University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Masters Theses   Dissertations and Theses

July 2020

A Stylometric Analysis of Climate Change Fiction

Nina Lorenz

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2

Part of the Arts and Humanities Commons

Recommended Citation

A Stylometric Analysis of Climate Change Fiction

A Thesis Presented

by

NINA LORENZ

Submitted to the Graduate School of the University of Massachusetts Amherst in partial

fulfillment of the requirements for the degree of

MASTER OF ARTS

May 2020

English

A Stylometric Analysis of Climate Change Fiction

A Thesis Presented

By

NINA LORENZ

Approved as to style and content by:

_____

Malcolm Sen, Co-chair

_____

TreaAndrea Russworm, Co-chair

_____

Daniel Sack, Member

_____

Daniel Sack, Graduate Program Director
English

_____

Randall Knoper, Head of Department
English

ABSTRACT

A STYLOMETRIC ANALYSIS OF CLIMATE CHANGE FICTION

MAY 2020

NINA LORENZ, B.A., UNIVERSITY OF STUTTGART

M.A., UNIVERSITY OF MASSACHUSETTS

Directed by: Professor TreaAndrea Russworm and Professor Malcolm Sen

This work sets out to analyze stylistic changes in Anthropocene fiction over the past 60 years. The starting point for the analysis has been Rachel Carson, and the presumed beginning of the Anthropocene in the 1960s. The primary insight gained reveals the connections within these novel and relations of similar writing about climate change thereby contributing to the field of Environmental Humanities in a fundamental way, as so far, climate change fiction has only been investigated through a topic centered focus.

The corpus compiled for scrutiny here extends to over 84 novels from these years. These novels have been selected based on a dual approach, looking at the secondary literature as well as a crowdsourced approach in looking at *Good Reads'* cli-fi lists. The resulting texts are then analyzed with stylo, an R package that has been specifically created for stylometric analysis by humanists. The results are visualized in a network that allows easier interpretation and leads to an understanding of more detailed questions about the nature of the connection between works, the inspiration and representation of a specific genre of writing. Moreover, the thesis looks diachronically at clustering based on time and topic. Understanding the ways in which authors address and have addressed climate change is one indicator of how climate change is and has been comprehended.

In terms of the digital approach applied here, the basis is a distant reading approach covering a larger number of novels and rather than close reading them, the task is to find patterns that extend throughout. However, for a thorough analysis, scalable reading is applied to

contextualize and investigate the results in more depth. Overall, the results are meant to establish a baseline for discussing climate change fiction in the Anthropocene which although gaining more scholarly attention still is understudied. The hope is to not only gain insight but to generate visualizations that will provide a helpful resource for fellow scholars.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

THE ANTHROPOCENE, THE TEXT, AND IMAGINARIES: AN INTRODUCTION TO

CLIMATE CHANGE FICTION THROUGH CONTEXT

The Anthropocene is a curious time. Both propelled by humanity and dictating the

contemporary condition for humanity, the Anthropocene is the framework that supersedes all.

With regards to novels this translates into all fiction becoming Anthropocene fiction, maintains

scholar Adam Trexler. The notion goes back to Roland Barthes who states that both author and

reader of a text are products of the cumulative texts, life experiences, and world knowledges they

have acquired in their lifetimes. In other words, they write, read and interpret based on everything

they have ever learned. Thus, in the Anthropocene, writing cannot exist outside of climate change

and is always a narrative about the climate crisis. To comprehend the Anthropocene is one key to

understanding the potential for imaginaries that shape life at the present moment. Before

analyzing fiction in the Anthropocene, as is the objective in this thesis, a definition of the

Anthropocene is required.

According to the reference work *Merriam Webster*, the "Anthropocene" is "the period of

time during which human activities have had an environmental impact on the Earth regarded as

constituting a distinct geological age" ("Anthropocene", *Merriam Webster*). This vague

description is characteristic for the keyword that has become so common in academic disciplines.

The origin of the term aimed at describing a new phenomenon measured by differences in the

earth's geology, but the precise account has proven much more complex. Introduced by geologist

Paul Crutzen, the Anthropocene was initially dated to the beginning of the invention of the steam

engine, the start of Industrialization in Great Britain in 1784. However, disciplinary conventions

in geology demand that the proof come "from a remote, pristine location in the Southern

Hemisphere that occurs at the same time as the north" to ensure a global change in opposition to a

localized event (Turney, Palmer, Maslin). Under these circumstances Crutzen has since revised

his initial thesis, as there was no significant spike in carbon dioxide and methane in geological strata dating to suggest the industrialization was the onset. Together with Steffen, Crutzen proposed a later era, the Great Acceleration, "when human impacts on our planet suddenly intensified and became global in extent," seemed logical due to its rapid growth of technology and capitalism. This period is less favored for spikes in carbon dioxide or methane but instead focuses on traces of nuclear testing. "However, while its trace will remain in geological records, the evidence of large-scale shifts in Earth System functioning prior to 1950 is weak" (Steffen et al.). Such findings have eliminated the most promising contenders that were 1945, the end of the world war, and the first nuclear test, the "Trinity" detonated in New Mexico. At present that "golden spike" which is "the location of a global marker of an event in stratigraphic material" proving the planet's altered geological system has not yet been found (Lewis and Maslin).

With a peak in nuclear testing and the cold war, the early 1960s have a high potential to be the turning point, which succeeds the nuclear testing of the 50s and is likely to have created a lasting impact. So far, the markers of stratospheric change have been found predominantly in the northern hemisphere. Trying to find a spike in both hemispheres, Lewis and Maslin found that atmospheric carbon dioxide "reaches a maximum in the mid- to high-latitude Northern Hemisphere at 1963–64 and a year later in the tropic". And a more recent discovery based on radioactive indicators "suggests it began in 1965" (Turney, Palmer, Maslin). It appears then that the most logical and promising onset is in the 1960s and thus will pose the onset for the analysis within this thesis. However, it should be mentioned that scientists have followed different strategies for finding markers through other events, such as the plague (Ruddman), animal extinction, and most notably colonialism. The Orbis spike in 1610 marks changes in the atmospheric carbon dioxide concentration that is linked to the establishment of colonialism and its consequences. The Orbis spike is highly interesting, but it will not be the basis of investigation for this thesis as the premise is an investigation of fiction writing which only occurs in a meaningful frequency later.

While geologists are researching the stratospheric and atmospheric markers to define the Anthropocene, academia has used the term and defined it in the context of history and culture. Scholars outside of geology adopt the term Anthropocene in order to point out the human agency in the current climate crisis, failing to acknowledge the problems that coincide with claims of responsibility and ultimately blame. Specifically, it is not all of humanity that equally has destroyed the planet's ecosystems, writes Dipesh Chakrabarty on the limitations of thinking about man-made climate change through the term Anthropogenic. He criticizes the overemphasis of human agency, with his arguments based on the planetary history and the different scales of time that are the earth's "deep time" versus the time humanity has impacted the planet. He concludes that the term implies a form of control over the climate on this planet that is not exclusively human; it conflates the impact of different responsibilities and it overlooks non-human agents and planetary development.

Considerations like this render the Anthropocene in its cultural dimensions no less evasive and at times problematic than in its geological definition. In fact, language at present appears unable to fully express to a satisfying degree all the complexities of the Anthropocene (Clark 2013). And still, a plethora of writing and theories exists, trying to create "manageable, if arguably evasive, frames for predictability and conceptualization" (Clark 17). A philosophical ontology that aids in the theorization of, among other things, climate change is Timothy Morton's *Hyperobjects*, which "refer to things that are massively distributed in time and space relative to humans" (Morton 1). These hyperobjects "are nonlocal; in other words, any 'local manifestation' of a hyperobject is not directly the hyperobject" (Morton 2). Moreover, "hyperobjects provoke irreductionist thinking, […] in which ontotheological statements about which thing is the most real (ecosystem, world, environment, or conversely, individual) become impossible" (Morton 19). Morton identified the problem as one of scale both temporally and spatially. An appropriate account therefore requires one to zoom out to exceed the writing and imaginative capabilities of humans. Translating this into the stakes for climate change fiction and thinking about the

Anthropocene, Morton's approach dictates not that there can never be accounts of the hyperobject, but there can never be a universally true account of it. Navigating these dimensions, Timothy Clark proposes to address these hyperobjects by "mapping out the contradictory implications of reading the same text or issue on several very different scales of space and time" (24). The repetition in various contexts is the closest approach to scaling of dimensions. This cannot solve the irreductionist nature of the hyperobject that is climate change; however, it appears a viable and tangible strategy in spite of this elusive phenomenon. Looking at climate change fiction is a way to productively map a scalable account that repeats the contradictory nature of climate change. The work of climate change fiction, in other words, is addressing the unseen and often unknown in both the future and the present. Fiction has the capacity to "provide a space in which to explore the complexity of sustainability as an ongoing, never fulfilled aspiration" (Johns-Putra et al. 6). Literature then plays on the border of human comprehension in order to "address our sentimental, affective, habitual and non-cognitive comportment towards the world" (Johns-Putra et al. 115). In other words, Clark proposes that the novel allows what-if scenarios, essentially revisiting complex questions .

The focus on the connection between climate and fiction for this thesis is influenced by the theories that have analyzed the imaginary of the Anthropocene. Amitav Ghosh, Elizabeth DeLoughrey, and Ursula Heise, among others, have theorized the relation of the novel, fictional writings, and climate change. In *Imagining Extinction*, Ursula Heise uses the sixth mass extinction as a backdrop for analyzing the ways cultural frameworks shape humanity's capacity to think and act. She analyzes environmental conservation efforts both in policy and science to emphasize how "imaginative webs" and "image-making" drive even these seemingly fact-based disciplines and the actions taken to preserve biodiversity (Heise 5-11). Imagination in this context is classified by its limitations and its inability to reach a point of exceeding one's contextual and cultural knowledge. Refocusing this approach Amitav Ghosh discusses the structural limitations that have arisen in literature since the Great Acceleration. He links the focus of literature on

moral individual narratives to the religious, imperialist, and capitalist structures in society at present "for let us make no mistake: the climate crisis is also a crisis of culture, and thus of imagination" (Ghosh). While acknowledging that there is climate change fiction Ghosh argues that these are still only a marginal part of all literature written today. This is rooted in a literary tradition that overshadows the exceptional with the everyday. Imagination adheres to present systematics and when strong enough to question and criticize these structures, it attacks first at the level of politics and state instead of the climate crisis. When understanding imagination as the capacity to form new ideas or concepts in an unexplored manner it appears Ghosh and Heise are right. The large bulk of climate change narratives present no revolutionary concepts. In this light the present scholarship and literature appears stuck in what Elizabeth DeLoughrey calls the"Anthropocene cultural imaginary" that is representative of the global (white) north. DeLoughrey includes Postcolonial works and indigenous representations of climate change allegories. Imagination in DeLoughrey's approach is similar to Heise and Ghosh; however she analyses works within a different imaginary, united by a set of values unlike the ones addressed by Ghosh and Heise. To summarize these different imaginaries, when it comes to climate change the problem appears to circle back to limitations and lack scalability of imagination in face of the hyperobject. It is unlikely that the next few years will see a radical new mode of thinking about the climate crisis, but perhaps this is not needed. If the acceptance of an imaginary is the frame, climate fiction is a form of play. A form of playing with the unknown and unseen and thus constitutes a form of world-making. Playing in this sense is different from imagining as it does not center on unique creative process but essentially is a flushing out of impact scenarios, of stakes, providing an arguably "safe" space for authors to play and replay any possible present of future scenes in the climate crisis. The space is safe as the consequences are, despite all realistic potential, still imagined.

The respective depictions of the imaginative structures inevitably lead to a discussion of their potential of addressing matters of climate change. Hulme argues "climate - as it is imagined

and acted upon - needs to be understood, first and foremost, culturally and that the environmental humanities can enrich and deepen such an understanding" (175). Grasping climate change, is to view all its dimensions from the visible changes to what Rob Nixon calls slow violence, "a violence that occurs gradually and out of sight, a violence of delayed destruction that is dispersed across time and space" (2). Culturally understanding climate change in other words responds not only to the individual narratives but also relates them to an audience that has the potential to be transformed. To this end Antonia Mehnert writes hopeful "climate change fictions play an important role in shaping not only our understanding of this unprecedented crisis but also in guiding our responses to and actions against it" (16). While it may be a romanticized notion of the novel sparking immediate activism, a potential for flushing out the different responses to the changing ecologies can point towards reasonable adaptive behavior and possible solutions to individual problems. Understanding the Anthropocene through modes of affect and responsibility is a popular trend that must be viewed and dissected, concurs Adelaine Johns-Putra. Johns-Putra points to a notion that circulates through environmental humanities, the balancing of accessibility and questions of stakes versus the avoidance of oversimplification for the sake of identification.

In his empirical study Matthew Schneider-Mayerson partly confirms Johns-Putra's intuitions about the effects of climate change fiction while adding a refined concrete understanding to the debate. In the experiment readers responded to the most popular cli-fi novels published after 2000, as readers are likely to have read at least a small fraction. The results reveal readers of climate change fiction as usually self-identified liberal millennials. While the study's finding does not exclude conservative readers of these novels, "ecocritics and environmentalists who hope that climate fiction might convert conservative climate deniers, [...] might be disappoint[ed]" (479). The implication is less radical, readers with an increased awareness might be nudged into "to tak[ing] individual, consumer, and political action" (Schneider-Mayerson 480). The study confirms the results of similar studies conducted after the boom of apocalypse movies in the 2000s. In her study of viewers of the movie *Age of Stupid* Rachel Howell found that

the audience was in an overwhelming majority (77%) alarmed about the current state of the planet and 60% cited this as the reason for seeking out the movie. Both studies show the audience as open to learning about climate change effects prior to engaging with the narratives. In terms of content, Schneider-Mayerson found temporality to become a thought-provoking dimension that resonated with readers. The negotiation between looking ahead and discussing the present successfully illustrated the invisible changes occurring as a result of climate change and left their readers with a renewed receptiveness for challenges resulting from slow changes. This is especially true for familiar settings that create identification with a reader and thus a closer emotional attachment to the fate that befalls these places. Beyond that, it is the stories of success that readers want to engage with, as Schneider-Mayerson found, "positive emotions tended to refer to a pleasant if vague orientation toward the future (e.g., "inspiring"), negative emotions were often intense, immediate, and self-directed" (489). Such a tendency has been proven time and time again, explains Eco-psychologist Nikki, people "want to, and can, live the story. Good stories, stories that will help propel us toward a sustainable world, compel imitation" (Harré 52-53). By implication, this means "the psychological tendency to avoid stories that deliver negative emotions means that well-intentioned authors who vividly depict the catastrophic consequences of climate change may actually be hindering their goal of heightening environmental consciousness" (Schneider-Mayerson 490). This does not exclude hardship as a result of climate change but the most impactful narratives depict a brighter future and any relatable action to get there. Rachel Howell explains how "[p]eople need to believe that they can do something about the problem and that it is worth doing something" (3). In other words, the narratives that instill cause for hope yield the most tangible impact in subsequent behavior rendering them worth studying. Following this argument, the necessity to further explore climate change fiction becomes indisputable.

To fully understand the potential of climate change fiction, a clarification and definition is necessary, starting with a disassociation from Anthropocene fiction. Introduced by Adam

Trexler, the latter term is used for fiction in the Anthropocene and rests on the assumption that the Anthropocene as a condition influences all cultural production and therefore all literature that is written in this time. This leaves the term too unspecific as it would encompass all novels of the present – a task that cannot be achieved here. Therefore, not all novels that are anthropogenic novels are expedient to the results of this study. More importantly here, the items processed need to have a tangible climate change motif. This thesis will use the term climate change fiction for novels containing any known relations and consequences of the Anthropocene, changing climates and so forth where this content and purpose are driving factors for the narrative. To concretize these examples, the overview of content structures that Adam Trexler provides may help to further clarify the differences between Anthropocene fictions and climate change fictions. Working with these categories serves as an illustration but should be seen as providing a comprehensive view of all possible Anthropocene fiction.

The first group of topics, "Place: Deluge, Floods and Absence," is a broad subsection, concerning works in which a climactic event or simply a natural disaster leaves the characters struggling either in the aftermath of or living through said event. Floods and the motif of water pose long-standing fears that threaten humanity throughout many imaginaries. However, the collection for this thesis is not limited to floods but equally contains narratives of droughts and manmade Armageddon such as nuclear extinction. While environmental social scientists like Andrew Ross and Mike Hulme emphasize the important distinction between weather and climate, these do not help in identifying human caused events. In the age of the Anthropocene these will be treated as effects of the climate crisis and therefore as climate fiction. Similarly, the "eco-nomic novels" which "account for the nonhumans and distributed populations of the Anthropocene" (Trexler 172) are part of the corpus. This subgenre then accordingly includes questions of animality and gives room to reconsider a human-centered perception of the world. In these novels, the displacement of a natural order is mourned and the disequilibrium between actors and those affected is brought into focus. When defining climate fiction through content, it

8

follows narrations relating impacts of climate change on human or non-human is a climate change novel and this is true regardless of the scientific accuracy or proof.

To juxtapose this definition, stories that digress from natural events and displacement are Anthropocene fictions. This subsumes novels discussing or questioning science and research, which also feature struggle with culture, information and ethics. The narratives are focused on the politics and bureaucracy and potentially contradict popular science, governments or companies. Examples within these types of novels address fears "that scientists unwittingly smuggled contestable value judgments about how the world (supposedly) is, and how it ought to be, in the guise of 'objective' and 'rational' representations of human biology or the nonhuman world" (Castree 152). However, they may also subvert or challenge belief systems and speak for example to Post-colonial ecologies (DeLoughrey). When looking at DeLoughrey's analysis of anthropogenic imaginaries, it becomes clear that these are certainly flushing out and playing with the values of society. However, they do not show a tangible textual representation of climate change and are thus examples of anthropogenic fiction.

Lastly, this section needs to address the special role of science fiction, sci fi, or cli-fi. The genre's traditions lend themselves to scalability, as Canavan and Robinson explain, "SF is our culture's vast, shared, polyvocal archive of the possible; […] and begins to imagine what our transformed planet might eventually be like for those who will come live on it" (17). The convention of imagining (and playing with) the future gives the genre an advantage at speculating an informed future that is rooted in technological and scientific developments. In the preface to a collection of climate fiction short stories, Elizabeth Kolbert declares, "the characters in these stories are made up and the situations invented. The events haven't happened and, in a strict sense, never will. But the science behind these tales is all too real" (Preface). This claim to scientific accuracy combined the seemingly distant fantasy of science fiction works as a mirror for present tense observations that cannot omit the transformed lived experience of climate change. The argument here focuses specifically on science fiction as futuristic imaginations at a

9

moment in time where the future appears doomed. It is important to acknowledge the value of science fiction in this context not solely for the many books that unite the genre and a climate change narrative but also because the genre is less often praised in an academic literary context. Among others, Amitav Ghosh has been criticized for his focus on literature as opposed to popular fiction thereby excluding a wide range of relevant writings (Shelby, Milner and Burgman). Ghosh's call for more climate change literature is based on a definition of literature similar to Claire Colebrook's, "literature is literature only with a certain non-biodegradability: unlike formal systems or even highly unremarkable texts that appear and then disappear, literature sustains itself" (177). Alleging that the current book market does not live up to this standard – and there is hardly a counter argument, the amount of undoubtable longstanding literary works is small, usually it extends to writers like Kim Stanley Robinson, Margaret Atwood, Barbara Kingsolver, Ghosh himself, Ursula Le Guin, Paolo Bacigalupi, and at times Cormac McCarthy. Certainly, this is not a long list when comparing it to the overall classics of literature that are so familiar. Further explaining his criticism Ghosh states "cli-fi is made up mostly of disaster stories set in the future […]. The future is but one aspect of the age of human-induced global warming: it also includes the recent past, and, most significantly, the present" (Ghosh). His argument is framed by his aspirations for imagination and a desire for revolutionary writings.

In contrast, literary scholar Shelby Streeby maintains "people of color and Indigenous people use science fiction and other speculative genres to remember the past and imagine futures that help us think critically about the present and connect climate change to social movements" (Streeby 5). While Ghosh's understanding of cli-fi is, in a way accurate, his negation of their climate change narratives as simplistic and one-sided accounts disregards their value as accessible contributions to climate change communication. To relate the importance of this debate to this thesis, the wide definition of climate change novels is meant to include popular fiction, perhaps more so than literature as the interest are the wide-ranging implications for thinking about climate change which includes an overwhelmingly high number of science fiction and speculative fiction.

10

A simplification and representation cannot be completely overcome, hence, the hyperobjectivity of climate change that Morton describes, applies to all aspects of writing among these questions. The irreductionist nature of the hyperobject that is climate change creates a paradox, which must be addressed culturally and yet cannot be addressed comprehensively. This leaves only one option for the literary scholar, to study any and every piece of writing that could become relevant and assemble the pieces. Accordingly, one needs both literature that stands the test of time, and popular fiction that focuses on immediate singular aspects. Mapping a timeline of both, constants and a progression of characteristic struggles at different points in time, the premise of this thesis is a firmly held believe that neither is more important than the other. The writings are working together to create a whole that reflects the hive mind of a society and its imaginary.

An inductive study, the point of investigation is not to explain or even prove a given theory. Instead this thesis posits a first exploration into the patterns within climate change fiction. The initial question was a simple one: how do we write about climate change? Developing the curiosity into an actual research question this thesis asks whether or not style for these climate change novels creates viable subgroups or whether they changed diachronically. Beyond that it challenges the factual difference between serious fiction, speculative fiction and science fiction. These premises created a set of frames for compiling a dataset. The first limitation excludes non-fictional writings, assuming that the style of journalistic, scientific, and academic writing respectively is largely stagnant. Whereas fictional writings are free to use different stylistic approaches that manifest in the writing and create a measurable signal. To isolate a signal among the noise, it is vital to create further restrictions on the corpus rendering data within a certain frame. This frame was set to a shared topic (climate change), a shared language (English), a shared medium (fictional text), and a shared time frame (1960-2019).

Deciding to apply a stylometric analysis though not straightforward results from a curiosity into the methods that create imaginaries. The goal is to have a closer understanding of the limitations and therefore contribute to knowledge about the approaches taken to the climate

crisis. Scholarship that recognizes the structures of imaginaries helps to open new methods yet unexplored in speaking about the climate crisis. Moreover, the stylometric approach was preferred to the content analysis firstly, because current scholarship by Adam Trexler and more recently Gregers Andersen are addressing these questions. Though their analyses are not quantitative, they include a remarkably large number of novels. The insight to be gained from applying topic modelling would thus be minimal, and the contribution is arguably not of a higher quality simply because it is computational. In fact, it may be potentially misguiding as computational and statistical analysis are not irrevocable truth tellers. The value of using quantitative methods is a way of finding structures, patterns and an attempt to bring order to an inherently diverse field and topic. Seeing the strength of the method in its scalability, it is a simplification and cannot replace the work that has been accomplished in the field already. This is therefore an acknowledgement of the considerable brilliance that has already gone into writing, researching, and thinking about climate change. Secondly, the stylometric analysis has begun to overlap in unexpected ways with the content of the novels, highlighting a correlation between style and content not foreseen. These correlations will be explained in more detail after introducing the methodology and detailing the metrics of the study in the next chapter. It is meant to provide transparency but also as a guideline for scholars that are not yet familiar with quantitative textual analysis. Finally, the fourth chapter will relay the results and conclusions to be drawn from the data.

CHAPTER 2

TALKING STYLE: THE BASICS OF STYLOMETRIC ANALYSIS

Aiming at a clarification of stakes, this section provides transparency for the paradigms

of stylometric analysis, including the corpus creation. As the methods are text based and rely on

quantitative textual analysis, it seems appropriate to begin by recalling the fundamentals of these

approaches. A name that resonates in this field is Franco Moretti, known for his approaches to

distant reading. According to Moretti the digital affords a new type of literary critique that allows

large scale analysis which, rather than focusing on specific texts, detects patterns across many

texts. The practice of distant reading and "the great unread" (Margaret Cohen) offers a substantial

framework for situating texts in a larger set of other writings. Moreover, this allows an equal

influence from works that are not considered part of the canon, meaning all works included in a

corpus have equal weight on the results. This nomothetic approach to studying literature allows

new ways of thinking about literature analysis. Digital humanist Matthew Jockers gives an

example of potential questions for investigations into these corpora:

- Whether there are stylistic patterns inherent to particular genres

- Whether style is nationally determined

- Whether and how trends in one nation's literature affect those of another

- The extent to which subgenres reflect the larger genres of which they are a subset

- Whether literary trends correlate with historical events

- Whether the literature that a nation or region produces is a function of demographics,
  time, population, degrees of relative freedom, degrees of relative education, and so
  on

- Whether the literature is evolutionary

- Whether successful works of literature inspire schools or traditions

- Whether there are differences between canonical authors and those who have been traditionally marginalized
- Whether factors such as gender, ethnicity, and nationality directly influence style and content in literature

It is apparent that these are radically different from questions conventionally asked by literary scholars. These baseline ideas provide a helpful guide for beginners in quantitative textual analysis or even digital textual analysis and have inspired the considerations for this analysis. Looking at climate fiction these become questions such as whether or not style in climate fiction changes over time, whether there is an inherent style to climate change fiction and whether style differentiates serious "canonical" climate fiction from popular examples. With the quantitative analysis as vital part of looking at the texts in question some preliminary decisions had to be made for the novel selection.

Compiling a valid corpus becomes the first critical paradigm as it sets the quality for the results. Therefore, the creation of the corpus is dependent on its quality of data, that is useable texts, and equally if not more on the relevance of the data. The first limitation restricts the timeframe to novels written and published post 1960. As has been discussed in Chapter 1 the Anthropocene provides the framework for these novels. Moreover, this is a reasonable beginning for data collection, with a meaningful sample size. The second restriction regarding content becomes harder to navigate, seeing as climate change fiction is a rather vague term that ideally is decided on an individual level. Yet this would require a prior reading and analysis of too many novels, therefore it has been decided to use the internet's "hive mind". The starting point was provided by Good Reads, a social cataloging website that is dedicated to books and their reviews, whose participants have created a list of climate change fiction novels (determined by vote). Additional information was provided by scholarly works of Amitav Ghosh, Antonia Mehnert, Gregers Andersen, Shelby Streeby and Adam Trexler. Starting with these lists ensures the inclusion of popular climate change fiction, and while these are not free of bias nor

comprehensive, it omits the personal bias of a singular interpretation. The list of novels compiled from these sources were filtered to novels that are relevant for the context, that is narrative non-fiction was excluded as were novels outside of the specified time span. Lastly, novels that have been first published in a language other than English have been eliminated. This is an unfortunate but necessary restraint of the analysis, as the focus is on style and these novels would render the translator's style rather than the author's. Additionally, this would open a question of representation, as there certainly are several climate change fictions in various languages. Therefore, this limit is imposed for validity purposes. Results that come from this corpus then, can only address conclusions that apply to the English-speaking world. Lastly, the corpus is limited to popular fiction and literature, which often neglects marginalized writers and communities. While there is some scholarly work dedicated specifically to this matter (Streeby), the inclusion of novels discussed in these works does not rectify this imbalance. Any conclusion that is thus drawn from this corpus speaks to a generalized theory of popular canonized climate change fiction not to idiosyncratic cases that are no less vital to the study of the field.

Overall, after narrowing down the list according to the aforementioned criteria, of the 258 originally selected novels 84 have been obtained and pose the basis in this thesis[1]. The dataset contains 7 novels from the 1960s, 6 from the 1970s, 9 from the 1980s, 9 from the 1990s, 10 from the 2000s, and a total of 43 from the 2010s. Preparing these texts for analysis required some minimal preprocessing. Mostly, this was limited to conversions into txt a standard procedure in digital humanities. Depending on their format, different protocols applied.[2]

Once the novels were in the appropriate format, the stylometric analysis could begin. Before the technical steps are detailed, a few words about the method. Stylometry is the analysis of style and rests on the assumption that texts and authors have a measurable signal other than

---

[1] At this point I want to thank my interns Rhane Mazzola and Nicolas Ribadeneyra for tireless efforts to find and convert such a large nu

[2] Instructions and details can be found in Appendix C.

content, based on their language use. The first stylometric analyses employed these techniques for authorship attribution, some of the most known attributions are the *Federalist Papers* to Madison more than Hamilton (Mosteller and Wallace) and in a more recent example J.K. Rowling was exposed as the author of *The Cuckoo's Calling* (Juola). However, the uses of stylometry have further expanded to forensic linguistics, genre detection (Al-Yahya 2018), co-authorship detection (Eisen et al 2017), chronology and gender.

In its simplest form the method reveals how closely related the texts in question are in terms of style. Texts recognized as close are not necessarily by the same author, rather linguists assume that language and written style are influenced by multiple factors. The statistics measured in stylometry are an indication of similarities and cannot reveal a causal relation such as authorship, plagiarism, mentoring or generational conventions; the conclusions drawn require further knowledge and insight of the researcher.

Computationally, style is measured by the words that are in the text. The most frequent words that occur are function words, used not for their meaning but for correct grammar. These function words have the potential to influence style subconsciously in a surprisingly meaningful way. Function words are a closed word category, that is the number of words within this group is traditionally fixed and limited. The opposite is true for open word groups, that can be easily appended and thus provide an unlimited word pool. Because of their variant nature, special vocabulary makes up only a small portion of our language yet is likely one of the first transferred variables when copying or emulating writing. Therefore, lexemes are not a reliable signal for stylistic analysis. Rather it is the function words, that are used with less caution that provide strong, quantifiable signals. Computer linguist Kestemont relates the merits of working with function words as follows:

- All authors writing in the same language and period are bound to use the very same function words. Function words are therefore a reliable base for textual comparison

- Their high frequency makes them interesting from a quantitative point of view, since we have many observations for them

- The use of function words is not strongly affected by a text's topic or genre: the use of the article 'the', for instance, is unlikely to be influenced by a text's topic

- The use of function words seems less under an author's conscious control during the writing process

As illustrated by this list, function words provide strong signals in investigations of style. Other methods such as n-grams, named entities, and Part-of Speech (POS) tags are examples of measurements that are equally interesting when investigating style (Eder, Piasecki, Walkowiak 3). N-grams have been shown to generate robust results, as they increase the amount of usable data points (Kestemont). For this thesis, most frequent words, n-grams, and the POS have been used; the most relevant plots are shown in the next section, while the remaining results are posted in the appendix. After choosing the parameters, the technique needs to be chosen. These can be grouped into "explanatory (or unsupervised) machine learning techniques, supplemented by simple visualizations such as dendrograms or scatterplots, and supervised techniques, said to be very accurate, albeit counterintuitive" as Eder, Piasecki and Walkowiak, creators of the stylo package explain (4). Unsupervised techniques are used in Principal Components Analysis (PCA), Cluster Analysis, or Consensus Trees and are popular because they compound the differences between texts into a singular number that is easier to read for the final outcome (Eder, Piasecki, Walkowiak 4). The basic understanding here is that interesting outliers or large clusters will form and offer themselves up for further investigations. In this thesis these methods are used in the creation of consensus trees and PCAs that visualize the stylistic difference between texts by distance.

For this thesis the analysis was conducted with the R package stylo, a package specifically created by computational linguists for the analysis of style. For a first overview of the data, consensus trees were created for each decade which helped to finetune the parameters. The

resulting data was then visualized in PCA, Cluster Analysis, or Gephi, a program specifically designed for network graphs, creating a more pleasing and legible graph. Similar to the consensus tree, a cluster analysis illustrated more clearly whether or not parameters were well-adjusted as novels written by the same author in a series should be grouped together. The setup allowed for actual identification of interesting patterns that were robust, meaning they reoccurred throughout the analysis. Choosing this method as an appropriate yet new approach to studying climate change fiction uses the affordances of stylometic analysis to "find similarities and differences between single documents and groups of documents, as well as tracing annotated documents over a timeline" (Eder, Piasecki, Walkowiak 4). Limitations within this frame of analysis are first based on the statistical sampling, viewing the corpus as statistical sample cannot be representative over all climate change fiction but rather representative for the specified criteria. Overall, the findings in this method require both computational linguistic skills and knowledge of the corpus in order to interpret the findings which unlike traditional analytics cannot simply be broken down into numbers. The results from this analysis trace patterns, which are indicative of a trend within the corpus and give insight into the writing styles of the novels.

CHAPTER 3

ALL ABOUT THE DATA: LOOKING AT THE NOVELS AND RELATING THE FINDINGS

To relate the results as logically as possible, this section will walk through each decade first and then relate the patterns to the ensuing decades to progressively increase the number of analyzed novels. Beginning with the first novels from the 1960s, a consensus tree of these early novels can be seen in Figure 1. The algorithm groups together stylistically similar texts, meaning distance is the most important measure here. The results are based on the 500-700 most frequent words, which are determined in three iterations over the corpus. The statistical distance model applied here and for all other consensus trees was Eder's delta, which offers the additional benefit of weighting more frequent words stronger than words that occur less frequently. Upon inspection three, potentially four, directions outline themselves: the first one is J.G. Ballard's *The Drought* and Rachel Carson's *Silent Spring*; the second direction connects Harry Harrison's *Make Room! Make Room!* and *The Left Hand of Darkness* by Ursula Le Guin; the third consists of Philip Dick's *Do Androids Dream of Electric Sheep?* and *Three Stigmata of Palmer Eldritch*; lastly, Frank Herbert's *Dune* creates a fourth node. This split is repeated with the part-of-speech tagged corpus and appears to be a robust trend within the data. Superficially, the reason is clear, novels similar in style are on the same branch. The distance between branches speaks to the extent of the difference in style. In the case of Dick, this is an obvious and logical conclusion as one author would have a consistent style (when compared to others). Identifying similarities between Carson and Ballard sparks curiosity. *The Drought* is primarily concerned with scarceness of water and fits into the post-apocalyptic, experimental science fiction. In contrast, *Silent Spring* relays the effects of pesticides and insecticides on ecologies; traditionally it is seen a science book. The common denominator in the content, other than destruction of environment, is the agency of large

companies and industry as such. At this point such an inquiry into content is only an observation and cannot determine a cause for similar style.

Looking at Harrison and Le Guin gives more insight into the consistency of this pattern. *Make Room! Make Room!* is a science fiction novel set in an overpopulated New York, in which Harrison explores effects of unchecked growth on society and culture through the eyes of the protagonist. In *The Left Hand of Darkness*, the novel's setting on imagined planets with imagined races is vastly different. Le Guin explores a fantasy that ultimately investigates a society through the main character, showing how the novels share an important skeleton for their stories. Thus, at least for these two clusters, it seems that content and style are connected, which so far has not been proven nor been extensively discussed in stylometric scholarship.

The cluster of Dick's novels of Dick, another science fiction writer, is comprised of *Do Androids Dream of Electric Sheep?* a post-apocalyptic fiction that is set in a world destroyed by nuclear war. The novel discusses multiple themes, among them robots, androids, war, and species extinction. *Three Stigmata of Palmer Eldritch* tells the story of life in an uninhabitable, hot world, from which the only escapes are intergalactic travel and hallucinatory drugs. Lastly, *Dune* describes a power struggle for the wasteland of a fictional planet. A key token is a drug which grants abilities such as foresight and multidimensional awareness. In terms of content it is logically placed between *Three Stigmata* and *The Drought*. The conclusion to be drawn from this data, indicates a strong correlation between style and content, as well as distinct subgroups in climate fiction.

Contextually, the identified clusters are not surprising considering the 1960s are a time of change with the cold war threatening to drive the world to nuclear extinction. Paradoxically, at the same moment the space race results in Neil Armstrong's moon landing, crossing the final frontier. Lastly, increasing CO2 concentrations due to fossil fuel emissions are proven by Charles Keeling and the Moana Records. The novels and the clusters reflect the upheaval and challenges of their time in three distinct groups concerned respectively with pollution of the environment by

unchecked companies, existence threatening war, and an escape into space. However, it is unclear how these developments influence word choices, as the measures are primarily based on function words more than unique vocabulary.
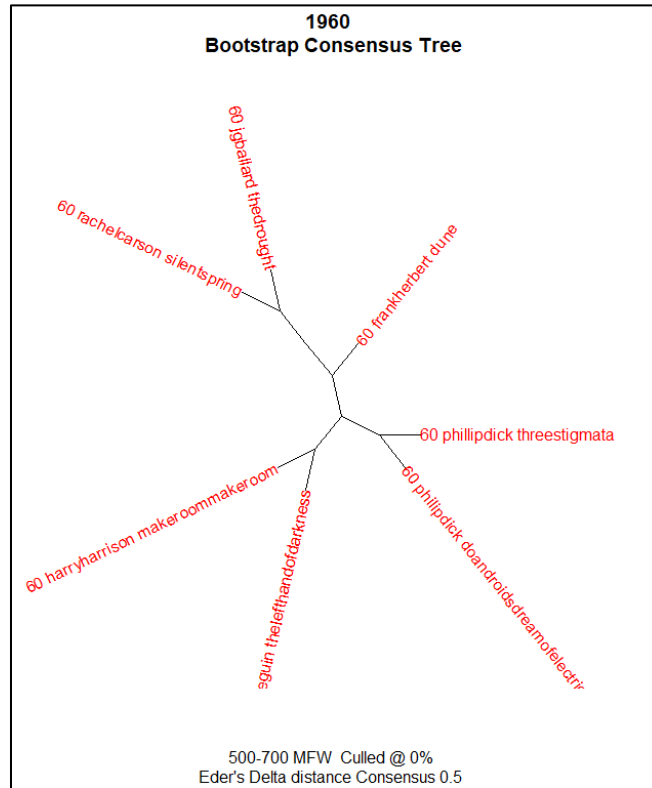
Figure 1: Consensus Tree of the 1960s corpus.

To further understand the characteristics of these subgroups more data might help; hence a glance at the results for the following decade posits a reasonable and logical next step. Here the following consistent clusters have emerged: Edward Abbey as a sole branch, a connection between Kate Wilhelm and Chelsea Quinn Yarbro, one cluster separating Ursula Le Guin's *The Word for World is Forest*, and one connecting Le Guin and John Brunner. Abbey's *The Monkey Wrench Gang* stands alone in this graph as the only non-science fiction work that tells the story of a group of activists rallying against a destructive system. *False Dawn* and *Where Late the Sweet Bird Sang* are both post-apocalyptic science fiction novels. In *False Dawn* an environmental collapse drives the narrative and the ensuing fight for survival of two characters. Similarly, *Where*

Late the Sweet Bird Sang is set in an apocalypse that drives characters to fight human extinction with cloning.

The adjacent subgroup of *Lathe of Heaven* and *The Sheep Look Up* displays a content-style correlation as well. *The Sheep Look Up* is set in a destroyed earth; however, unlike in the previously named novels the focus rests on human agency of this destruction and offers a true anthropogenic account of climate change. In comparison, *Lathe of Heaven* uses a main character who has the power to change the world's condition with his dreams. The questioning of society's violence towards the planet and each other unites both novels, though, the match is not as obvious or direct as seen in the previous example of *False Dawn* and *Where Late the Sweet Bird Sang.*

The last branch, *The Word for World is Forest* is surprisingly separated from Le Guin's other work. Still the split remains consistent even after (reasonable) changes in the parameters and is therefore significant. Though the novel fits the science fiction genre, it supersedes the obvious space narrative and engages questions of racial equality with its imagined colonization of space, as a metaphor for eco(in)justice.

After analyzing the respective corpora, the divides between clusters appear consistent. Neither of these analyses provided actual close readings of the respective novels, and they did not address the characters, the temporal aspects or even a close account of the settings; yet the simple content summary substantially correlates with the style, a remarkable observation at this point. Moreover, the development of a more nuanced understanding of climate change surges. Science in this decade starts paying critical attention to ecosystems around the world and establishes links between atmospheric carbon dioxide concentrations, melting glaciers, and sea level rise.
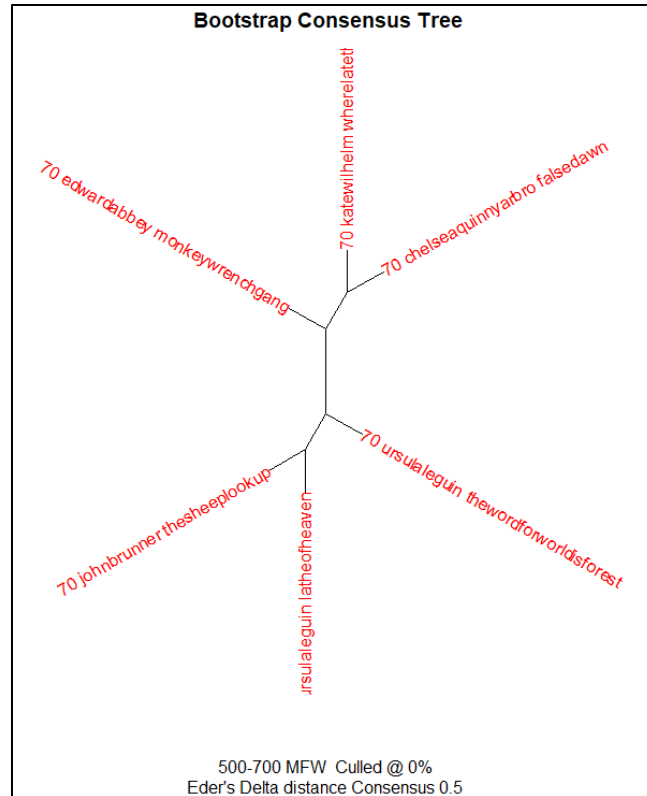
Figure 2: Consensus tree of the 1970s corpus with three iterations over the 700 most frequent words.

For more clarity on whether these previous splits are consistent and whether they form robust subgroups Figure 3 presents a consensus tree that contains both the novels of the 1960s, and the 1970s. Indeed, the algorithm connects Ballard and Carson while leaving Dick as a separate branch. Moreover, Herbert is located between the two, though closer to Dick. With the added information, Harrison and Le Guin, though they are still part of the same branch, are distanced from each other. This branch continues to confirm the content-style theory as all novels discuss and critique society's actions towards individuals or the environment. In Le Guin's *The Left Hand of Darkness* and *The Word for World is Forest* it is a critique of settler colonialism, race, and eco justice while in *Make Room! Make Room!* the critical eye is focused on the environmentally destructive practices of a society.

23

The Abbey, Wilhelm and Yarbro branch reveals the struggle to survive in a hostile system. In Wilhelm and Yarbro this is a hostile ecology, while in Abbey's novel this is the pollution through capitalism.

Looking to Dick and Herbert, the novels' positions appear not too different from the previous cluster, however the fantastical element in their survival struggles appears to map onto a stylistic connection. This connection as standalone identifier is not strong, and therefore questionable. Similarly, Carson and Ballard, despite their common theme of industrial pollution, are substantially different in terms of their narrative frames. The stylistic connection, therefore, will not hold in ensuing discussions.
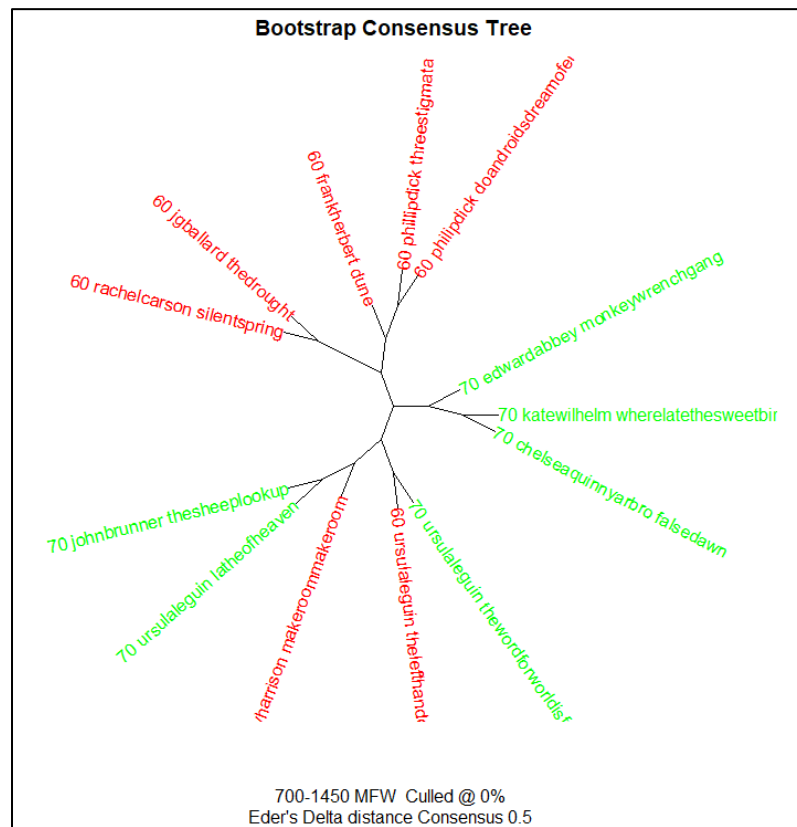

Figure 3: Consensus tree of the 1960s and 1970s corpus.

Following these implications Figure 4 shows a plot of the Principal Components Analysis (PCA) for the 1960s and 1970s. A PCA calculates the data points after summarizing the most important features for analysis into two principal components. PC1 combines the most critical

factors, in this case 17.3%, while PC2 combines other features. Looking at the plot below, it becomes clear that *The Drought* and *Silent Spring* digress strongly from each other and the remaining novels. The reason for the algorithm previously relating them in the consensus tree is likely a similar distance from 0.

Beyond that, the plot confirms the subgroups visualized in the consensus tree of Figure 3. Le Guin, Brunner, Dick and Herbert in fact are alike in terms of the first variable but differ in the second group of features applied. Le Guin's *Lathe of Heaven* is closer to Brunner's *The Sheep Look Up* than to her own remaining works. The style of this novel and the focus on earth appears to substantially differ from the abstracted themes of the other two novels. With further data it will become clear that these two novels are more at the core of this corpus of climate fiction and therefore more representative of its features. Dick and Herbert form a very close cluster, as do Wilhelm and Yarbro. From the reading of these graphs, the deduction must be that subgroups exists for Dick and Herbert; Wilhelm and Yarbro; Le Guin; and Le Guin and Brunner. Which means by the end of the 1970s there exist three strong stylistic directions within climate fiction writing. It is not wise to treat Ballard and Carson as a uniform subgroup, nor to join Abbey and Harrison into fixed clusters. Instead more data would be required to see if there are novels that relate to these cases.
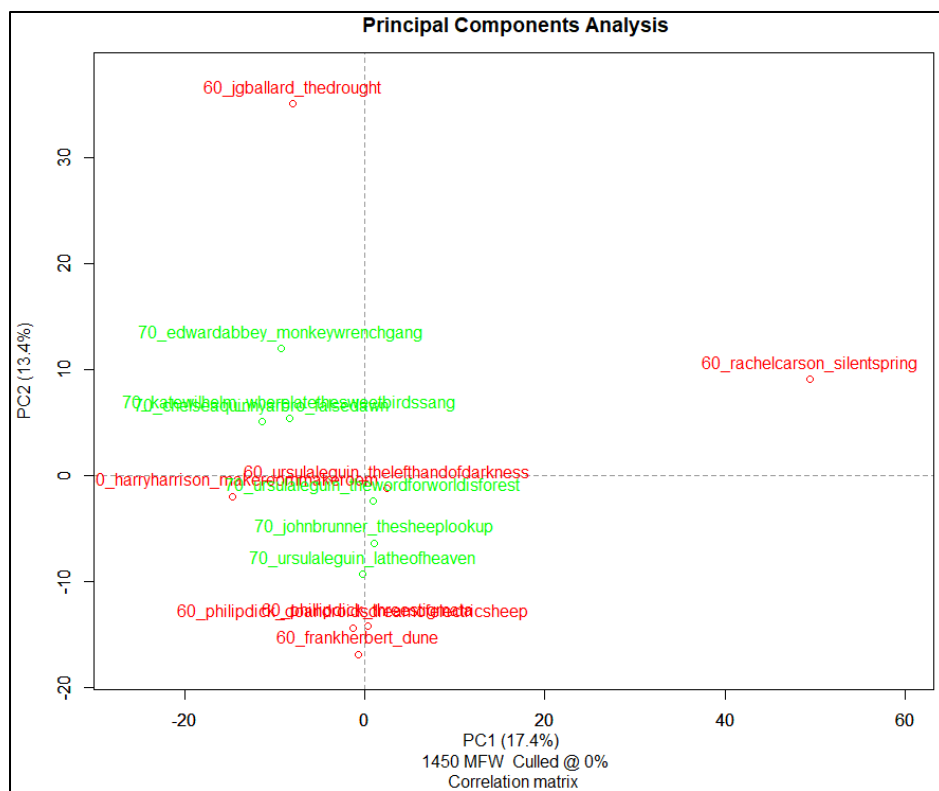
Figure 4: Correlated PCA of the 1960s and 1970s corpus.

Moving to the 1980s corpus, Figure 5 maps these novels in a consensus tree, first without

other noise and later in comparison to the other data. The corpus is split into three subgroups and

two outliers, the first outlier being a novel by Ballard who, as is known from the previous data,

has a special style. In terms of content, *The Crystal World* presents an environmental apocalypse,

the setting in this case is West Africa and the characters see the literal crystallization of the

landscape in form of actual diamonds and crystals. The second outlier is Jean Auel's *The Clan of*

the Cave Bear, a prehistoric fiction, which addresses questions of solastalgia and interspecies

relations as it tells the story of a human girl meeting and joining a group of Neanderthals after an

earthquake left them homeless.

While these appear to be the singular outliers, the largest group features only Octavia

Butler's trilogy *Lilith's Brood,* a science fiction series that like *The Clan of the Cave Bear*

contains interspecies relations, essentially a cluster of outliers by one author. The story is set in

both space and on earth and relates the journey of a woman who survived human extinction on

26

earth and was kept alive by aliens. Eventually, it becomes a question of interspecies procreation and a return to earth.

The second subgroup combines two dissimilar authors Alan Dean Foster with *Slipt* and performer Ben Elton's *Stark*. Yet, the stories harmonize, as both feature villainous corporations guilty of environmental destruction and conspiracy against parts of society with the story's main character fighting to stop these polluters. Though their content is related, the correlation of style is surprising, and with a look at the larger corpora the group is not consistent. The same case applies to Kurt Vonnegut's *Galapagos* and George Turner's *Drowning Towers*, which immediately contradict the pattern with divergent content. *Galapagos* is the story of an unusual evolution of a new related species after all of humanity has died except the small population of an island. In contrast, *Drowning Towers* is the story of survival in the light of government corruption and climate change events that render life for the poor increasingly hard or impossible. The dissimilar nature of both author's style and content of this subgroup raises questions, either the content style relation holds no more or the cluster will not remain consistent. Overall, it can be noted that novels of this decade feature an introduction of species, extinction and interspecies relations, previously missing from the corpus. The topics discussed in the texts mirror the growing attention to greenhouse gas emissions and the international task forces created to understand the implications of these pollutants.

Bootstrap Consensus Tree

80 octaviabutler adulthoc

80 octaviabutler dawn

80 octaviabutler imago

80 jgballard thecrystalworld

80 jeanauel theclanofthecavebear

80 alandeanfoster slipt

80 benelton stark

80 georgeturner drownin

80 kurtvonnegut galapagos

700-1000 MFW Culled @ 0%
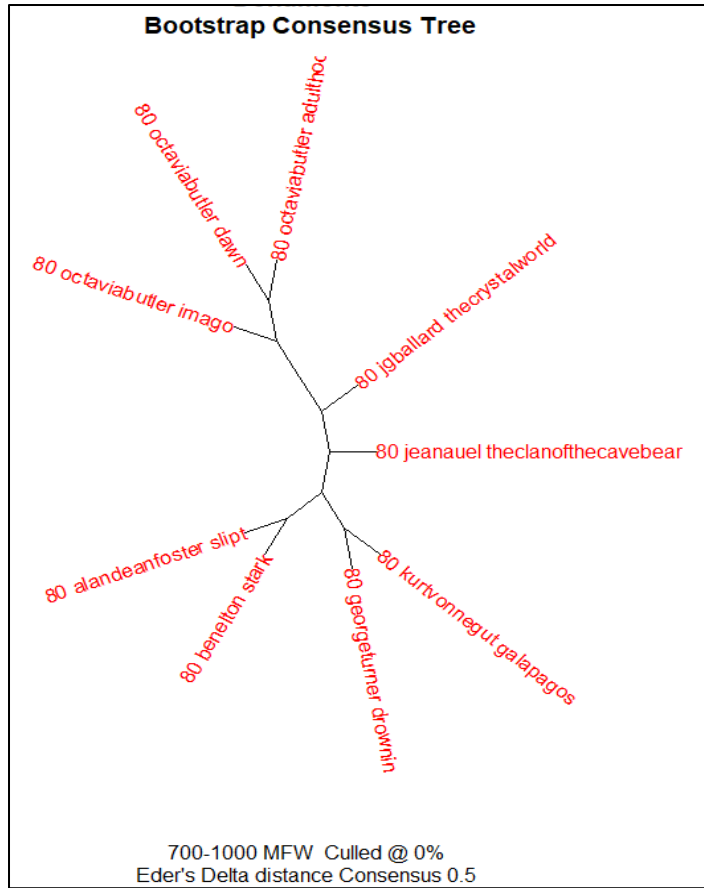Eder's Delta distance Consensus 0.5

Figure 5: Consensus tree of the 1980s corpus.

The observations previously mentioned did not all hold up in the larger corpus; as expected, the consensus tree reclassified the groups. With a growing number of novels to choose from, the groups diverge from the previous illustrations in instances where the position was not definite. Unfortunately, this is the case for a substantial number of examples, seen in novels that switch position depending on the parameters set. Still, constants present themselves; works of the same author keep their positions, which explains the clusters of Butler, Ballard, Dick, and Le Guin.

Beyond these expected clusters Wilhelm and Yarbro, Foster and Harrison, Carson, and Herbert remained in their positions. The changes that occurred with the introduction of the 1980s corpus pertain mostly to this corpus itself. Foster and Elton have been separated into different branches. Similarly, Vonnegut and Turner are separated, though still on the same branch. Vonnegut is now more closely positioned to Le Guin, who is writing about different worlds and

28

imagined species, which appears fitting, yet if this content were a determining factor Butler

should be found in the same group. To resolve this questionable clustering PCA gives

information about the confidence of this model.
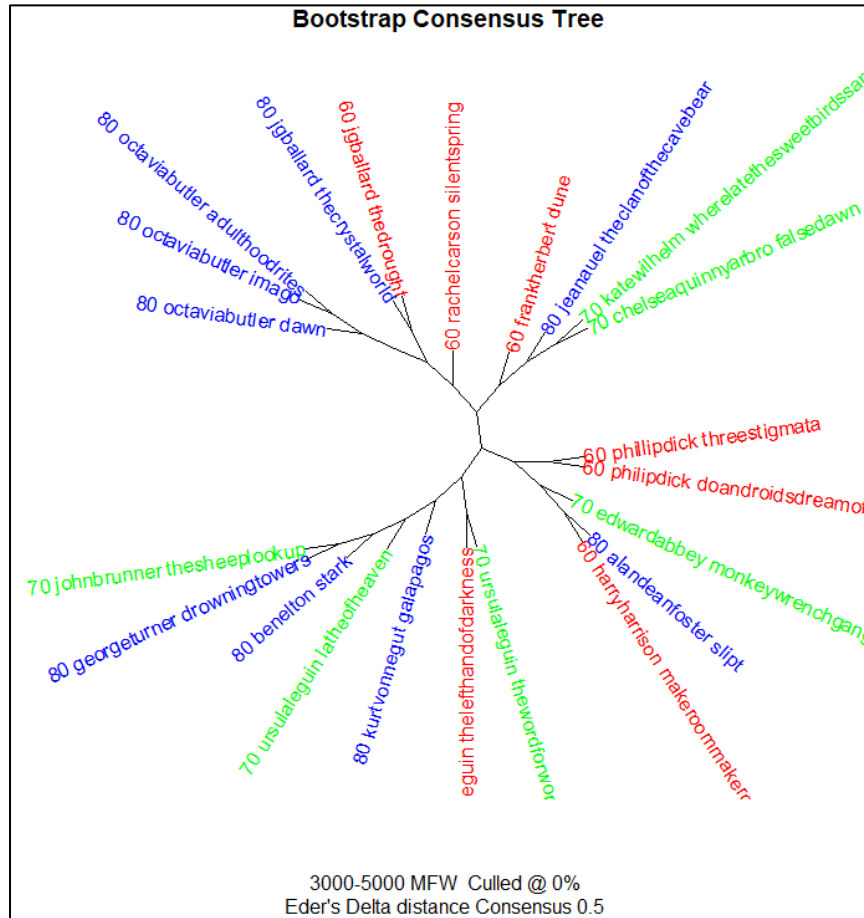
Figure 6: Consensus tree of the 1960s, 1970s, and 1980s.

With the PCA the image is reaffirmed, Vonnegut is in a cluster with Le Guin while

Butler is not only separate from this cluster, but seperated from the total corpus as stylistically

distinct. Similiarly, Carson and Ballard are differentiated while the majority of the novels are

clustered around the center. This trend is logical, from statistics it is known that with an increase

in data the number of normal cases increases, while allowing for a few outliers. Statistically

speaking, these exceptions are the interesting cases as they predict where models are less

applicable, and for the purposes of this thesis the protocol is not as straightforward. However, it

can be deduced that the majority of the novels share a baseline of variables that make them alike,

at the very least through an imagined discussion of climate fiction. In conclusion, the majority of climate change novels by the 80s share a repertoire of words and style and there exists a model case of the climate change novel. Special cases are Rachel Carson, Ballard's novels, Butler's *Lilith's Brute* Trilogy, *The Clan of the Cave Bear, Dune* and *The Monkey Wrench Gang*. At this point, the content cannot be named as the main distinguishing factor, it must truly then be a unique approach to their writing that situates these novels apart from the bulk of novels. Going forth, the corpus will be increased to include the following decade.

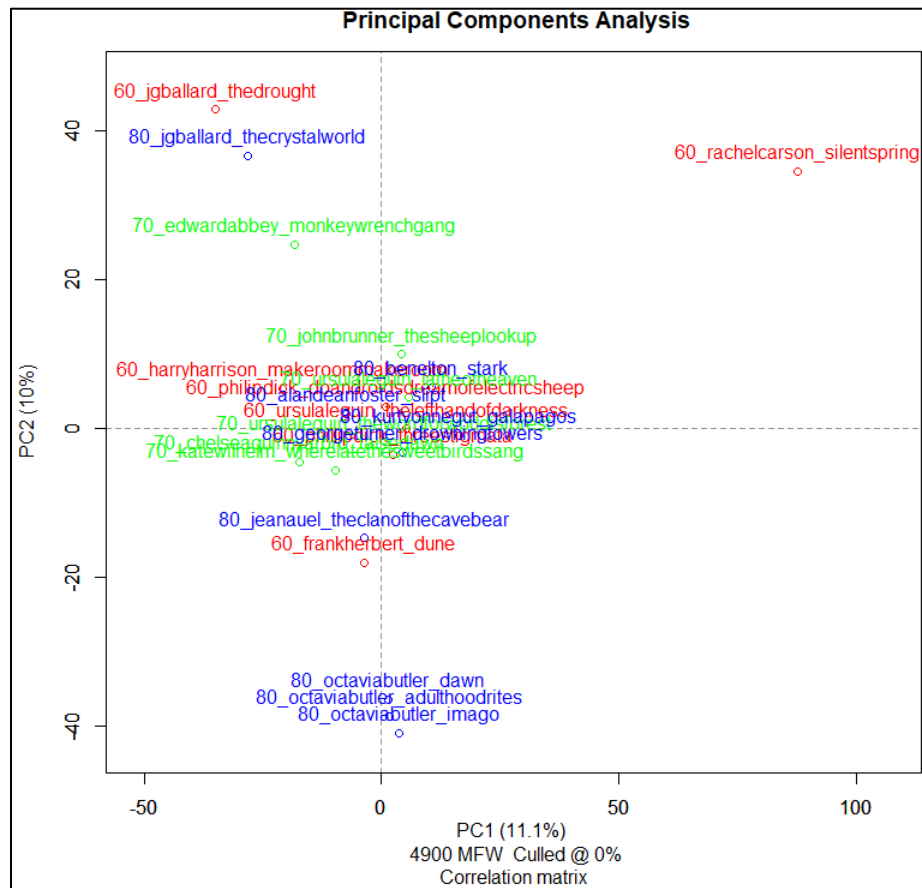Figure 7: Correlated PCA of the 1960s, 1970s, and 1980s corpus.

Three authors – Octavia Butler, Kim Stanley Robinson, and Ben Elton – dominate the 1990s corpus, and logically form subgroups in the consensus tree. The two other authors represented in the corpus are Starhawk and Bruce Sterling forming their own branch. Overall, the climate change references become more explicit in this decade. *Heavy Weather* manifests this

trend with a story of an ecosystem changed through greenhouse gases and the ensuing catastrophes. The second node of this cluster, *The Fifth Sacred Thing* presents the reader with an ecotopia that emerges in the U.S. after the apocalypse has hit. Similarly, Butler's book series is set in the U.S. after the climate crisis has become a life-threatening phenomenon. All three novels see the formation of a cult that changes the societal structure's and norms, therein reinforcing the hypothesis of a style and content correlation. Yet, Kim Stanley Robinson's cluster does not neatly fit this theory. While the novels in Robinson's *Mars Trilogy,* which details humanity's expansion, settling, and terraforming of other planets after earth has become uninhabitable, is a logical and expected group. His novel *Antarctica* is an eco-thriller with a character investigating life on Antarctica and stands out compared to the trilogy.

Ben Elton's novel *Gridlock* illustrates the effects of a government conspiring against public transportation in favor of the automobile industry. *This Other Eden,* centers around a company that sells domes to isolate people from environmental disaster, profiting off the climate crisis. Both, in other words, target institutions within the climate crisis. To relate these observations to previous decades, the case could be made that the groups are coherent in both style and content, in this case offering increasingly detailed depictions of the climate crisis, however this is not expected to hold up in the larger corpus.

Bootstrap Consensus Tree

90 kimstanleyrobinson redmars
90 kimstanleyrobinson gr
90 kimstanleyrobinson antarctica
90 octaviabutler parableofthetalens
90 octaviabutler parableofthesower
90 benelton gridlock
90 benelton thisothereden
90 brucesterling heavywe
90 starhwak theriffithsacredthing

3000-5000 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5
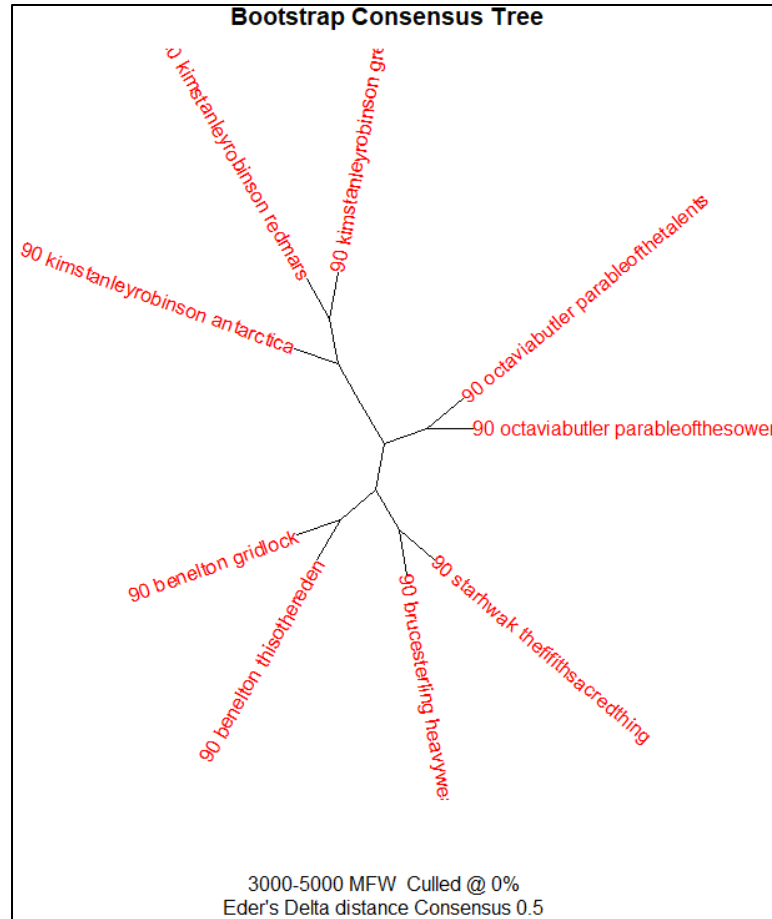
Figure 8: Consensus Tree of the 1990s corpus.

The consensus tree from the 1960s to the 1990s presents a nuanced graph that gives no room for the identification of patterns. The depiction brings no enlightening information that would add value to the debate and hence has been added to the Appendix but will not be discussed further at this time. To potentially view a more distinguished pattern, Figure 9 illustrates the PCA of this data. For this visualization Carson and Ballard have been removed, as their exceptional nature has been recorded already. The deletion allows an emphasis on the remaining corpus and consequently a more in-depth analysis. It should be noted that the PCA values are significantly lower compared to the previous plots, thus attesting to a lower consensus of the variables.

In this visualization author groups, meaning works by the same author are recognized as stylistically almost identical, this includes Ben Elton, Philip Dick, Octavia Butler's *Lillith's*

32

Brood, and Kim Stanley Robinson's *Mars Trilogy. Lathe of Heaven* continues to be classified as distinct from Le Guin's other works, as is the case with Robinson's *Antarctica*, which means the authors wrote differently, a skill that is not shared by most writers. Beyond that the PCA reiterates previous observations, such as aligning Herbert with Dick, and Le Guin with Harrison; the same applies to the 70s and 80s corpora. At closer investigation the patterns do hold up, which classifies them as meaningful.
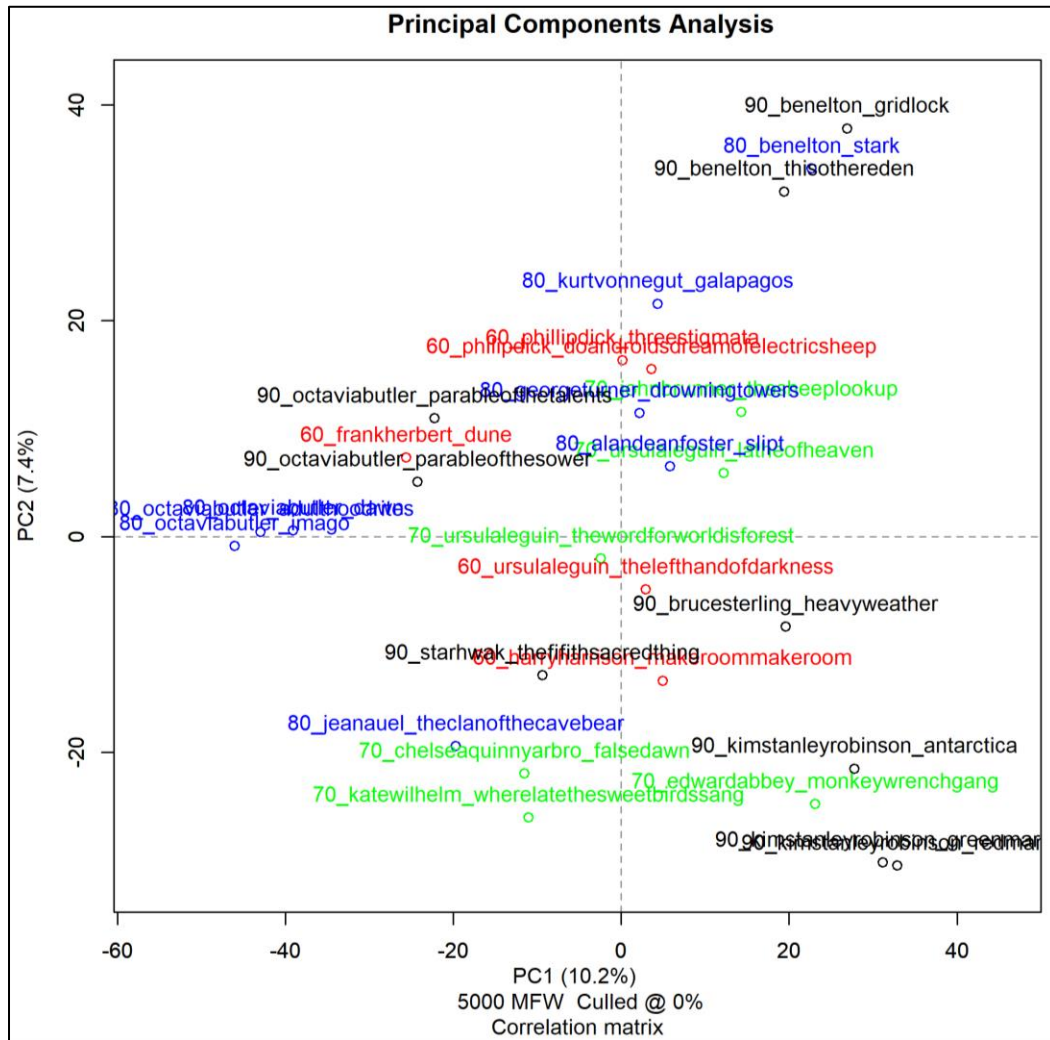

Figure 9: PCA from the 1960s to the 1990s with the removal of Carson and Ballard.

Moving further through the decades, Figure 10 depicts the novels of the 2000s in order to become familiar with the texts included in this corpus. The first cluster is comprised of Stephen Baxter's series *Flood* and *Ark. Flood* tells the story of sea level rise and the ensuing floods in

33

London, focusing on the effects on individuals in light of this escalating catastrophe. The sequel *Ark* shares the desperate attempts to find modes of survival and continues the narrative of *Flood.* Paul McAuley's *The Quiet War* differs in genre and narrative frame with a setting in the far future after climate change has forced humanity to leave earth and form new habitats on other planets, thus giving way to a power struggle between governmental forces. The following separate note of Ray Hammond's *Extinction* follows the unfolding of catastrophe, set in the near future, the climactic events challenge humanity and reveal the unpreparedness of the systems that are meant to regulate life as is, leaving individuals to find their own solution in the apocalypse. The following branch singles out Cormac McCarthy's *The Road,* a post-apocalyptic narrative about a father and a son traveling through a deserted America. The bottom of the tree ends in two clusters, the first one consisting of Margaret Atwood's *MaddAddam* Trilogy here presented by the first two books *Oryx and Crake* and *The Year of the Flood.* The story relates the before and after of an apocalypse first from the view point of an isolated character who actively participated in the coming of the apocalypse, and in the sequel through the eyes of activists, expecting the apocalypse but not actually creating it. The final group shows once again the strong influence of science fiction with Susan Pfeffer's *Life as We Knew It* a YA novel about a family surviving the apocalypse resulting from a new position of the moon. On a branch attached to Pfeffer is Kim Stanley Robinson's *Fifty Degrees Below*, one of the first novels to scrutinize strategies of mitigation and adaptation and *The Remnant* by Tim LaHaye and Jerry Jenkins, part of a religious science fiction series set in nuclear Armageddon. From this overview, no clear directions become visible. However there is insight to be gained as the novels become more explicitly political which coincides with global efforts to address strategies of mitigation and adaptation best exemplified in President Bush denouncing the US participation in the Kyoto Protocol, claiming it would hurt the American economy. The novels appear to point to diverging power struggles as well as to the individuals who are left to fend for themselves.

Bootstrap Consensus Tree

00 paulmcauley thequietwar
00 stephenbaxter flood
00 stephenbaxter ark
00 rayhammond extinction
00 cormacmccarthy theroad
00 margaretatwood theyearoftheflood
00 margaretatwood oryxandcrake
00 susanpfeffer lifeaswekne
00 kimstanleyrobinson fiftydegreesbelow
00 timlahayejerryjenkins remnant

3000-5000 MFW Culled @ 0%
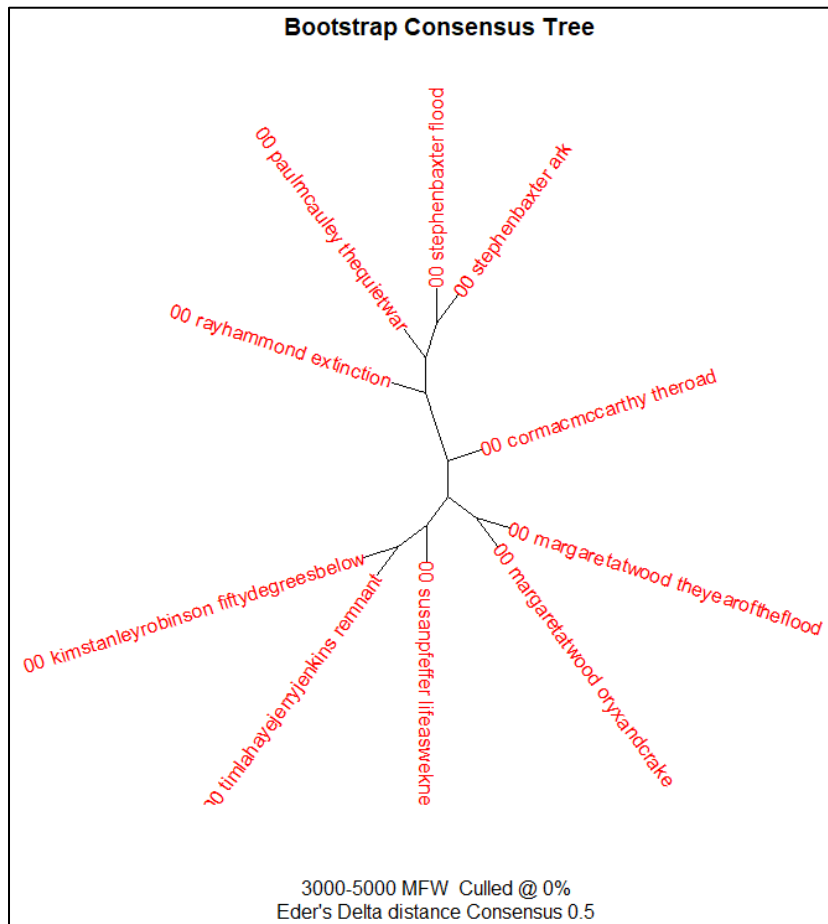Eder's Delta distance Consensus 0.5

Figure 10: Consensus Tree of the 2000s.

Further situating the changes of the decade within the corpus, the PCA in Figure 11,

stripped of Carson and Ballard, highlights the main clusters. The graphic sets McCarthy apart

from the main corpus, and Elton's novels are similarly distinct from majority of novels. The style

of these two authors is accordingly uncommon when compared to other climate change fiction

available at the time. Le Guin and Atwood's works, written with the most overlapping style are

located at the center of the plot. These novels are the prototypical basis against which novels are

compared. The climate change novels of the 2000s appear to either gravitate toward the heart of

the corpus or in hard science fiction clusters with Kim Stanley Robinson and Bruce Sterling.

Unfortunately, the visualization gives no additional clarification at this point. To see the

connections between novels more clearly, a network with the information from stylo is presented

in Figure 12.

35

Principal Components Analysis

PC2 (6.6%) (y-axis)

00_cormacmccarthy_theroad

70_edwardabbey_monkeywrenchgang

70_chelseaquinnyarbro_falsedawn
70_katewilhelm_whererelatethesweetbirdssang
60_harryharrison_makeroommakeroom

60_frankherbert_dune 00_somtowsucharitkul_starship
00_stephenbaxter_flood
00_stephenbaxter_ark
60_frankherbert_dune ...wak_thefifithsacredthing
80_jeanauel_theclanofthecavebear
90_brucesterling... robinson_redmars
00_paulmcauley_thequietwar

70_ursulaleguin_thewordforworldisforest
...leguin_thelefthandofdarkness
00_margaretatwood_theyearoftheflood
stanleyrobinson_fiftydegreesbelow
00_susanpfeffer_... ...andeanforster... oryxandcrake
80_octaviabutler_adulthood_ ...octaviabutler_parableofthesower
80_octaviabutler_dawn
70_ursulaleguin_latheofheaven
60_philipkdick_doandro...
80_georgeturner_drowningtowers
90_octaviabutler_parableofthetalents

80_kurtvonnegut_galapagos

90_benelton_thisothereden
80_benelton_stark
90_benelton_gridlock

PC1 (8.6%)
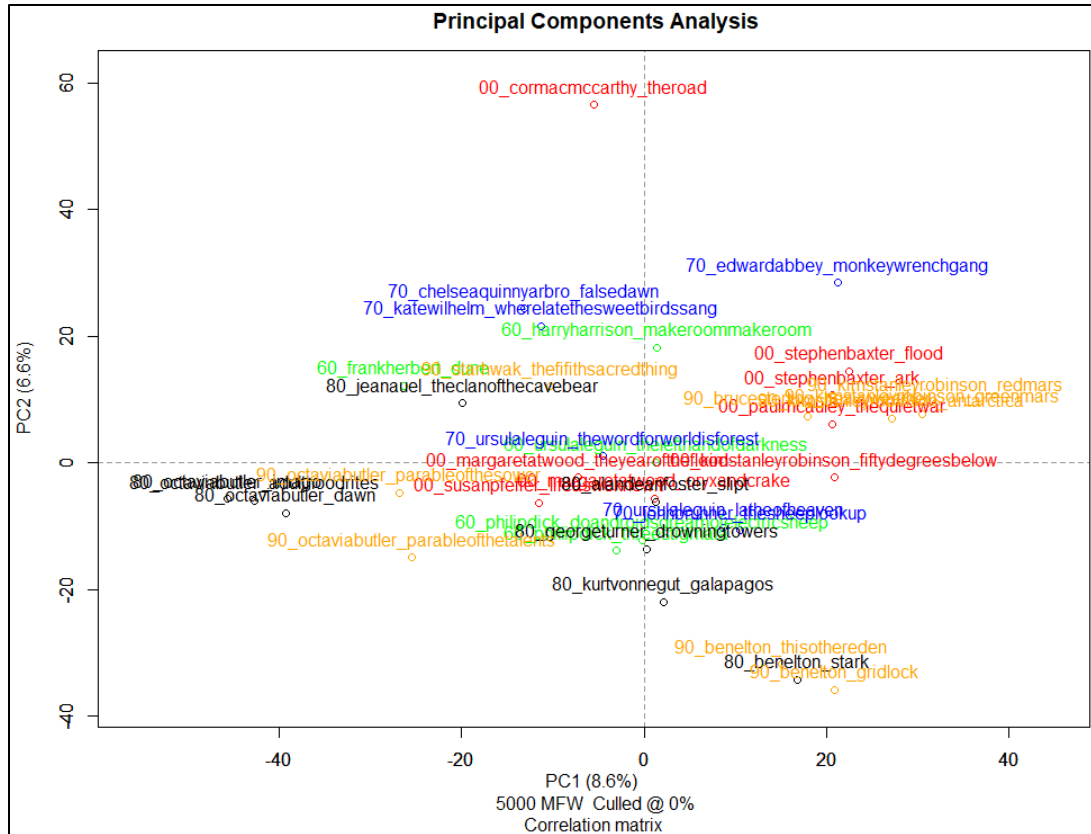5000 MFW Culled @ 0%
Correlation matrix

Figure 11: PCA including all corpora until 2000 without Ballard and Carson.

Before diving into the analysis of the network below, some facts regarding the nature of

the network illustration will be helpful. Unlike the consensus trees, distance is not a meaningful

unit in these networks, the space seen in the graph is purely aesthetic. The most important

information is expressed through connection measured by the variable "degree," a node with a lot

of connections has a high degree. Related to the measure of degree are various types of centrality.

For the purposes of this analysis the relevant variation is *betweenness centrality*, the literal

position of being between two other nodes, which measures the importance of a given node based

on the amount of connections that pass through it. A novel with a high betweenness centrality not

only connects to other stylistically similar texts but connects multiple layers of the network.

Specifically, the node with the highest betweenness centrality by far was Margaret Atwood's

Oryx and Crake. In this visualization is signified by the size of the node. Seeing as the novel was

written later than most of the corpus, the betweenness indicates its intertextuality, meaning *Oryx*

36

and Crake is likely influenced by the other texts rather than having influenced these texts. The reverse may apply to an earlier example, Brunner's *The Sheep Look Up,* which has a high betweenness centrality and hence appears to have both given inspiration to other authors while producing a text that relates to the existing canon.

Beyond the betweenness centrality the modularity, a measure for the clustering within a network, is a staple when analyzing network as it gives information about the subgroups of a network. The Gephi algorithm detects the strength of connections between data points and groups them accordingly, which is represented in different colors in the network below. The clustering shows four categories, as has been seen in previous consensus trees already. These appear to cross genre boundaries exemplified by the purple cluster including Elton, Carson but also Le Guin and Atwood. A new grouping links McCarthy and Ballard, both of which have been marked as separate from the majority of works. The network can therefore be read as a clue into subgenres of climate fiction. To view the total and final picture, a last analysis that includes the latest works in climate change fiction will complete this analysis.
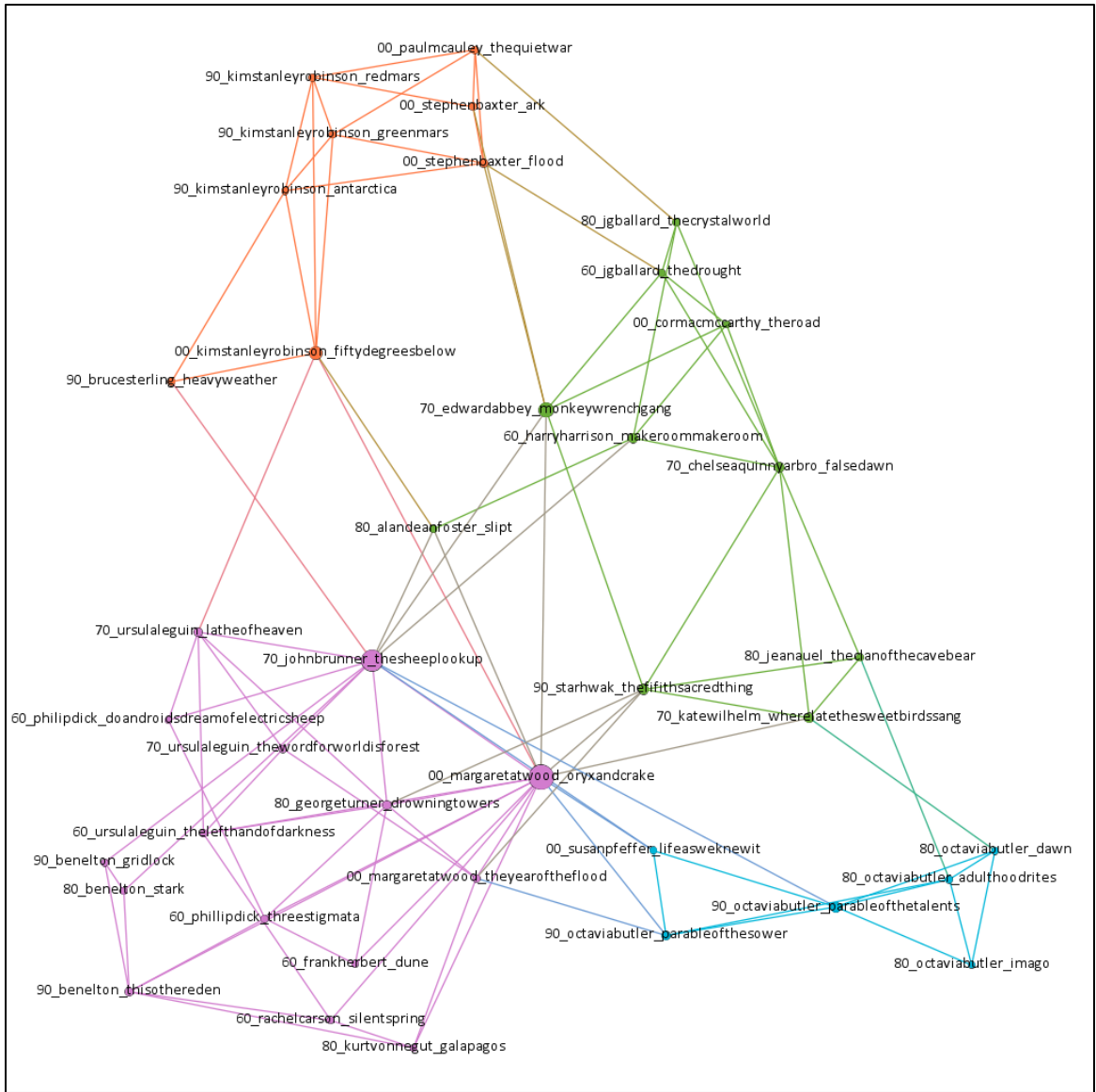
Figure 12: Network for works ranging from the 1960s to the 2000s.

The 2010s are marked by an exponential increase in climate change awareness through political action like the Paris agreement in 2015 where participating nations vowed to keep rising temperatures under 2 degrees Celsius. Aside from this, it is entirely common to have a large part of climate science in the public discourse with major newspapers daily publishing at least one article related to climate change, beyond that, they translate major findings into digestible and accessible information. Meanwhile, the effects of climate change are increasingly felt, most dramatically after natural disasters. Resulting from the access to information and widely

addressed issues of climate change, the decade sees increasing consumer actions criticizing and stepping out of the neoliberal market. Other popular activism organizations that illustrate this trend are *Occupy Wall Street* and of late *Fridays for Future*, initiated by Greta Thunberg. Remembering these developments, the rise in climate change fiction is unsurprising.

Listing all the novels would be too lengthy; a spotlight on selected works will help in a familiarization with the corpus. Standout works include Barbara Kingsolver's *Flight Behavior*, a narrative about a housewife and her encounter with a unique symptom of climate change, displaced monarch butterflies. Beyond Kingsolver, Amitav Ghosh has recently published the novel *Gun Island* that centers on questions of displacement and transition of life in the climate crisis. Contrastively, the eco-thriller *The Ice* by Laline Paull is set in the arctic and tells the story of a friendship among melting glaciers and species extinction. Ben Lerner's *10:04* relates the present condition marked by impending doom, pointing to the loss of imagination and hope through the eyes of a New York writer. The satire *Solar* by Ian McEwan becomes a question of authority, who reserves the right to guide humanity through this crisis? The last novel to be pointed out here is *The Collapse of Western Civilization* by Naomi Oreskes and Erik Conway which merges science and historical fact with the style and format of a novel, it analyses different interest groups within the current crisis and relates the impact of the ideologies pursued. With this snapshot alone, a notable diversification becomes apparent in climate change novels.

The corpus at this time is not dominated by any one style or genre but offers a plethora of approaches to the crisis. Figure 13 shows a split into four groups with one group (here blue) being significantly smaller in comparison to the rest. It includes *Rain Birds*, *The Marrow Thieves*, and *Convergence*. The latter two fit in the categories YA, combining science fiction and minorities, while *Rain Birds*, with its focus on the experiences of two ecologists and their dilemmas in life, appears to be the odd one out. Yet the connection remains robust in the larger corpus, hence their style must indeed be similar.

Betweenness centrality helps to identify Kingsolver's *Flight Behavior* as the most influential node in the network and therefore the most representative novel of this corpus. The three other highly connected nodes are *The Ice, Gun Island,* and *Maddaddam*; all three are part of different subgroups and become important nodes within the respective clusters. Seeing as Ghosh's novel has been published in 2019 it is less likely to have influenced the remaining novels in the green cluster, instead it appears to join what Ghosh has termed "serious fiction" or literature. Atwood's final book of the *Maddaddam* series shows the author as front runner of the orange cluster, which consists of speculative fiction and science fiction with the element of a fantastical future. Contrastively, the purple cluster around the *Ice* contains novels that address apocalyptic scenarios with tropes like glacier melts, ice ages, floods, and heatwaves. Overall, it can be said that the decade is again split into four categories that seem to among other things be consistent in style and frame.
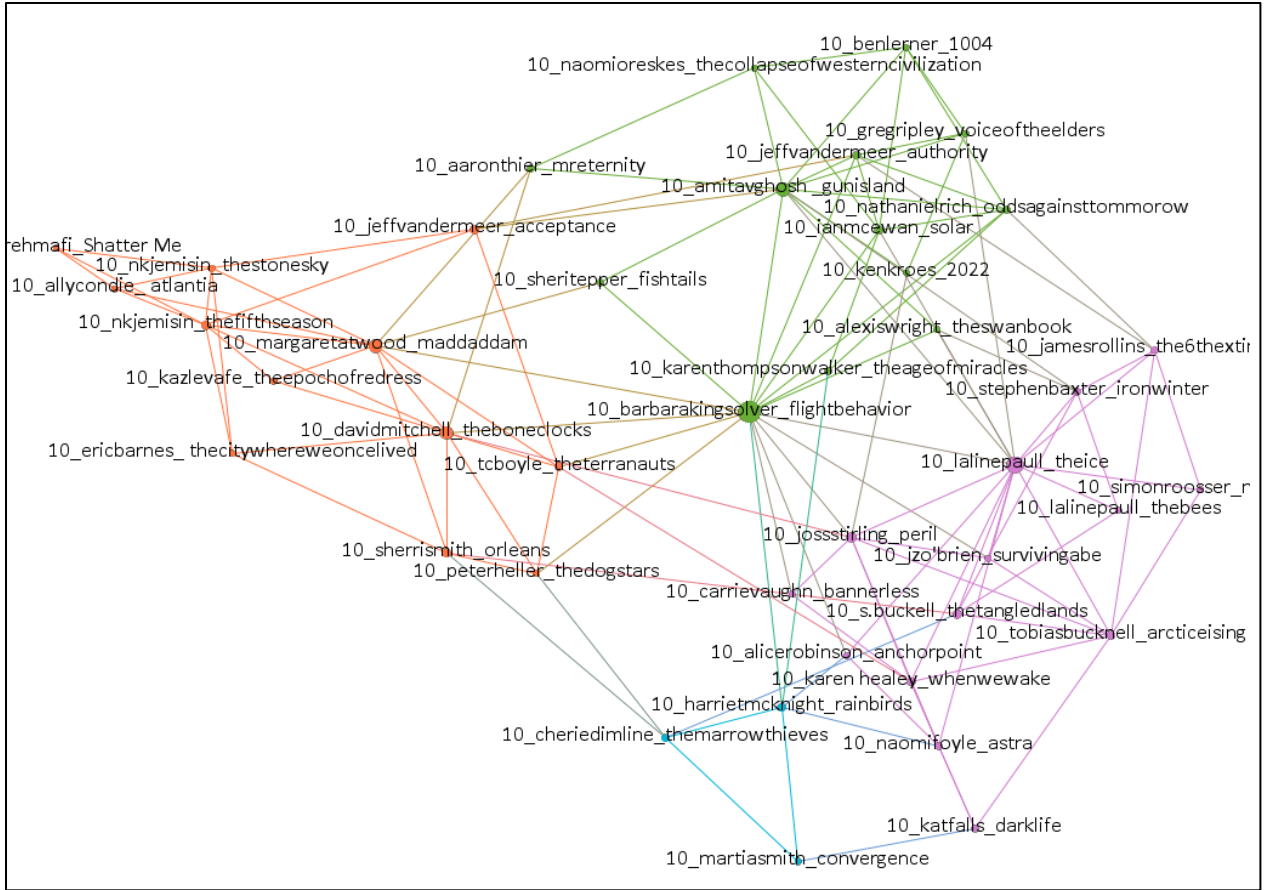
Figure 13: Network of the 2010 corpus.

Lastly, a look at the entire corpus is in order, to verify some clusters as robust and view the works that stand out over all decades. Figure 14 shows 8 modularity classes with weighted edges where the thicker the edge the stronger the connection. One of the largest clusters in light green contains the works of Atwood, Carson, and Vonnegut. The strong connection between Carson and Oreskes comes as no surprise with both texts adhering to scientific facts in their writing. Secondly, the pink cluster separates Octavia Butler once again from the rest of the corpus, reinforcing the notion of this robust finding. Though the distance is not representative of the stylistic distance measures Butler's works are little connected to other novels, differentiating the cluster from the remaining corpus. Moving clockwise, there is a small brown cluster connecting Elton's texts, with its insulation confirming the previously viewed difference seen in the PCAs. The comics writing style thus must be different from the traditional style of other

authors. The light blue cluster almost completely replicates the branches seen in the consensus trees of the 80s and 90s. Generally, a certain affinity to science fiction and space as the last resort can be found in this cluster with the *Mars* trilogy and Le Guin's works. Moreover, the cluster is united by the narrative element of a post-apocalyptic earth and scrutiny of society's behavior in the aftermath. In comparison, the orange cluster with Ballard and Baxter involves imagery of natural disaster more explicitly and depicts the apocalypse as it unfolds. In the networks' center, Ghosh, Kingsolver, Paull and Dick form a cluster that maintains connections to most of the surrounding groups. For example Dick is linked to LeGuin, and towards the lower half of the cluster Kingsolver, Paull, and Gosh connect to Atwood and the green cluster, as well as to the purple cluster. The purple cluster unites stories of the present condition in the climate crisis, everyday characters and their battle for surviving the apocalypse or a natural disaster.

Lastly, the forest green cluster sets apart a small number of recent YA science fiction novels that feature a strong female lead, who warrior-like fights against the system using her unique powers. Overall, the clusters in this network reflect trends that have been observed in consensus trees and PCAs, only a few novels have shifted or have been reclassified. It cannot be denied that meaningful patterns have emerged from this corpus that hopefully inspire further inquiry and investigation.
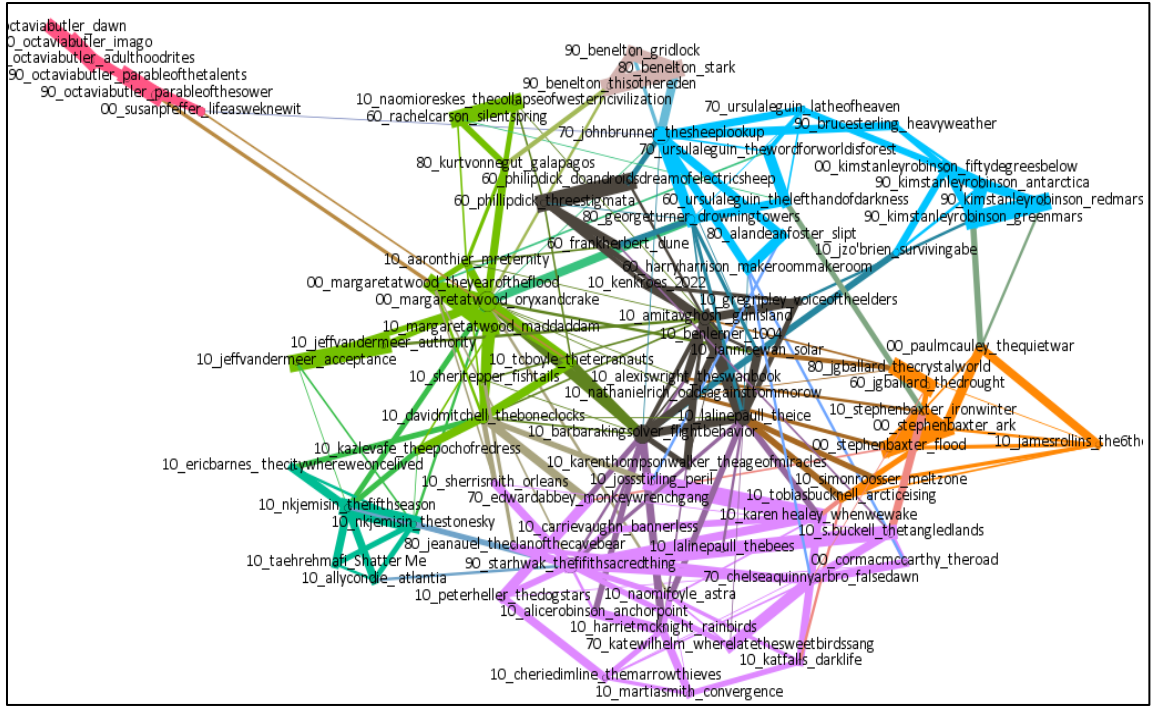
Figure 14: Network of the complete corpus with weighted edges.

CHAPTER 4

CONCLUSIONS

This thesis has set out to computationally investigate climate change fiction with stylometric analysis. With the corpus created for analysis, though not comprehensive, a wide range of these texts have been found and analyzed. Overall, there has been little to no indication that style varies based on the time or decade a text was created in. However, there are occurrences that correlate with developments of their times, seen in the example of *Gun Island* or *Rain Birds* for the 2010s, where an account of everyday life in the crisis becomes a new type of climate change fiction. Additionally, the study has shown that style can be both unique to one author and related to others through narrative set up and content. It appears content has a tangible influence on the style and the tradition that an author enters when writing their novel.

While within most decades four subgroups emerged, the complete works divide into eight subgroups, two of which are author specific, leaving 6 clusters that are grouped based on their styles. Beyond these trends, individual insights have become apparent as well. Among these is the realization that though Rachel Carson has pioneered in many ways the attention and activism in climate change, her writings have had less impact on ensuing novels than expected. Ballard's novels are interesting outliers that merit further inquiry and comparison, potentially with Stephen Baxter's works as there exists as strong connection between the two. Moreover, Octavia Butler provides a number of outstanding works that merit inquiry and comparison to understand how they are different to critically assess lessons to be learned from her approaches. Her works pose an interesting contrast to Ursula Le Guin's novels, seeing as they are similar in terms of genre and narrative frame a close reading of the two is likely to be fruitful in answering questions like, are the differences a result of an increase in knowledge, are they based on choice of narrator, or are they based on racial differences of the main characters?

Another interesting pattern this analysis has pointed to is the opposition of serious fiction, popular fiction, and science fiction. This delineation as Ghosh has termed it, does not hold up in the data as such, however, there are clusters such as the dark green or the Elton cluster that appear less connected and central in their overall contributions to this canon. Therefore, an opposition of these clusters may render some indicators on characteristics unique to each cluster.

More than pointing to patterns and raising new questions, the analysis has given some answers about particularly representative works that have emerged from this corpus. Atwood's *MaddAddam* Series, Barbara Kingsolver's *Flight Behavior*, John Brunner's *The Sheep Look Up*, Starhawk's *The Fifth Sacred Thing*, Laline Paull's *The Ice,* Amitav Ghosh's *Gun Island,* David Mitchell's *The Bone Clocks*, Susan Pfeffer's *Life as We knew It ,* Octavia Butler's *Parable of the Sower* and *Parable of the Talents,* George Turner's *Drowning Towers,* Naomi Foyle's *Astra,* S Buckell's *The Tangled Lands,* Stephen Baxter's *Flood,* and Nathaniel Rich's *Odds Against Tomorrow* are the ten most central novels of this corpus. They provide a list of ten essential climate change novels for any person teaching or interested in climate change fiction[3].

With these things in mind, hopefully there are a few contributions to future scholarship, which could focus on a closer analysis of key novels and in-depth comparisons between them. An expansion to include works not available in this analysis would potentially provide more insight where the algorithm has been uncertain. I hope to give inspiration to other digital humanists to take on the hyperobject climate change so that future projects will prosper.

---

[3] The ranking is based on their betweenness centrality.

CORPUS DATA

Book Title	Author	Year
Silent Spring	Rachel Carson	1961
The Drowned World	J. G. Ballard	1962
The Drought	J. G. Ballard	1964
Dune	Frank Herbert	1965
The Three Stigmata of Palmer Eldritch	Philip K. Dick	1965
Make Room! Make Room!	Harry Harrison	1966
Do Androids Dream of Electric Sheep?	Philip K. Dick	1968
The Left Hand of Darkness	Ursula K. Le Guin	1969
Dune	Frank Herbert	1969
Lathe of Heaven	Ursula K. Le Guin	1971
The Word for World is Forest	Ursula K. Le Guin	1972
The Sheep Look Up	John Brunner	1972
The Monkey Wrench Gang	Edward Abbey	1975
Where Late the Sweet Birds Sang	Kate Wilhelm	1976
False Dawn	Chelsea Quinn Yarbro	1978
The Clan of the Cave Bear	Jean M. Auel	1980
Galapagos	Kurt Vonnegut	1985
Lilith's Brood	Octavia E. Butler	1987
Slipt	Alan Dean Foster	1987
Drowning Towers	George Turner	1987
The Crystal World	J. G. Ballard	1988
Stark	Ben Elton	1989
Gridlock	Ben Elton	1991
Red Mars	Kim Stanley Robinson	1992
Snow Crash	Neal Stephenson	1992
Green Mars	Kim Stanley Robinson	1993
Parable of the Sower	Octavia E. Butler	1993
The Fifth Sacred Thing	Starhawk	1993
This Other Eden	Ben Elton	1993
Heavy Weather	Bruce Sterling	1994
Antarctica	Kim Stanley Robinson	1997
Parable of the Talents	Octavia E. Butler	1998
The Remnant	Tim LaHaye, Jerry B. Jenkins	2002
Oryx and Crake	Margaret Atwood	2003

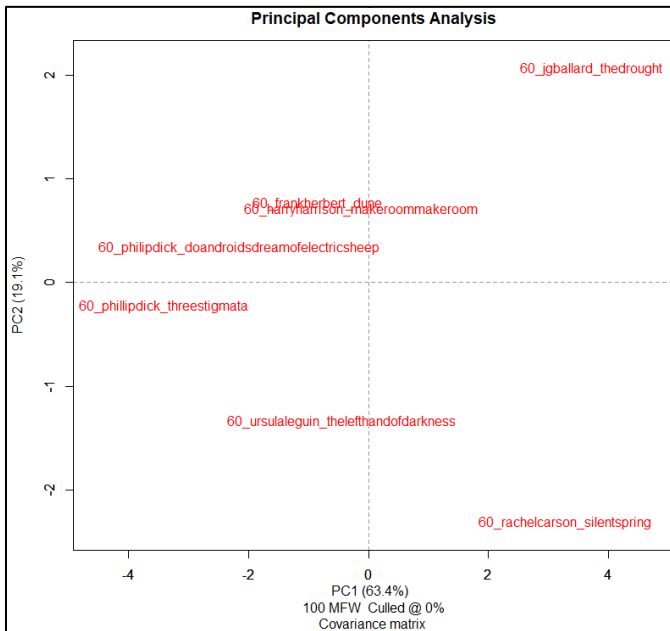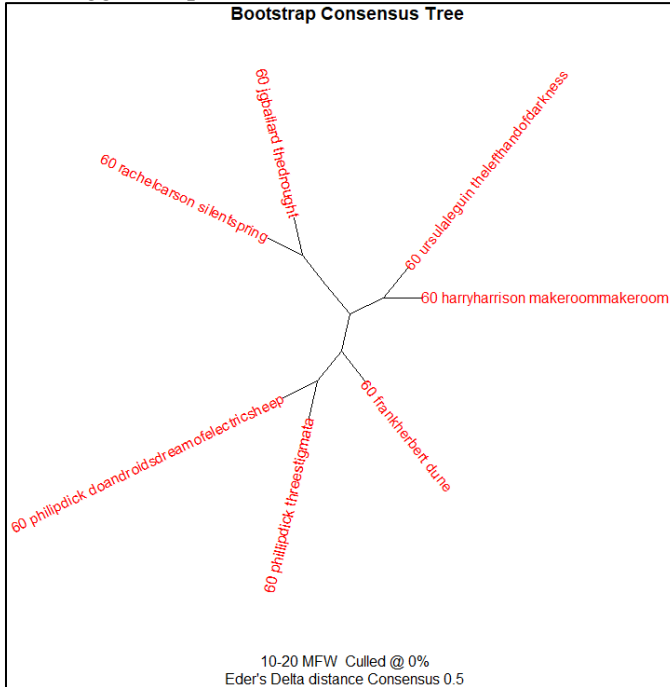|		
Fifty Degrees Below	Kim Stanley Robinson	2005
Extinction	Ray Hammond	2005
The Road	Cormac McCarthy	2006
Life as We Knew It	Susan Beth Pfeffer	2006
Quiet War	Paul McAuley	2008
Flood	Stephen Baxter	2008
Ark	Stephen Baxter	2009
The Year of the Flood	Margaret Atwood	2009
Mr. Eternity	Aaron Thier	2016
The Swan Book	Alexis Wright	2013
Anchor Point	Alice Robinson	2015
Atlantia	Ally Condie	2014
Gun Island	Amitav Ghosh	2019
Flight Behavior	Barbara Kingsolver	2012
10:04	Ben Lerner	2014
Bannerless	Carrie Vaughn	2017
The Marrow Thieves	Cherie Dimaline	2017
The Bone Clocks	David Mitchell	2014
The City Where We Once Lived	Eric Barnes	2018
Voice of the Elders	Greg Ripley	2018
Rain Birds	Harriet McKnight	2017
Solar	Ian McEwan	2010
Nature's Confession	J.L. Morin	2015
Surviving Abe	J.Z. O'Brien	2014
The 6th Extinction	James Rollins	2014
Acceptance	Jeff VanderMeer	2014
Authority	Jeff VanderMeer	2014
Peril	Joss Stirling	2017
When We Wake	Karen Healey	2013
The Age of Miracles	Karen Thompson Walker	2012
Dark Life	Kat Falls	2010
2022	Ken Kroes	2015
The Bees	Laline Paull	2014
The Ice	Laline Paull	2017
MaddAddam	Margaret Atwood	2013
Convergence	Marita Smith	2017
The Fifth Season	N. K. Jemisin	2015
The Stone Sky	N.K. Jemisin	2017
Astra	Naomi Foyle	2014

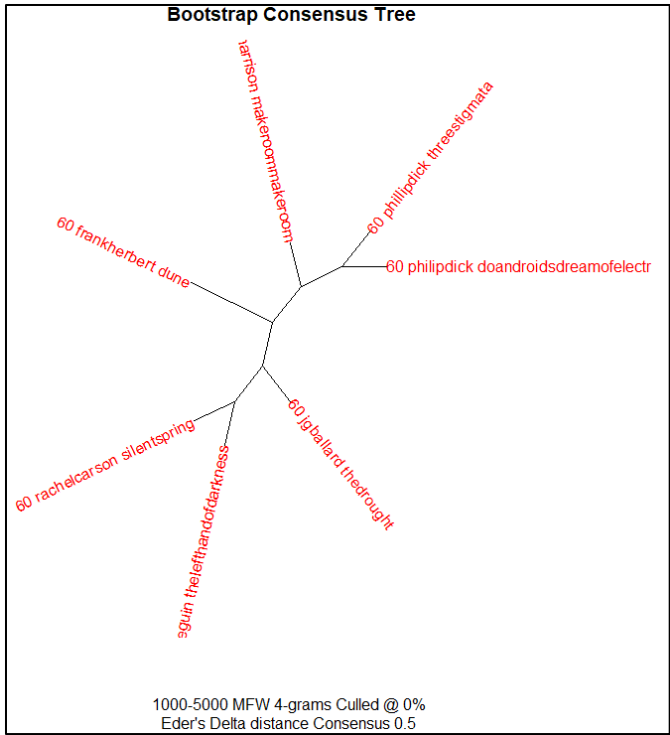|		
The Collapse of Western Civilization: A View from the Future	Naomi Oreskes	2014
Odds Against Tomorrow	Nathaniel Rich	2013
Fish Tails	Sheri S. Tepper	2014
Orleans	Sherri L.Smith	2013
Melt Zone	Simon Rosser	2014
Iron Winter	Stephen Baxter	2012
Haline	Sundeep Ahuja	2014
The Terranauts	T. C. Boyle	2016
Shatter Me	Tahereh Mafi	2011
Arctic Rising	Tobias S. Buckell	2012
The New Atlantis	Ursula K. Le Guin	2013

ADDITIONAL DATA VISUALIZATIONS

Further materials on the 1960s:

POS tagged corpus:

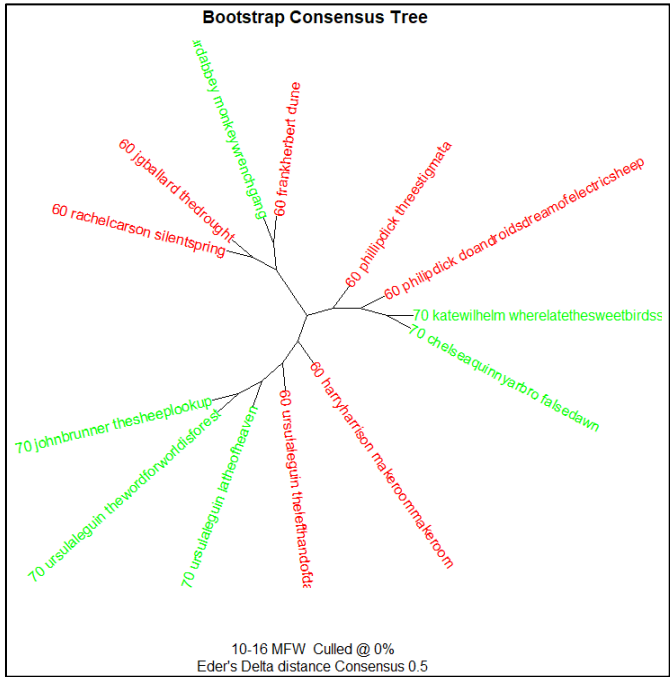Bootstrap Consensus Tree

10-20 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5

Principal Components Analysis

PC1 (63.4%)
100 MFW Culled @ 0%
Covariance matrix

Corpus with 4-grams

Bootstrap Consensus Tree

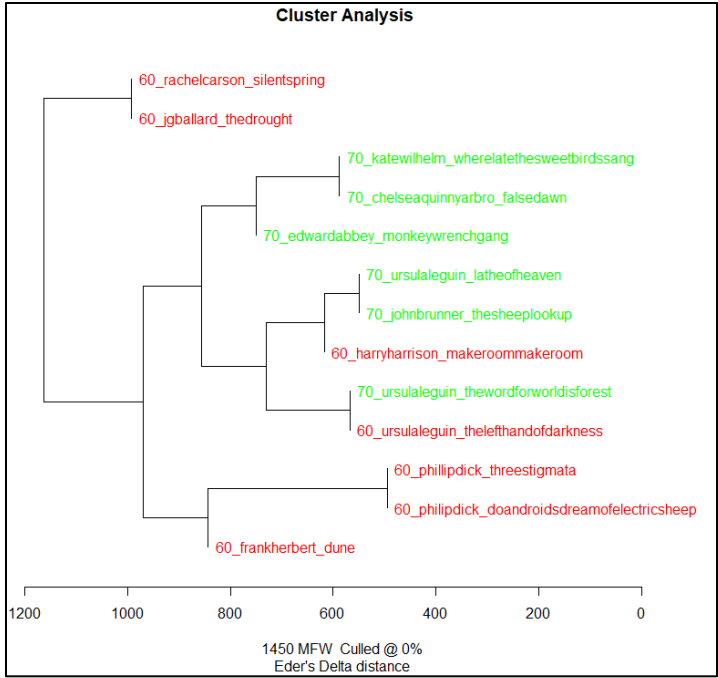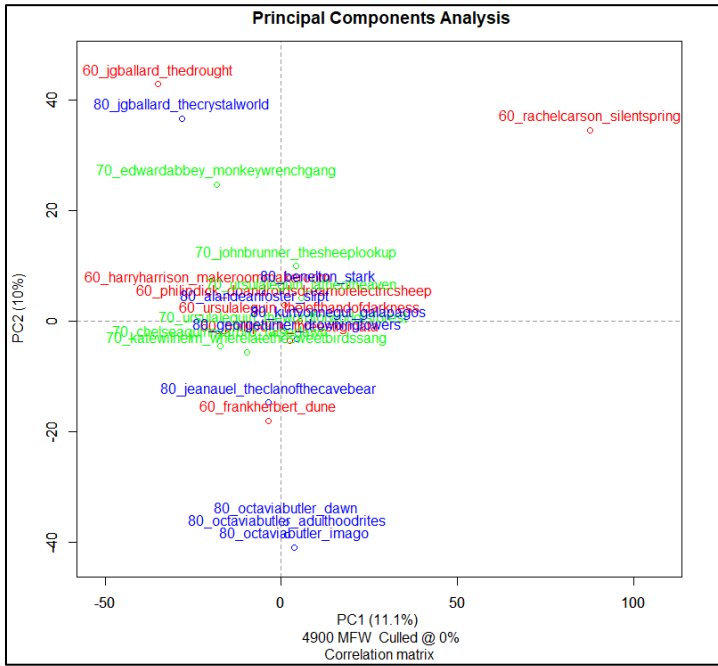1000-5000 MFW 4-grams Culled @ 0%
Eder's Delta distance Consensus 0.5

Further materials on the 1970s:

Comparison between on the POS tagged corpus

Bootstrap Consensus Tree

10-16 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5

Cluster Analysis

60_rachelcarson_silentspring

60_jgballard_thedrought

70_katewilhelm_wherelatethesweetbirdssang

70_chelseaquinnyarbro_falsedawn

70_edwardabbey_monkeywrenchgang

70_ursulaleguin_latheofheaven

70_johnbrunner_thesheeplookup

60_harryharrison_makeroommakeroom

70_ursulaleguin_thewordforworldisforest

60_ursulaleguin_thelefthandofdarkness

60_phillipdick_threestigmata

60_philipdick_doandroidsdreamofelectricsheep

60_frankherbert_dune

1200   1000   800   600   400   200   0

1450 MFW Culled @ 0%
Eder's Delta distance

Further materials on the 1980s:

Principal Components Analysis

60_jgballard_thedrought

80_jgballard_thecrystalworld

60_rachelcarson_silentspring

70_edwardabbey_monkeywrenchgang

70_johnbrunner_thesheeplookup

60_harryharrison_makeroom 80_brucesterling_stark

60_philipdick_doandroidsdreamofelectricsheep
80_alandeanfoster_slipt
60_ursulaleguin_thelefthandofdarkness
70_ursulaleguin_lefthandofdarkness
70_chelseaquinnyarbro_falsedawn 80_georgerrmartin_dyingofthelight
70_katewilhelm_wherelatethesweetbirdssang

80_jeanauel_theclanofthecavebear
60_frankherbert_dune

80_octaviabutler_dawn
80_octaviabutler_adulthoodrites
80_octaviabutler_imago

PC2 (10%)

-50   0   50   100

PC1 (11.1%)
4900 MFW Culled @ 0%
Correlation matrix

Further materials on the 1990s:

POS tagged 1990s corpus

Bootstrap Consensus Tree

10-20 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5

Bootstrap Consensus Tree

3000-5000 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5

Bootstrap Consensus Tree

10-20 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5
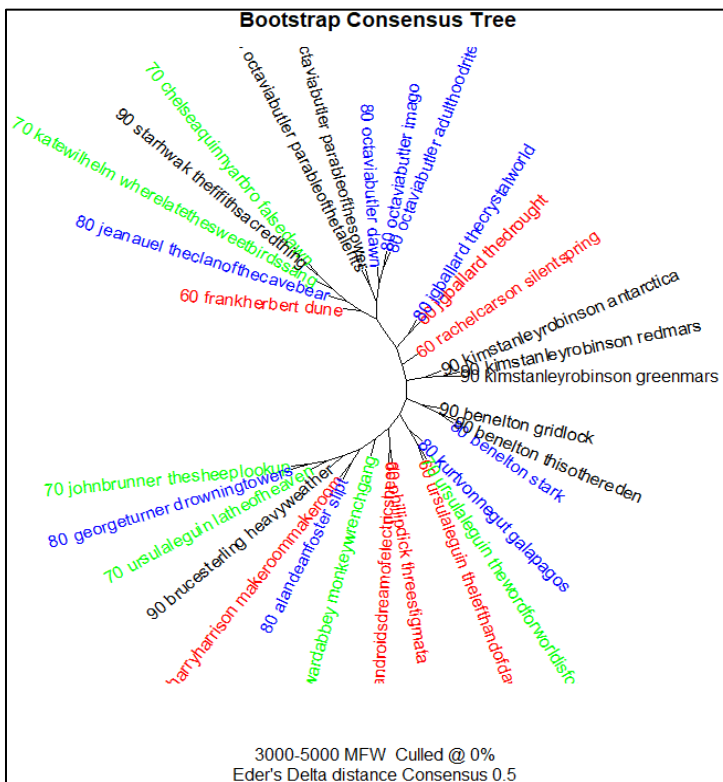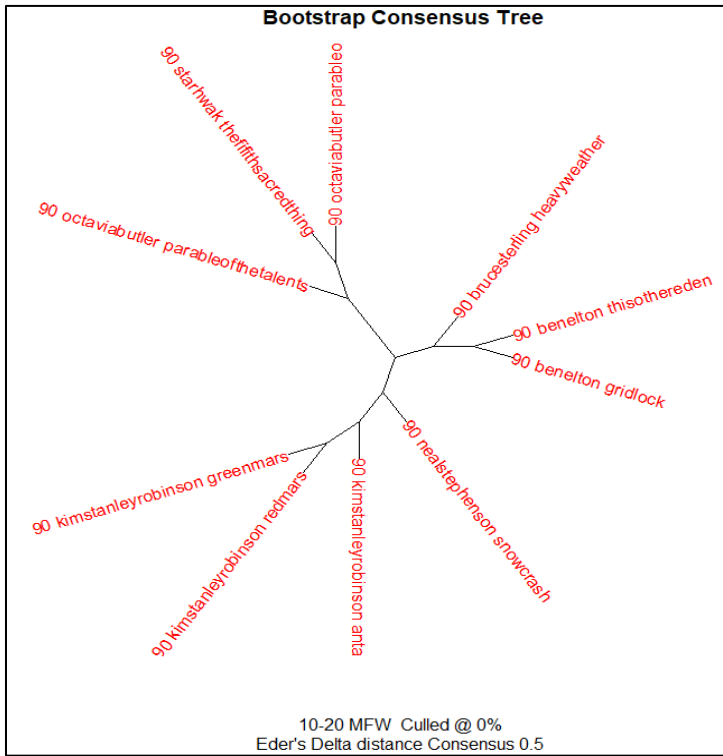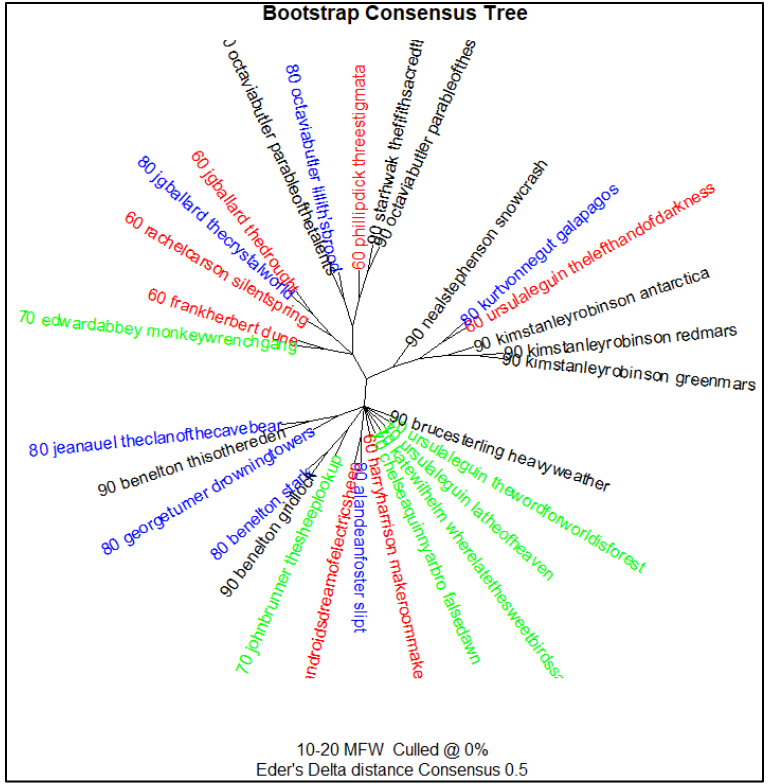
POS tagged bi grams comparison 1960s-1990s

10-20 MFW 2-grams Culled @ 0%
Eder's Delta distance Consensus 0.5
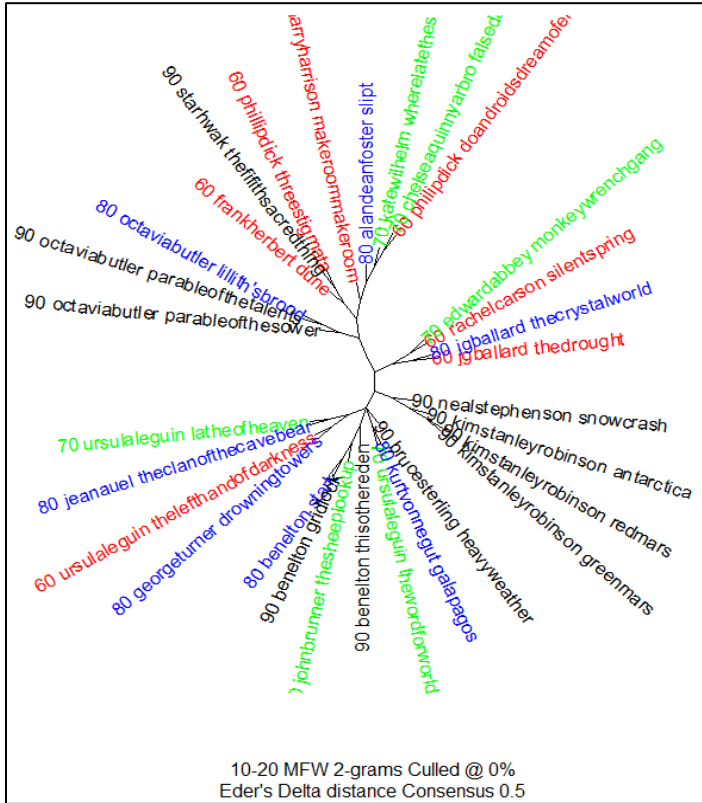
MFW Cluster analysis of the corpus for the 1960s-1990s corpora

Cluster Analysis

Further materials on the 2000s:

Bootstrap Consensus Tree

Further materials on the 2010s:

Bootstrap Consensus Tree

1000-2000 MFW Culled @ 0%
Eder's Delta distance Consensus 0.5

Principal Components Analysis

PC2 (5.9%)

PC1 (7.8%)
5000 MFW Culled @ 0%
Correlation matrix
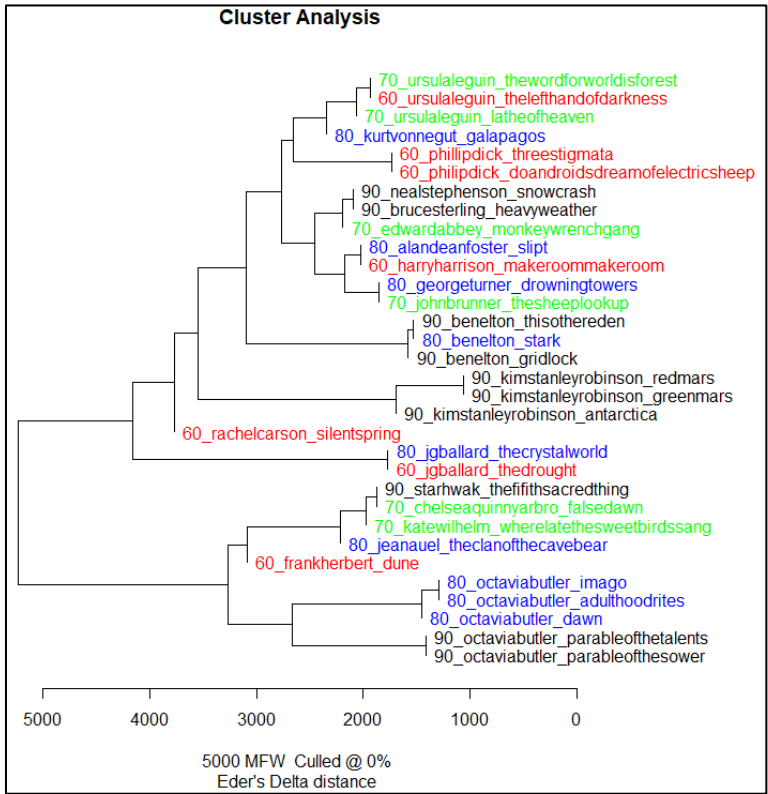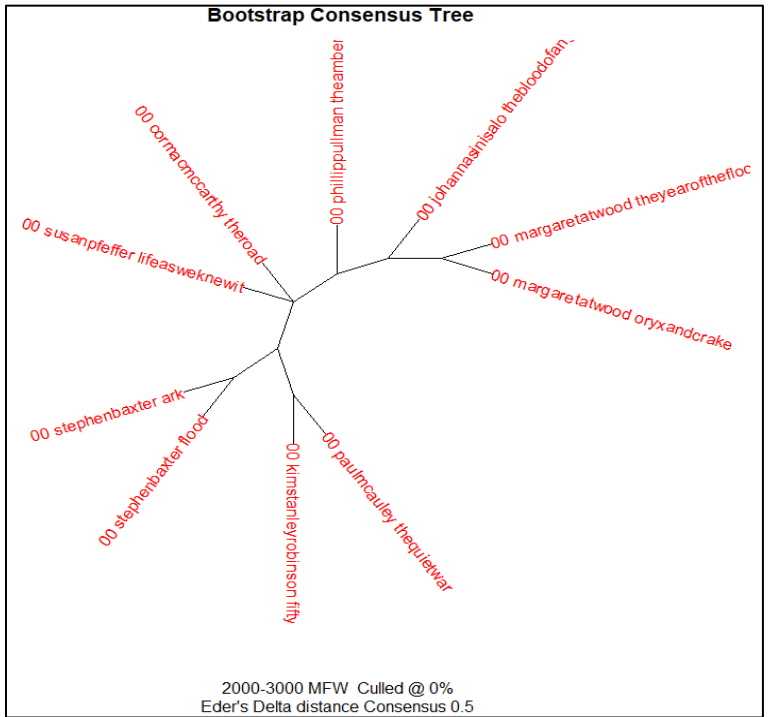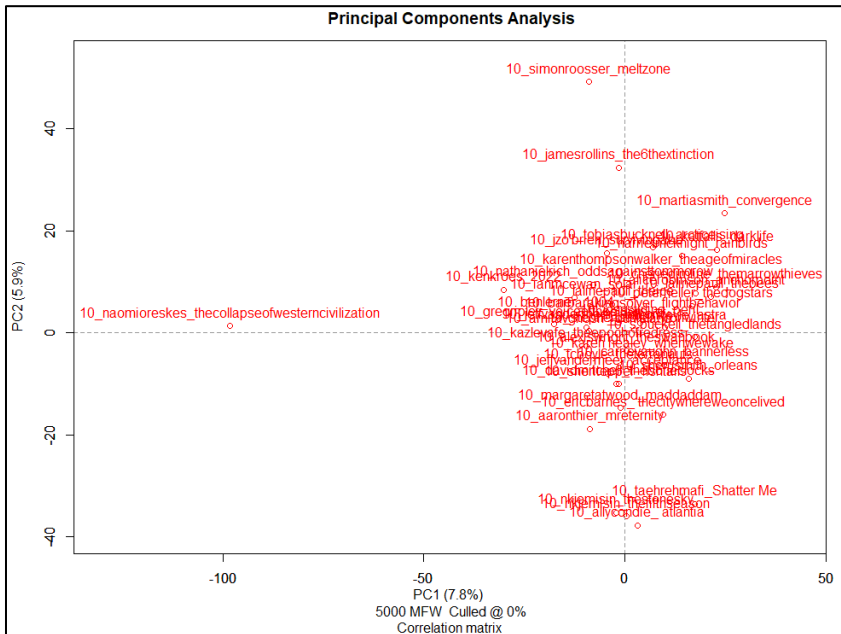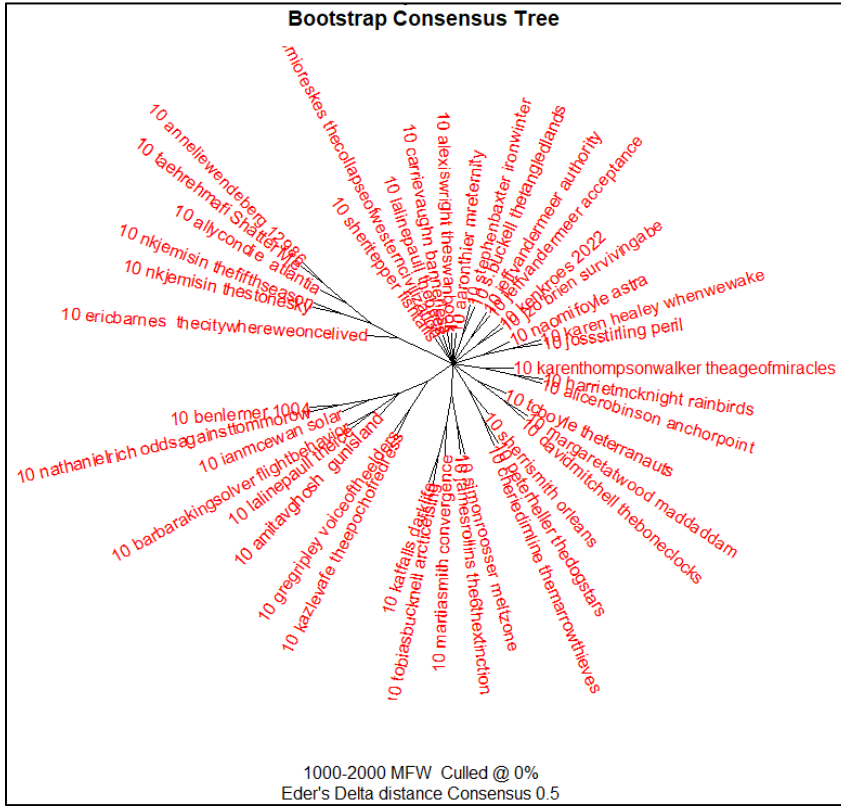
APPENDIX C

CODE AND INSTRUCTIONS

Instructions for converting the novels into txt files

There are three possible scenarios for your original file, either a digitally *born pdf,* which means it was written on a computer and you will be able to highlight lines within the document (follow instructions 1a), a non-digitally born pdf usually scans (follow instructions 1b), or the file is an eBook in the form of an epub (follow instructions 1c). Be sure to name the files consistently in the format YY_author_title.extension we will need this format for the analysis in R later.

1.a Conversion for digitally born pdf files

Step 1: Download the program xpdfreader the

extension pdftotext (https://www.xpdfreader.com/download.html) make sure both of them are in the same folder.

Step 2: In your PowerShell, copy the following code and change the highlighted sections to the folder containing the files you are looking to convert. Remember you can either find your current location with the command "pwd". To find the file path you need you can use file explorer, select the right folder and copy the lineup top that shows the current directory.

```
cd ***INSERT YOUR PATH***
$files = get-item *.pdf
foreach ($f in $files)
{***INSERT YOUR PATH*** $f}
```

1.b Conversion for non-digitally born pdf files

Step 1: Split the pdfs into individual pages, use any program you like (ex. Unity PDF). Save them like this YY_author_title_pageX.extension

Step 2: Convert each page in image format png/tiff using Imagemagick (Download

here: https://imagemagick.org/script/download.php). In your PowerShell, navigate to the folder

that contains the split pdfs with "cd" and your path. Then copy and run

these commands in PowerShell:

```
$files = get-item *.pdf
foreach ($f in $files)
{magick convert –density 300 $f –depth 8 –alpha off –background
$f.Name.replace(".pdf", ".tiff")}
```

Step 3: Convert the images to txt files, by copying and running his code in your PowerShell.

```
$files = get-item *.tiff
foreach($f in $files)
{tesseract –l eng $f $f.Name.replace(".tiff", ".txt")}
```

1.c Conversion for epubs

Step 1: For this file format is easiest to fall back on online solutions, so take your pick on which

version you like best: https://epubee.com/convert-epub-to-

txt.html or https://www.bitrecover.com/blog/epub-to-txt-converter/

2. Tagging the Corpus

Step 1: If you have not yet installed Python, be sure to install the appropriate version for your

computer. Open the Python launcher IDLE and copy the code below into a python file.

Step 2: Create two folders, one named "Corpus" containing your files and an empty one named

"Tagged". In your shell navigate to the location containing both subfolders.

Step 3: In your shell run the command "python ***path to your file***/FILENAME.py".

```
import os, re, nltk
democounter=1

def tagger(source, destination):
    with open("Corpus\\" + source, "r", encoding="UTF-8") as f:
        text=f.read()
        tokens=nltk.word_tokenize(text)
        #print(tokens)
        tags=nltk.pos_tag(tokens)
        for x, y in tags:
            destination.write((y) + " ")

    for file in os.listdir("Corpus"):
        destination=open("Tagged\\" + file, "w")
tagger(file, destination)
        democounter+=1
    print("Done.")
```

BIBLIOGAPHY

Definition of ANTHROPOCENE. www.merriam-webster.com/dictionary/Anthropocene.
  Accessed 12 Feb. 2020.

Al-Yahya, Maha. "Stylometric Analysis of Classical Arabic Texts for Genre Detection." *The
  Electronic Library,* vol. 36, no. 5, 2018, pp. 842–55. doi:10.1108/EL-11-2017-0236.

Andersen, Gregers. *Climate Fiction and Cultural Analysis: A New Perspective on Life in the
  Anthropocene*, Routledge, 2020. Routledge Environmental Literature, Culture and Media.

Canavan, Gerry, and Kim Stanley Robinson, editors. *Green Planets: Ecology and Science
  Fiction*, Wesleyan University Press, 2014.

Castree, Noel. "Nature." *The International Encyclopedia of Geography: People, the Earth,
  Environment, and Technology / Editor-in-Chief, Douglas Richardson ; General Editors, Noel
  Castree, Michael F. Goodchild, Audrey Kobayashi, Weidong Liu, Richard A. Marston,* edited
  by Douglas Richardson et al., vol. 41, Wiley Blackwell, 2017, pp. 1–26.

Chakrabarty, Dipesh. "Climate and Capital: On Conjoined Histories." *Critical Inquiry,* vol. 41,
  no. 1, 2014, pp. 1–23. doi:10.1086/678154.

Chris Turney, Jonathan Palmer, Mark Maslin. *Anthropocene Began in 1965, According to Signs
  in the World's 'Loneliest Tree'.* 1 Jan. 2018, phys.org/news/2018-02-anthropocene-began-
  world-loneliest-tree.html. Accessed 30 Jan. 2020.

Clark, Timothy. "The Deconstructive Turn." *symplokē,* vol. 21, 1-2, 2013, p. 11.
  doi:10.5250/symploke.21.1-2.0011.

Colebrook, Claire. "Framing the End of the Species." *symplokē,* vol. 21, 1-2, 2013, pp. 51–63.
  doi:10.5250/symploke.21.1-2.0051.

Crutzen, Paul J. "Geology of Mankind." *Nature,* vol. 415, no. 6867, 2002, p. 23.
  doi:10.1038/415023a.

DeLoughrey, Elizabeth M. *Allegories of the Anthropocene*, Duke University Press, 2019.

Eder, Maciej. "Visualization in Stylometry: Cluster Analysis Using Networks." *Digital
  Scholarship in the Humanities,* vol. 32, no. 1, 2017, pp. 50–64. doi:10.1093/llc/fqv061.

Eder, Maciej, et al. "An Open Stylometric System Based on Multilevel Text Analysis." *Cognitive
  Studies | Études cognitives, no.* 17, 2017, doi:10.11649/cs.1430.

Feldman, Anna, et al., editors. *Proceedings of the 3rd Workshop on Computational Linguistics
  for Literature (CLFL)*. Gothenburg, Sweden: Proceedings of the 3rd Workshop on
  Computational Linguistics for Literature (CLFL). Stroudsburg, PA, USA, Association for
  Computational Linguistics.

Ghosh, Amitav. *The Great Derangement: Climate Change and the Unthinkable.* Paperback edition, The University of Chicago Press, 2017. The Randy L. And Melvin R. Berlin Family Lectures.

Harré, Niki. *Psychology for a Better World: Strategies to Inspire Sustainability*, Dept. of Psychology University of Auckland], 2011.

Heise, Ursula K. *Imagining Extinction: The Cultural Meanings of Endangered Species*, The University of Chicago Press, 2016.

Howell, Rachel A. "Lights, Camera … Action? Altered Attitudes and Behaviour in Response to the Climate Change Film the Age of Stupid." *Global Environmental Change,* vol. 21, no. 1, 2011, pp. 177–87. doi: 10.1016/j.gloenvcha.2010.09.004.

Hulme, Mike. "Climate: Living Lexicon for the Environmental Humanities." *Environmental Humanities,* vol. 6, 2012, pp. 175–78. environmentalhumanities.org/arch/vol6/6.9.pdf.

Jockers, Matthew Lee. *Macroanalysis: Digital Methods and Literary History / Matthew L. Jockers*, University of Illinois Press, 2013. Topics in the Digital Humanities.

Johns-Putra, Adeline. *Climate Change and the Contemporary Novel*, Cambridge University Press, 2019.

Johns-Putra, Adeline, et al., editors. *Literature and Sustainability: Concept, Text and Culture*, Manchester University Press, 2017.

Kestemont, Mike. "Function Words in Authorship Attribution. From Black Magic to Theory?" *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL),* Gothenburg, Sweden, Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)edited by Anna Feldman et al. Stroudsburg, PA, USA, Association for Computational Linguistics, pp. 59–66.

Maibach, E., Rosenthal, S., Kotcher, J., Bergquist, P., Ballew, M., Goldberg, M., Gustafson, A. *Climate Change in the American Mind: April 2019.* 1 Jan. 2019.

Lewis, Simon L., and Mark A. Maslin. "Defining the Anthropocene." *Nature,* vol. 519, no. 7542, 2015, pp. 171–80. doi:10.1038/nature14258.

Margaret Cohen. "Narratology in the Archive of Literature." *Representations,* vol. 108, no. 1, 2009, pp. 51–75.

Mehnert, Antonia. *Climate Change Fictions: Representations of Global Warming in American Literature / Antonia Mehnert*, Palgrave Macmillan, 2016. Literatures, Cultures, and the Environment.

Milner, Andrew, and Burgmann, JR. "Climate Fiction: A World-Systems Approach." *Cultural Sociology,* vol. 12, no. 1, 2018, pp. 22–36. doi:10.1177/1749975517725670.

Milner, Andrew, and J. R. Burgmann. "A Short Pre-History of Climate Fiction." *Extrapolation,* vol. 59, no. 1, 2018, pp. 1–23. doi:10.3828/extr.2018.2.

Moretti, Franco. *Distant Reading*, Verso, 2013.

Morton, Timothy. *Hyperobjects: Philosophy and Ecology After the End of the World / Timothy Morton*, University of Minnesota Press, 2013. Posthumanities 27.

Nixon, Rob. *Slow Violence and the Environmentalism of the Poor.* 1. Harvard Univ. Press paperback ed., Harvard Univ. Press, 2013.

Paul J. Crutzen, Eugene F. Stoermer. *The "Anthropocene"*. Global Change Newsletter2000.

Ruddiman, William F. "The Anthropogenic Greenhouse Era Began Thousands of Years Ago." *Climatic Change,* vol. 61, no. 3, 2003, pp. 261–93. doi:10.1023/B: CLIM.0000004577.17928.fa.

Schneider-Mayerson. "The Influence of Climate Fiction: An Empirical Survey of Readers." *Environmental Humanities,* vol. 10, no. 2, 2018, pp. 473–500.

Steffen, Will, et al. "The Trajectory of the Anthropocene: The Great Acceleration." *The Anthropocene Review,* vol. 2, no. 1, 2015, pp. 81–98. doi:10.1177/2053019614564785.

Streeby, Shelley. *Imagining the Future of Climate Change: World-Making Through Science Fiction and Activism / Shelley Streeby*, University of California Press, 2018.

Trexler, Adam. *Anthropocene Fictions: The Novel in a Time of Climate Change / Adam Trexler*, University of Virginia Press, 2015. Under the Sign of Nature: Explorations in Ecocriticism.