia, et al., 2020; Penido et al., 2013). The local VNs’ maturity showed some VNs’ maturity are as high as their terrestrial counterparts. The results support that the ancient Mars was “episodically warm” at least and that the ancient Mars had surface runoff long enough to carve the Martian surface in some regions.

5.2 Future Work and Research Directions

This research work has provided some unique insights that the traditional geoscience methods cannot offer. The general approach in this research work is to establish the spatial
associations between the VNs' properties and climate conditions on Earth, then apply them to Mars to infer early Mars climate. The VNs’ properties in this research include the stream junction angle, which is the fingerprint of the fluvial processes, and the hierarchy relation of elevation cells, which is the imprint left by fluvial landform evolution.

This research can be taken a step further in a number of directions in the future. One direction is to improve the SPADE. First, the SPADE treats data isotopically and Euclideanly. To improve the SPADE, I can use directions to weight the squared difference of value pairs, because many geographical processes are anisotropic in nature. Also, I need to add more distance metrics to meet the requirements of different research fields. Especially, I can construct a weighted matrix by using the topological distance within the SPADE, because the adjacency is more important than the Euclidean distance in human geography. Second, if the dependent variable of SPADE is homogeneous (i.e., the values of the dependent variable are the same), the SPADE measuring the association between spatial variables would offer invalid information. The future research includes finding a solution to this problem and testing the performance of the SPADE.

The second direction is to investigate the influence of geologic factors to the characteristic of fluvial landform. In the chapter 3 and 4, I assume that the Martian underlying surface is homogeneous to simplify the research question. Based on the previous research, the underlying surface of Mars should have heterogeneity (Ehlmann & Edwards, 2014). Also, in the chapter 3 and 4, the terrestrial data showed that the associations between the characteristics of fluvial landform and geologic factors (e.g., lithology) are significant and are not much weaker than the associations between climatic factors and characteristics of fluvial landform. For this reason, it is important to investigate the influence of the geologic factors to the junction angles and watershed maturity and to apply the model from the terrestrial data to Martian dataset.
The third direction is to combine the spatial analysis of landform properties with General Circulation Models (GCMs) (Turbet & Forget, 2021; Wordsworth et al., 2015; Wordsworth et al., 2021) and physically-based landform evolution model (Barnhart et al., 2009; Boatwright & Head, 2019) to test the ancient Martian climate hypotheses. The estimated climatic conditions from landform properties and modeled climatic conditions from GCMs with different configurations will be compared. The configurations of GCM will be ranked based on the associations between their modeled results and estimates from Mars landform properties. The initial conditions of high-ranked GCM will likely be true initial conditions. The modeled climatic conditions form high-ranked GCMs will also be the input of the physically-based landform evolution. The output of the physically-based landform evolution can be compared with the Martian landform properties. The comparison results can be backpropagated to the GCMs and physically-based landform evolution model for calibrating their configurations.
CHAPTER 6
RESPONSE TO COMMITTEE’S COMMENTS

6.1 Overview

In the oral dissertation defense, the committee members provided some suggestions about the uncertainty of data and modeling results, philosophical thinking of research method, and alternative explanation for results. I will use this chapter to address their concerns and suggestions. In this chapter, I will discuss the influence of map projection to junction angle extraction, uncertainty of estimation model, philosophical thinking about analogizing in planetary research, and alternative explanation for the immature Martian watershed.

6.2 The Uncertainties of Junction Angle Influenced by the Map Projection

In Chapter 3, the terrestrial and Martian junction angles are extracted under the sinusoidal projection, which is an equal area projection. Since the junction angles are determined by the directions of two inflow valley network polylines, it is necessary to compare the differences of extracted junction angles under different map projections. To compare the differences, I extract the junction angles from North American HydroSHED river networks data under Mercator projection(a conformal projection) and Sinusoidal projection(an equal-area projection) seperately.
The North American HydroSHED river networks data are selected in this experiment because the data cover larger range of latitudes and have less amount of streams comparing with the data amount in the NHDPlusV2 Dataset, which was used in Chapter 3. The larger coverage area could increase the credibility of the comparison results between extracted junction angles, and low amount of data could reduce the computation time.

The basic statistical characteristics of extracted junction angles under different map projections are shown in the Table 15. Although the statistical characteristics between junction angles under two different projections are similar, some junction angles are significantly different. For example, the junction angle at (100.926° W 57.571° N), the junction angle is 61.79° under the Mercator projection and is 77.91° under the Sinusoidal projection. The difference is not caused by the junction measuring part of the junction angle extraction algorithm (which is based on the ratio of sectional area to the area of the circle, see Figure 10), but is caused by the distance distortion under different projections in polyline simplification. As I mentioned in Chapter 3, the stream polylines are simplified by the Douglas-Peucker algorithm (Douglas & Peucker, 1973), which used a distance threshold. The algorithm is implemented by the ArcGIS and use the planner distance as the threshold. The same geodesic distances varies under different map projections. For this reason, the simplified polylines may have difference, leading to differences in polyline simplification and thus resulting junction angles.

To verify this idea, I used a hybrid method by first simplifying the streams polylines under Sinusoidal projections and then converting the simplified polylines to Mercator projection for the rest steps of junction angle extraction. The basic statistic characteristics of the hybrid method are shown in Table 15. The junction angle at (100.926° W 57.571° N) is 77.89° using the hybrid method and is 77.91° using the original method under the Sinusoidal projection. The results showed that the major influence factor in my junction angle extraction is the distance distortion in line simplification distance threshold under different map projections.
projection. The sinusoidal projection, which is what I used, preserves area and has the less distance distortion than the Mercator projection, especially in the high latitude area, so the sinusoidal projection is a more suitable than the Mercator projection in using the junction angle extraction algorithm described in Chapter 3 (based on fractional area of a circle, see Figure 10).

Table 15: Comparisons of junction angles under different map projections

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercator projection</td>
<td>90.58</td>
<td>90.00</td>
<td>35.59</td>
</tr>
<tr>
<td>Sinusoidal projection</td>
<td>89.19</td>
<td>90.00</td>
<td>34.01</td>
</tr>
<tr>
<td>Hybrid method</td>
<td>89.30</td>
<td>90.00</td>
<td>34.17</td>
</tr>
</tbody>
</table>

6.3 The Uncertainties of Models

In the Chapter 3, I used the Ordinary Least Squares (OLS) and the $k$-NN methods to build model between terrestrial junction angles and terrestrial climatic conditions. Although I used upper and lower limit of parameters (e.g., evaporation, discharge rate, and scale) to estimate the upper and lower range of the length of “warm” duration, I did not consider the uncertainties contributed by the estimation models. The estimation models (OLS and $k$-NN) also have the uncertainties. The uncertainties could be expressed by the confidence interval. The confidence interval could not only provide the upper limit and lower limit of estimated value but also be used to verify whether the model is effective. If the confidence interval is too large, then the credibility of models is low.

In the OLS, the estimated value have a confidence interval. Normally, the 95% confidence interval is used in the research. The confidence interval provide the upper limit and lower limit of the estimated Martian Mean Annual Precipitation ($MAP$). This confidence interval
of estimated value could also add the upper limit and reduce the lower limits of length of “warm” Mars.

Since the k-NN is not a model-based method, the confidence interval can not be estimated directly. I suggest that use a testing dataset to extract the distribution of errors. Then, use the estimated $MAP \pm 2 \sigma_{of\ errors}$ as the confidence interval.

I will address these issues when I submit this manuscript for publication.

6.4 Philosophical Thinking about Analogizing in Planetary Research

Although the landscapes and environments on different planets are different, the physical laws behind the geologic processes on different planets are the same. This provides the theoretical basis for analogizing in planetary research. Analogizing about different geologic processes (e.g., volcanism, erosion, impact, tectonism) on different planets has been done by different research projects and has played increasingly important role in understanding planetary geology (Anglés & Li, 2017; Bridges et al., 2004; Wheatley et al., 2019). Not only that, analogizing landforms on Earth and other planets has also deepened our understanding of terrestrial processes (Baker, 1993).

With more and more planet exploration missions, the observation and resolutions of data collected about other planets have increased and improved, so this provides an unprecedented opportunity to compare and explore the spatial pattern of characteristics between Earth and other planets from a spatial perspective. In this chapter 3 and chapter 4, I analogized the characteristics of Martian valley network and characteristics of terrestrial streams to infer processes and conditions on Mars. The spatial statistics method can be used to test hypothesis and examine whether the inferences of hypothesis are self-consistent. The empirical
method is built based on the assumption of that dendritic VNs were formed by surface runoff. The assumption is widely accepted by both “warm” and “cold” Mars supporters.

Of course, there are also shortcomings of planetary research from the spatial perspective. Since the comparative planetary analogy from the spatial perspective is the intensely data-driven, it does not directly tell us geologic processes, human interpretation and reasoning is still needed in order to explain and make sense of the pattern revealed by data and spatial statistics. It also highly relies on the selected terrestrial data and research assumptions. For example, Galofre, Jellinek, et al. (2020) recently proposed a hypothesis that Martian VNs were formed by subglacial and fluvial erosion. Although some terrestrial streams were formed by the subglacial erosion, the conclusion in this dissertation research cannot exclude or support Galofre, Jellinek, et al. (2020)’s hypothesis because the research assumptions are different.

6.5 Immature Watershed or Outliers

In the chapter 5, the watersheds with extreme low maturity (maturity values are out of 2 $\sigma$ range) are considered as the outliers. The reason is that the filling function can only reconstruct most hierarchy relations of cells, but cannot reconstruct all the hierarchy relations of cells. Also, very few watersheds are extremely immature. The small number of watershed with low maturity is consistent with this interpretation.

It is possible that the extremely immature watersheds in the results are not caused by the failed reconstruction but by the short surface runoff duration length. There are two alternative explanations. First, the extremely immature watersheds are located in the area where other processes may have operated to disturb (or “frustrate” (Craddock & Howard, 2002)) the self-organization of fluvial processes, e.g., impact cratering and volcanic eruption,
so the watershed did not have the enough time to fully evolve and reach high maturity. It is possible because only small VNAs have the extreme immature watershed. Secondly, the fluvial processes were controlled by the local processes, e.g., hydrothermal processes (Gulick, 1998), and some watersheds had much less surface runoff than others, so some watersheds are extreme immature. These alternative hypotheses could be tested in the future research.
WORKS CITED


