Review of Climate Research and Funding 1993 ~ 2017: A Multinomial Logistic Regression Approach

Y. Odeyemi¹, M. Pollind¹, R. Peeler², K. Nozawa³, D. Vesely², A. Page², C. Rakovski³, and H. El-Askary³,⁴,⁵,⁶

¹Computational and Data Sciences Graduate Program, Schmid College of Science and Technology, Chapman University, Orange, CA 92866, USA
²Research Group, VOXX Analytics, Orange, CA 92843, USA
³Schmid College of Science and Technology, Chapman University, Orange, CA 92866, USA
⁴Center of Excellence in Earth Observing, Chapman University, Orange, CA 92866, USA
⁵Department of Environmental Sciences, Alexandria University, Moharem Bek, Alexandria 21522, Egypt

Received 9 May 2019; revised 10 June 2019; accepted 24 June 2019; published online 30 June 2019

ABSTRACT. This research builds a multinomial regression framework to conduct a meta-analysis of trends in climate research and funding as related to the state of affairs in the last twenty-five years in this area of research. We used a climate research query-based strategy searching the Web of Science, National Science Foundation, Australia Department of Environment and Energy, African Development Bank’s African Climate Change Fund, the Asian Development Bank Climate Change Fund and Australia’s Department of Environment and Energy databases to perform quantitative and qualitative trend analysis. Data were harvested using a web scraper and filtered for the 1993 ~ 2017 window. Comparative analysis was carried out to evaluate the climate research output per continent. Also, we evaluated the role funding plays in the climate research outcomes. Different text processing and mining techniques were used to extract information and data needed for trend analysis and statistical modeling. The text processing revealed trends such as major keywords, key opinion leaders, and individual country’s contribution, monthly and yearly spread of published articles in the climate research domain. From these trends, we engineered some of the variables to build a multinomial regression model to further understand future trends in the climate research space. It is probabilistic in nature with the assumption of no inter correlation between variables, hence outputs are more significant. We found that funding for climate research has been on a steady increase in the last twenty-five years, with the US and European investing hundreds of millions of dollars in alternative and renewable energy. Lastly, the multinomial logistic regression assesses the impact of number of investigators, abstract word count and institution types on the class of grant awarded by NSF.

Keywords: climate change fund, climate research, EU-LIFE program, multinomial logistic regression, National Science Foundation, natural language processing, Paris Climate Agreement

1. Introduction

The socio-economic impact of climate change has been well established over the years with different regions of the world experiencing erratic temperature variation, increase in hurricanes, floods, disease outbreaks, flora and fauna extinction, coral bleaching etc (El-Askary et al., 2014; Perry and Morgan, 2017). The impact of climate change is usually long-term in nature and often lead to slow recovery of the local inhabitants. A case study is the 2005 hurricane Katrina which cost $125 billion in damages to the city of New Orleans in the US which is yet to recover from the havoc of that hurricane.

To address the global issues related to climate change, the United Nations Framework Convention on Climate Change (UNFCCC) established the Paris agreement in 2015 to deal with greenhouse-gas-emissions mitigation, adaptation, and finance (Carattini et al., 2018). The agreement has representatives of 196 member countries as signatories to the agreement. This puts Europe at the epicenter of climate change discourse (Business Insider, 2017). Likewise, at the continental level, countries implemented programs, enacted laws and provided funds and resources for better understanding of climate change modalities. Several governmental organizations were set up to promote different climate related initiatives. In North America, the National Science Foundation (NSF) is responsible for funding non-medical research, the European union’s executive Agency for Small and Medium sized Enterprises (EASME) LIFE program provides financial resources for climate related research in Europe. In Africa, the African Development Bank - Africa Climate Change Fund (AfDB-ACCF) extends financial support to climate affiliated studies. The Asian Development Bank Climate Change Fund (ADB-CCF) and Australia’s Department of Environment and Energy oversee funding of cli-
climate research in Asia and Australia respectively.

The AfDB-ACCF was established by the AfDB in 2014 and has over ten million euros in contributions from various donor governments. The ACCF’s main objective is to help African countries access larger amounts of climate finance and use funds provided more effectively; help African countries to account for climate change in the their growth strategies and policies, by means of upstream diagnostics and providing technical assistance; provide capacity-building in climate change and green growth for African countries and stakeholders at national and regional levels (ACCF, 2015). The Asian Development bank established the Climate Change Fund in 2008 to address the impacts and outcomes of climate change with the primary goal of strengthening support to low-carbon and climate-resilient development in developing member nations. The fund provides financing through four modalities: grant component of investments, technical assistance (stand-alone and piggy-back or linked to loan), and direct charge. It has four components: adaptation, clean energy development, reduced emissions from deforestation and forest degradation, and improved land use management (REDD+ and land use); and climate finance readiness (ADB-CCF, 2018). The EU LIFE program was set up in 1992 to contribute to the implementation, updating and development of EU environmental and climate policy and legislation by co-financing projects with European added value. The LIFE financial instruments have been implemented in five phases, LIFE I (1992 ~ 1995), LIFE II (1996 ~ 1999), LIFE III (2000 ~ 2004, extended till 2006), LIFE plus (2007 ~ 2013) and lately LIFE 2014 ~ 2020. The LIFE 2014 ~ 2020 financial instruments are being managed by the European Investment Bank manages the financial instruments; Natural Capital Financing Facility (NCFF) and Private Finance for Energy Efficiency (PF4EE). The EU through the LIFE program has financed approximately 4000 projects contributing over three billion euros to climate affiliated research (EU-LIFE program, 2014).

Access to funds for climate research has become challenging in recent times even though there have been substantial investments by governments and private initiatives to address some of the challenges synonymous with climate change. Funding plays a significant role in the generation and preparation of data needed for proper analysis and understanding of trends. The outcome of these analyses often shapes theory formulation, hypotheses and in some cases economic policies.

Conversely, the struggle for funding often lead to modification in the research being proposed by scientists (Okagaki and Dean, 2016). Grant governing bodies put premium on proposals that focuses on problem solving methods, such as mitigation of ozone layer depletion, carbon pricing/decarbonization. While it is expected that the grant bodies primary objective is to ensure that funds spent on research will yield immediate outcomes, targeting natural disaster causing indications has led to a streamlined perspective of causes and impacts of climate change.

Without a doubt, funding influences scientific outcomes, it is common practice for scientists to design experiments to investigate and analyze climate change associated indicators such as long-term temperature variation, greenhouse gas effect, mass coral bleaching and ocean acidification. Precursory studies are the mainstay of grant writing and help to cement scientific studies logic. However, focusing on previously studied climatic indication limits our perspective of causal effect relationship of climatic processes. A case study is the impact of climate change on coral reef ecosystems, previous studies did not account for the impact of mutation and adaptation in vulnerability of coral reef ecosystems. Interdisciplinary studies to ascertain new indication associated with coral reef vulnerability and ozone layer depletion are often not at the forefront of grant writing and such inquiry is often account for a smaller percentage of proposed studies. The advent of climate informatics and other predictive tools for data analysis have led to significant breakthroughs in understanding indications and thus leading to better research outcomes.

In this study, we carried out a meta-analysis of funding and general trends in the climate change research space. Also, we built a multinomial logistic regression to stratify the National Science Fellowship grants in the last 25 years.

2. Technical Approach and Methodology

Over the years we have had controversy surrounding the impact of climate change in the public discourse, which has often led to doubt in outcome of climate research (Weingart et al., 2000; Boussalis and Coan, 2016). This research shed light on the types of articles published in scientific literature and give an unbiased assessment of publications and funding of the climate research. This work reviews the spread of articles published in the last twenty-five years and a general outlook on financial grants provided by AfDB-ACCF, ADB-CCF, Australian Department of Environment and Energy, EU-LIFE program and NSF.

2.1. Data Collection and Extraction of Climate Related Articles

We mainly looked at a sub sample of articles scraped from Web of Science (WOS). Grants received from the National Science Foundation (NSF), EU-LIFE program, Australia Department of Environment and Energy were public available. We collected articles from WOS based on key terms associated with climate. We used keywords from the IPCC glossary of terms such as green house, drought, cyclone, glacier, holocene to search and filter for articles related to climate studies (Table 1). We collected data in a time frame window from 1993 to 2017.

For WOS we selected author id, article title, DOI, journal title, keywords, date of publication, publication id, abstract, authors’ full names, authors’ affiliation (universities, research organizations). Open access archived data from the NSF website was simply downloaded. The NSF set included the following: the award instrument, amount given to the grant, the principal investigator, when the grant was awarded, and the abstracts.
Table 1. Search Strategy Syntax for the Data Sources

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Strategy Syntax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web of Science</td>
<td>TI = (‘climate change’ OR ‘global warming’ OR ‘climate variability’ OR ‘precipitation’ OR ‘green house’) AND TS = (‘drought’ OR ‘glacier’ OR ‘latent heat flux’ OR ‘permafrost’ OR ‘nutrient cycle’ OR ‘sea surface temperature’ OR ‘earth science’) OR (‘surface temperature’ OR ‘transpiration’)</td>
</tr>
<tr>
<td>NSF</td>
<td></td>
</tr>
<tr>
<td>EU-LIFE Australia DE&amp;E</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Discretization of the Funding Amount

<table>
<thead>
<tr>
<th>Tier</th>
<th>Tier 3 (T3)</th>
<th>Tier 2 (T2)</th>
<th>Tier 1 (T1)</th>
<th>Tier 0 (T0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount</td>
<td>less than 50,000</td>
<td>50,001 ~ 500,000</td>
<td>500,001 ~ 5,000,000</td>
<td>Greater than 5,000,000</td>
</tr>
</tbody>
</table>

Table 3. Multinomial Regression of Discretized NSF Funding Classification (T1, T2, baseline: T3) Showing Estimate

| Predictors         | Estimate | Std. error | Z value | Pr(>|z|) OR |
|--------------------|----------|------------|---------|------------|
| num_investigators:1| 0.6500   | 0.03       | 24.25   | < 0.001    |
| num_investigators:2| 0.0500   | 0.03       | 1.29    | 0.20       |
| abstract_word_count:1| 0.0010 | 0.00       | 3.97    | < 0.001    |
| abstract_word_count:2| -0.0050 | 0.00      | -10.98  | < 0.001    |
| Type 2:1           | -0.2800  | 0.11       | -2.50   | 0.01       |
| Type 2:2           | -0.0600  | 0.12       | -0.54   | 0.60       |
| Type 3:1           | -0.4100  | 0.21       | -1.98   | 0.04       |
| Type 3:2           | -0.2800  | 0.22       | -1.26   | 0.21       |
| Type 4:1           | -0.1200  | 0.08       | -1.43   | 0.15       |
| Type 4:2           | 0.2300   | 0.08       | 2.89    | 0.004      |

Figure 1. (a) Distribution of research organizations based on Carnegie research classification, (b) Distribution of NSF award grant class.

2.2. Data Exploration, Aggregation and Prepossessing

We looked at descriptive statistics and data structure or exploration to understand the data from both statistical and computational perspective. The descriptive statistics showed the spread of the data set while the data structure shows the data type, composition and dimensionality. Alteryx software was used for aggregation to ensure that the integrity of the data is maintained for onward processing and analysis. The data was cleaned and preprocessed to remove redundant attributes, handle missing values, and remove duplicates. Data was cleaned and processed through Alteryx and R.

2.3. Text Processing and Mining

The abstract attribute was collapsed into several corpora, that is, all the abstracts was condensed to several bodies of text document. The corpora were processed by cleaning the html tags contained in the text (Meyer, 2008). The cleaned corpuses were standardized and modified by removing stop words (example; the, to, and, but, or etc.). The resulting texts were stemmed i.e. stripping of words to its root word (e.g. kicking to kick). After stemming, whitespaces (blank spaces) were stripped to condense the text. Finally, the corpora were preprocessed by transforming all uppercase to lowercase, removal of
numbers, stop word removal, punctuation removal, stemming and whitespace stripping (Meyer, 2008). After a successful text processing, the texts were tokenized and counted.

2.4. Exploratory Data Analysis: Distribution and Rank

Exploratory data analysis (EDA) was visually assessed based on key attributes spread of the trends in the climate change research. Attributes were ranked based on total number of published articles. The attributes of primary interest are journal name, authors’ name (and ID), article title, abstract, year of publication, affiliation and country. Due to the complexity of the data, rank distribution was limited to the “top ten” filter. EDA was done primarily on the amount awarded, year and types of award instruments such as standard grant, fellowship, continuing grant etc.

2.5. The Proposed Imputation Models

MLR is a modeling approach that extends the classical logistic regression technique to handle an outcome variable with k factors (k > 2). MLR selects a baseline category and incorporates (k - 1) logistic regressions to model the association between the log-odds of having an outcome of type versus the baseline type. Mathematically, that can be expressed as follows:

\[ \log \left( \frac{P(Y = s | X)}{P(Y = k | X)} \right) = X \beta_s, s = 1, 2, ..., k - 1 \]  

(1)

2.6. Variable Engineering

Feature engineering was done to select for appropriate predictors and response variables for the multinomial logistic regression model. Authors, university/affiliation, abstract, and the award amount were reengineered into the number of investigators, research organization type/university class (based on the Carnegie research university classification system R1, R2, R3 and non-university research organizations such as Woods Hole Oceanographic Institution were classified as R4), word count of the abstract and amount class (discretized into 3 tiers; T3: less than 50000, T2: 50001 ~ 500000, T1: greater than 500000 in US$) respectively (Table 2 and Figure 1(a) ~ (b)). The predictors include all the variables except the amount class which is the outcome/response variable with a nominal structure. Pearson’s correlation coefficient was used to assess for inter correlation (multicollinearity check) among the predictors, and was done to reduce potential overfitting (Figure 2):

\[ \log \left( \frac{P(Amount = Tier1)}{P(Amount = Tier3)} \right) = b_{11} + b_{12}(\text{res\_org\_type = R1}) + b_{13}(\text{res\_org\_type = R2}) + b_{14}(\text{res\_org\_type = R3}) + b_{15}(\text{abstract\_word\_count}) \]  

(2)

\[ \log \left( \frac{P(Amount = Tier2)}{P(Amount = Tier3)} \right) = b_{21} + b_{22}(\text{res\_org\_type = R1}) + b_{23}(\text{res\_org\_type = R2}) + b_{24}(\text{res\_org\_type = R3}) + b_{25}(\text{abstract\_word\_count}) \]  

(3)

Figure 2. Correlation matrix of the predictors (number of investigators, abstract word count and type of research organization based on Carnegie classification).

3. Results and Discussion

3.1. Exploratory Data Analysis: Distribution and Rank

Between the 1993 ~ 1997 and 1998 ~ 2002, 2013 ~ 2017 there were 14.62, 42.62, 26.55 and 70.38% increase respectively from the previous five years except for 2008 ~ 2012 where there was a 11.27% decrease from the previous five years. The highest number of published articles with approximately 45000 was recorded in June, while January recorded the lowest number of published articles (34000 approx.). The months of June and January recorded a 186.49 and 83.98% increase within the 1993 ~ 2017 timeframe (Figure 3).

Africa, Asia, Australia, Europe, North America, and South America contributing 3.66, 16, 9.55, 26.03, 33.78 and 10.17% of total number of publications, respectively. Africa’s minimal contribution to climate change research can be attributed to the non-existence of a unified climate change policies prior to the establishment of the AfDB-ACCF in 2014.

Africa had the highest increase in the number of climate related published articles 786.45% in a 25 years frame, with Egypt, Nigeria, Algeria, South Africa, Kenya contributing significantly in last 10 years (Figure 4). This is followed by Asia recording a 129.26% increase, China, Japan, South Korea, Iran and India for a significant portion of the total articles published in Africa. South America recorded the third highest number of climate related publications with 115.72% increase. Brazil, Argentina and Chile contributing a larger portion of the South

97
American publications in the last ten years. Europe had a 45.86% increase in the time frame. Consistently, France, Germany, UK, Italy, Spain and Netherlands were the top contributors to the climate scholarly work in Europe. North America recorded a 39.86 increase in published articles. US and Canada have been the major contributors to the total number of publications in North America. Lastly, Australia had the least increase (35.67%) in the total amount of climate affiliated publications. New Zealand and Australia were the two main contributors to climate research in continental Australia. North America and Europe accounted for the largest contributions to climate research in terms of grants and volume of climate related publications. Both continents continue to invest billions of dollars in climate programs, centralized data gathering platforms and exploration of alternative renewal energy to yield a better outcome. While other continents contribution to climate research outcome is significant, however, they are often limited by inadequate financial resources and weak economic policies.

3.2. Assessment of NSF Funding for Climate Research

Between 1993 and 2017, the NSF awarded more ‘standard grants' than any other award instrument, (66.9, 79.6, 73.4 and 73.2%), ‘continuing grant' 29.2, 14.4, 21.6, 21.9 and 35.9%, ‘fellowship grant' 6.0, 8.6, 6.7, 7.2 and 4.3%, and ‘cooperative agreement' recorded less than 5% in the last twenty-five years (Figures 5).

3.3. Amount Awarded by NSF per Award Instrument for Climate Research

The fellowship, continuing and standard grants award instruments had 56.0, 13.3, and 71.1% increase in the amount awarded by NSF while the cooperative agreement saw 82.3% decline between 1993 and 2017 (Figure 6).

3.4. Proportions of Articles published

PLOS ONE, Journal of climate, Geophysical Research Letters, Geophysical Research Atmosphere and Climate Dynamics were consistently among the top 10 journals in terms of number of publications. Other journals such as Scientific Reports, Science of Total Environment and Global Change Biology also published significant amount of publications. In 1993 ~ 2017, Journal of Climate 15.3, 9.0, 11.9, 9.5 and 9.8%; Geophysical Research Letters 11.7, 10.6, 11.2, 11.9 and 10.3%; Climate Dynamics 10.2, 9.7, 9.1, 9.1 and 8.8% of the total articles in top 10 journals by quantity, PLOS ONE had no publications prior to 2006 (year of establishment), it published 15.2% of the articles between 2013 and 2017 (Figure 7).

Figure 5. Distribution of NSF award instrument for climate related research the year 1993 ~ 2017.

Figure 6. Amount awarded by NSF Per Award Instrument for Climate Research (1993 ~ 2017).

3.5. Multinomial Regression Results

The results from the multinomial logistic regression modeling are presented below with details shown in Table 3. They allowed us to identify and assess the effect sizes of the significant predictors of the odds of getting grants of amounts in the ranges from 50 to 500 k and over 500 k versus getting a grant in the range of 0 to 50 k:

\[
\log \left( \frac{P(G \in (50k,.500k])}{P(G \in (0k,.50k])} \right) = -2.73 + 0.65NI + 0.001WC - 0.28I(T2) - 0.413(T3) - 0.12(T4) \tag{4}
\]

\[
\log \left( \frac{P(G \in (500k, Inf])}{P(G \in (0k,.50k])} \right) = -0.21 + 0.05NI - 0.005WC - 0.06I(T2) - 0.28(T3) - 0.23(T4) \tag{5}
\]

We have found that number of investigators, abstract word count and institution types 2 and 3 had significant effects (p-values of 0.001, 0.001, 0.01 and 0.04) on the log-odds of getting a grant in the range between 50 and 500 k versus a grant in the range of 0 to 50 k. In particular, each additional investigator increased the odds of getting a grant in the higher range vs the lower range by 92%, each additional key word in the abstract increased the odds of getting a grant in the higher range.
vs the lower range by 0.5%, the PI being associated with institution of type 2 increased the odds of getting a grant in the higher range vs the lower range by 76% compared to the case where the PI was associated with a type 1 institution, the PI being associated with institution of type 3 increased the odds of getting a grant in the medium range vs the lowest range by 5% compared to the case where the PI was associated with a type 1 institution. In simpler terms, an increase in the number of investigators combined with novel methods and the feasibility of a proposal increase the likelihood of receiving a grant.

We have found that institution type 4 had significant effects (p-value of 0.004) on the log-odds of getting a grant exceeding 500k versus a grant in the range of 0 to 50k. In particular, the principal investigator PI being associated with institution of type 4 increased the odds of getting a grant in the highest range vs the lowest range by 26% compared to the case where the PI was associated with a type 1 institution.

It is interesting to note that the only significant variable in both logistic regression models was the number of words in the abstract even though the direction of the effects were opposite, higher word count increased the odds for the medium grant amount interval and decreased the odds for the highest grant amount interval. Also, the number of investigators and institution types 2 and 3 were significant predictor only for medium grant amount with negative effects while institution types 4 was a significant predictor only for the highest grant amount with a positive effect.

4. Conclusions

The relative low contribution of African and South American countries to climate research may be due to the following factors: inadequate financial capital, incomplete enforcement of local environmental laws and weak coordinated and centralized data sharing platforms among member countries. It is important to note that Africa and South America are the continents with the least contribution to greenhouse effect gas emission, it might be a reason why it might appear as if these countries are apathetic to climate change issues. European and North American countries contribute the most to climate change research due to their long-standing record of enforcement and adherence to climate related policies. Also, these countries have the financial and human resources to address some of the challenges of greenhouse gas emission from industrialization as articulated in the 2016 Paris Accord. The US for example through the environmental Pollution Agency (EPA) enforces the Clean Air Act, the Clean Water Act and other federal environmental laws. In 2008, the Justice Department brought 34 civil cases for violations of the federal environmental laws, with polluters agreeing to pay $36 million in penalties.

From the data analyses, it is apparent that there has been increase in the funding of climate research in the last twenty-five years, however, there still a lot that to be done to achieve the objectives of the Paris Accord. Local authorities for example, can replicate global climate change policies and initiatives.
at a local level, to ensure more financial resources for scientists and regulators to help reduce carbon emission in the coming years. Also, the private sector can invest more in renewable energy research to help reduce local carbon footprint.

In the last twenty-five years, there has been steady increase in global funding of climate research, this can be attributed to the several sensitization and awareness programs set up by IPCC and other local environmental agencies to enlighten governments and the global population on the impacts of climate change on our environment. The results of this study show that more countries need to invest more in renewable energy research needed for a cleaner environment now and in the near future. The MLR model shows that variables such as number of investigators and PI’s research institution significantly influences the amount awarded for climate related research by the NSF. In particular, non-university research institution often has better prospect of receiving top tier (T1) funding than university research institution, this can be attributed to the scope of specific socioeconomic challenges the non-university research organization is trying to address. The collaboration of investigators with diverse background (area of expertise) often increases the chance of getting a T1 grant because it is assumed that the combined experience and skill sets of the co-investigators is an indicator of potential success in the research effort. There is still more to be done in tackling the challenges of climate change in terms of collaboration among countries, institutions and funding agencies. Conclusively, robust access to funding is pivotal to better research outcome in climate research, in turn leads to better understanding of climate change phenomenon. The knowledge from the research outcomes often shape economic policies and laws.

Acknowledgments. The authors acknowledge the National Science Foundation, European LIFE Action Program, African Development Bank Climate Change Fund, Asia Development Bank Climate Change Fund and Australian Department of Environment and Energy for providing access to curated data for this research effort, as well as the anonymous reviewers whose comments and suggestions helped to improve and clarify this manuscript.

References


