Profiling and Analyzing Climate Change Statements in IPCC Reports

Anonymous ACL submission

Abstract

We propose new methods to extract and profile the climate change statements from the Sixth Assessment Reports of the Intergovernmental Panel on Climate Change (IPCC). We represent the 9,728 statements from the latest IPCC reports (AR6) with associated uncertainty levels and glossary terms. We profile their distributions across different parts of the 6000+ page AR6 reports. We also present a few case studies centered around the glossary term "wetland", namely linking related statements across summary sections and chapter content, finding and profiling supporting references, and comparing them with large language models for statement summarization. We believe this work marks an initial step towards in-depth information extraction regarding climate change. It lays the groundwork for more advanced automated analysis of climate-related statements and broader integrative scientific assessments.

Input: "…contributing to detectable increases in local rainfall and coastal flooding associated with these storms. There is *high confidence* (Seneviratne et al., 2021) that anthropogenic climate change has contributed to extreme precipitation associated with recent intense hurricanes, such as Harvey in 2017. North American sea ice… "

Output: {**text:** There is *high confidence* (Seneviratne et al., 2021) that anthropogenic climate change has contributed to extreme precipitation associated with recent intense hurricanes, such as Harvey in 2017.
confidence level: high confidence,
likelihood level: none,
source: {wg: WGII, chapter: 14, section: 14.2.1},
key_terms: [anthropogenic, climate, climate change]}

Figure 1: Turning IPCC statements into structured text. (Top) An input text segment from the IPCC AR6 WGII report. (Bottom) System output – semi-structured representation of the corresponding statement.

1 Introduction

A scientific statement is a factual statement which prescribes or entails the conditions for its verification (Miller, 1947). Statements are often viewed as a basic unit of scientific discourse. With the scientific literature growing rapidly in volume, keeping track of a large set of related statements is a widely-recognized challenge across diverse fields such as biomedicine, public health and law (Achakulvisut et al., 2019; Li et al., 2021; Wuehrl et al., 2023; Surdeanu et al., 2010; Li et al., 2022). Climate change poses a complex and multifaceted challenge to modern science (Klenk and Meehan, 2015). Keeping track of the scientific literature and scientific statements is also a pressing need in this domain, where research findings should inform critical policy decisions and enhance public understanding (Kasperson and Stern, 2010).

The Intergovernmental Panel on Climate Change (IPCC) is the United Nations (UN) body for assessing the science related to climate change (IPCC, 2024b). One important output from IPCC are the integrative assessment report series – the latest sixth Assessment Reports (AR6) were released between 2021 and 2023 (Arias et al., 2021; Adler et al., 2022; Shukla et al., 2022). These integrative assessment reports present both opportunities and challenges for computational tools for science. On the one hand, the IPCC has developed protocols to recruit experts (IPCC, 2024a), evaluate a large body of literature (IPCC, 2024b), and encode uncertainties and consensus (Mastrandrea et al., 2010) – which may mean that the scientific statements are more robust, and resulting documents are easier to automatically parse than research papers in general. On the other hand, AR6 alone has 743 authors and review editors (IPCC, 2024a). The result of this large-scale collaboration is an assessment report AR6 totalling 10,000+ pages. We posit that the volume of information here is still too large for anyone to read and comprehend; therefore automated tools will be helpful but largely missing – since most work on NLP for science, including climate science (Callaghan et al., 2021), focuses on paper collections.

This work takes the first few steps towards extracting information from IPCC reports. First, we design and implement a tool to extract scientific statements from IPCC Working Group (WG) reports. An example of the extracted statements, containing its text, uncertainty (i.e., confidence and likelihood) levels, source and key terms, is shown in Figure 1. Second, we present a comprehensive profile of 9,728 statements across three IPCC WG reports (Section 4). Profiling these statements with confidence and likelihood levels provides insights into the robustness and reliability of the information, which is crucial for informed decision-making. Our analysis shows that WGII has a higher proportion of *high* and *very-high* confidence statements, and 36.34% of statements appear in different summary content rather than chapter content. Additionally, profiling the distribution of key terms in statements across the reports helps in understanding the thematic focus and terminological consistency. Lastly, we present three case studies that take the first steps towards linking related statements (Section 5)to highlight the connections between different parts of the reports; identifying supporting references (Section 7) to provide a deeper context for the statements; and comparing statement summarization with those by large language models (Section 6 to assess the effectiveness of automated tools in summarizing complex scientific information.

We hope that this constitutes a useful first step towards analysing other integrated assessment reports (Mach and Field, 2017), which include and obviously not limited to the Millennium Ecosystem Assessment (Assessment, 2005), the Global Energy Assessment (Global Energy Assessment, 2012), and upcoming assessment on AI that forms the foundation for UN AI Governance (UN Advisory Body on Artificial Intelligence, 2023). We will release the statements at https://anonymous.com.

2 Related Work

Statement or claim extraction from scientific documents is a crucial task across various domains. In the biomedical field, Achakulvisut et al. (2019) utilize a Bi-LSTM with a CRF layer for statement extraction, while Li et al. (2021) and Wuehrl et al. (2023) introduce methods for tagging relevant text segments and leveraging entity recognition for statement verification, respectively. In legal texts, Surdeanu et al. (2010) employ a hierarchically structured model for extraction. For the COVID-19 pandemic, Li et al. (2022) build a system to extract, structure, and monitor statements from various sources in real-time. Unlike them, our work addresses a new problem in building NLP tools for climate change (Stede and Patz, 2021).

For the climate change domain, there is research that focuses on extracting climate-related statements as datasets for downstream tasks such as fact-checking. The datasets include Climate Fever (Diggelmann et al., 2020), Climate Feedback (Walter et al., 2020), and Skeptical Science (Winkler et al., 2021). However, these sources derive statements from social media, news, websites, etc., not from IPCC reports.Specifically, Lacombe et al. (2023) provide a dataset by extracting statements from three IPCC Sixth Assessment Reports, aligning with our goal. However, their PDF extraction method misses some statements, introduces inaccuracies, and overlooks statements with likelihood levels. Additionally, their classification based on confidence labels is unreasonable, as experienced climate experts consider multiple factors beyond the statements themselves.

3 IPCC Reports and Scientific Statements Therein

The IPCC Sixth Assessment Report comprises three Working Group (WG) reports, which are released sequentially from 2021 to 2022. They are WGI (2021), focusing on the Physical Science Basis; WGII (2022), on Impacts, Adaptation, and Vulnerability; and WGIII (2022), on Mitigation of Climate Change. The structure of each WG report consists of a Summary for Policymakers, a Technical Summary, and a set of numbered chapters. The three WG reports have 12, 18 and 17 chapters, respectively. We leave analyzing the Special Reports and Synthesis Report of AR6 as future work.

Statements are one of the scientific building blocks of IPCC reports, with each statement clearly categorized by confidence levels and likelihoods, to provide a nuanced and comprehensive overview of climate impacts and risks. *Confidence* and *likelihood* levels, which are key metrics used by the IPCC, express scientific uncertainty. Confidence levels in the IPCC assessment process reflect the validity of a statement based on the type, amount, and quality of evidence supporting it, while likelihood levels denote the probability of the occurrence of an event or outcome, calculated through statistical methods and expert judgment. The IPCC provides a framework that details the confidence and likelihood levels (Adler et al., 2022). The framework structure can be found in Figure 5 in Appendix A.1. Confidence is assessed using a 5-level scale that includes the categories of *very low*, *low*, *medium*,

Source	WGI			WGII			WGIII			Total		
	Page	Para.	Word	Page	Para.	Word	Page	Para.	Word	Page	Para.	Word
SummPol	32	132	9,243	34	45	5,140	51	581	22,046	117	758	36,429
TechSumm	112	340	34,449	84	207	28,728	102	499	45,053	298	1,046	108,230
ChapSumm	42	207	26,764	61	428	46,157	43	254	30,734	146	889	103,655
ChapBody	1,740	4,419	732,260	2,341	6,427	833,961	1,599	6,240	757,115	5,680	17,086	2,323,336
Total	1,926	5,098	802,716	2,520	7,107	913,986	1,795	7,574	854,948	6,241	19,779	2,571,650

Table 1: Basic profile of IPCC AR6 Working Group (WGI, WGII, WGIII) Reports, containing the number of pages, paragraphs and words by content type: Summary for Policymakers (SummPol), Technical Summary (TechSumm), Executive Summary of Chapters (ChapSumm) and the remaining Chapter contents (ChapCont).

high, and *very high* confidence. The likelihood is divided into 10 scales, from *exceptionally unlikely* (0-1%) to *virtually certain* (99-100%).

We store the three reports by first chunking them into paragraphs (excluding figures, plots, etc.) and injecting the paragraphs into the Elasticsearch[1] database to facilitate search and analysis. An overall dataset profile is in Table 1, showing that the three WG reports collectively have more than 6,000 pages, nearly 20K paragraphs, and more than 2.5 million words. Various summary content, including the Summaries for Policymakers (SummPol), Technical Summaries (TechSumm), and Chapter Executive Summaries (ChapSumm), constitutes approximately 10% of the entire AR6. Note that Annexes, Atlas, and Front Matter are excluded from Table 1.

4 Extracting and Profiling Statements from IPCC AR6

We propose a method to automatically extract scientific statements from IPCC reports and represent each statement s as a faceted tuple:

$$s = \{t,\ c,\ l,\ o,\ w\}.$$

Here t represents the statement text; c and l represent the confidence and likelihood level associated with statement s, respectively – either of which can be absent; o specifies the source of s in the IPCC reports, including the relevant working group, chapter, and section; w refers to a set of key terms from IPCC Glossary that appear in the statement text. The extraction method are described below, and an example extraction result is in Figure 1.

4.1 Extracting Statement Text

The data source for extraction is the HTML webpages[2][3][4] for the AR6 WGI, WGII and WGIII reports. While the PDF versions of the reports are available, we find HTML parsing more reliable despite recent developments in PDF extraction tools (Bast and Korzen, 2017; Meuschke et al., 2023). We assume that each statement has a confidence level or a likelihood level tag. In the HTML file, such tags are in italics (e.g., high confidence). We split the whole reports into individual sentences and extract sentences with the italic confidence or likelihood tags as statements.

The extracted statements are processed by: (1) eliminating footnotes embedded within sentences using the HTML tag "_idFootnoteLink _idGenColorInherit"; (2) filtering out statements with fewer than 50 characters. Such short and abnormal statements arise from the limitations of our sentence segmentation method, which depends on ending punctuation marks like periods and question marks; (3) The executive summary HTML of Chapter 9 in the WGIII report is not available. Therefore, we conduct a manual examination of the corresponding PDF and include the statements. Furthermore, we encounter instances where a sentence includes multiple statements, e.g., "...wetland vegetation will increase with SLR (*high confidence*)... impacts than deeper estuaries with well-developed sediments (*medium confidence*)...." (cf. the third statement in Table 3). Such complex sentences condense multiple pieces of information into a coherent structure, offering a concise summary that facilitates comparison. If these elements are separated, their meaning becomes fragmented. In such instances, we treat them as a single comprehensive statement and assign the confidence or likelihood level of the last-mentioned tag.

While all 5 confidence levels are found in the dataset, the 10 likelihood levels are used less consistently. In the reports, a few new wordings are found in the tags that aren't in the set of pre-defined likelihood levels, such as 'high certainty'. We then manually merge all of the variants into the given 10 scales. Details are in Appendix A.4.

[1] https://www.elastic.co/
[2] https://www.ipcc.ch/report/ar6/wg1/
[3] https://www.ipcc.ch/report/ar6/wg2/
[4] https://www.ipcc.ch/report/ar6/wg3/

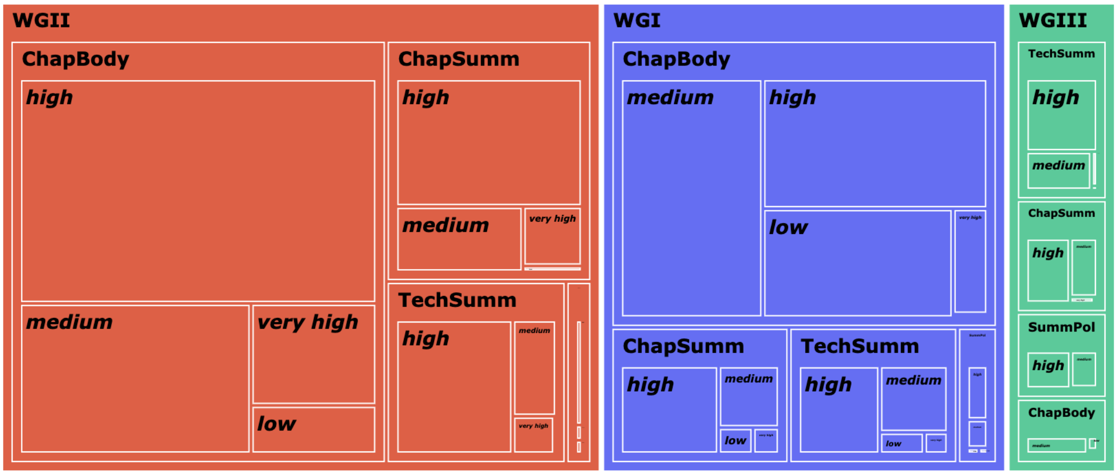

Figure 2: A treemap of statements by confidence levels (*very high, high, medium, low, very low*) and sources (ChapBody, ChapSumm, TechSumm, SummPol) for IPCC AR6 WGI, WGII, and WGIII reports. Block sizes correspond to the number of statements.

4.2 Detecting Glossary Terms in Statements

The IPCC includes a glossary at the end of its reports. This glossary comprises a collection of keywords accompanied by their respective definitions.

We collect terms from IPCC-glossary portal[5], by storing all the terms found in the glossaries of AR5, AR6, and the special reports published between AR5 and AR6. Variations of the same word, such as "aerosol" and "aerosols", are both present in the glossary. Given their grammatical and semantic similarities, we lemmatize and combine them into a single entity "aerosol". In total, we have identified 1,504 terms defined by IPCC that could potentially match the statement text.

To identify the presence of key terms in statements, we employ the `SpaCy`[6] tokenization and lemmatization tools from on both the terms and the statement text, then do the token-level matching. Additionally, we convert the statement text to lowercase and eliminate punctuation.

4.3 Overall statement profile

We obtained 9,728 statements, which is in excess of the 8,094 statements extracted by (Lacombe et al., 2023). We denote the 9,728 statements as set S; the subset of 8,581 statements with confidence levels as set $C = \{s \in S, \text{where } s_c \neq \phi\}$; the subset of 1,508 statements with likelihood levels as set $L = \{s \in S, \text{where } s_l \neq \phi\}$. Set C contains 3,117 statements from WGI, 4,620 from WGII, and 844 from WGIII. Set L includes 1,210 statements from WGI, 271 from WGII, and 27 from WGIII. There are 361 statements that include both confidence and likelihood levels. 91.2% of C and 89.2% of L contain at least one key term. The overall distribution aligns with observations on integrative assessment (Mach and Field, 2017) – confidence is most applicable when characterizing statements in WGII (on impacts, adaptation, and vulnerability) because cross-disciplinary evidence is often required for such inquiry. By contrast, likelihood is more common in WGI (on physical science) since statements could come from single lines of inquiry or similar inquiries whose likelihoods could be aggregated.

Figure 2 contains a breakdown of confidence levels across different parts of each WG report. In general, most of the statements in C are found within the chapter bodies. Over 90% of the overall statements have confidence levels above medium (i.e., *medium*, *high*, or *very high*). Specifically, *high confidence* is the most common confidence level for statements in most chapters, except for those in the chapter bodies of the WGI and WGIII reports. As for L, most of the statements are found in chapter bodies as well, and the majority of them have a *likely* label. A detailed distribution is shown in Table 4, Appendix A.1.

Figure 3 shows the frequency of key terms occurring in C and their distributions in the three WG reports. As expected, general concepts such as "climate", and "climate change" are dominant terms in the statements. Other popular topics in the climate change domain, like "emissions", "ecosystem", and "global warming", are also emphasized. More specific terms, such as "sea ice" and "RCP8.5", are covered by a certain number of statements as well. This reveals that our extracted statements

[5] https://apps.ipcc.ch/glossary/
[6] https://spacy.io/

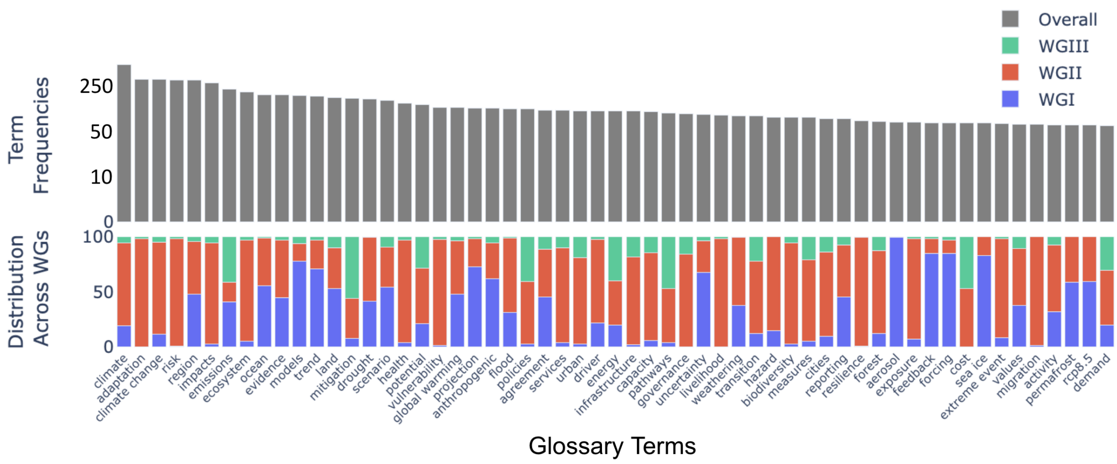

Figure 3: Frequency and breakdown of the top 57 terms (occurring 100 times or more) in set C of IPCC AR6 statements. Top:The number of times each term appears on log scale. Bottom: The proportion of each key term in each of WGI, II and III.

WGII 12.ES	Disruption in water flows will significantly degrade ecosystems such as high-elevation wetlands and affect farming communities, public health and energy production (*high confidence*).
WGII 12.3	Disruptions in water flows will significantly degrade or eliminate high-elevation wetlands (*high confidence*) (Bury et al., 2013; Dangles et al., 2017; Mark et al., 2017; Polk et al., 2017; Cuesta et al., 2019).

Table 2: One statement from D and one from N with a similarity score higher than threshold $\theta = 0.78$. Key terms in the statements are highlighted.

offer comprehensive coverage in summarizing a variety of climate change-related findings. Specifically, the most frequently occurring terms in the statements from different working groups closely align with the themes of each WG. Statements from WGI (The Physical Science Basis) focus on mitigation strategies and foundational terms such as "evidence", "global", and "anthropogenic". WGII (Impacts, Adaptation, and Vulnerability) emphasizes terms like "vulnerability", "risk", "impacts", and "adaptation". Meanwhile, WGIII (Mitigation of Climate Change) heavily utilizes terms such as "mitigation", "emissions", and "technological". These identified terms, especially when paired, can provide deep insights into thematic overlaps and interdependencies between different areas of climate science. For example, the combination of "emissions" and "mitigation" can highlight the direct relationship between the volume of emissions and the effectiveness of mitigation strategies. Furthermore, lower-frequency terms may uncover niche topics or emerging trends in climate science that have not yet reached mainstream recognition, presenting vital directions for our future research.

Glossary terms (and their combinations) can be further used to identify statements related to specific topics of interest. In the rest of this paper, we present a few case studies of linking, supporting, and comparing related scientific facts using statements.

5 Case Study 1: Linking Statements Across AR6

We define two statements to be **linked** if they convey similar meanings or ideas, and pertain to comparable contexts or topics. In this section, we link related statements across different parts of IPCC AR6. Since using the collection of 9000+ statements individually does not seem practical for readers of the IPCC report, we believe being able to identify topically similar and scientifically related groups is a tangible first step.

We build this case study around the glossary term "wetland" – chosen due to both its relevance in our geographical area and the fact that the number of related statements forms a small yet diverse set for building intuition and manual checking of validity. We focuses on two sets of statements: set N contains all 5,232 statements in Chapter bodies, and set D contains 12 statements containing "wetland" in the summary sections – SummPol, TechSumm and ChapSumm. Both are proper subsets of set C which contains all statements with confidence.

We compare these two sets of statements to highlight the contrast between the broader discourse and specific mentions in summaries. This comparison reveals the extent and consistency of wetland-related discussions across detailed and summary contexts, helping us understand the prominence of these issues.

Throughout this section, glossary terms are high-

WGII TS.D	Restoration of wetlands could support livelihoods and help sequester carbon (*medium confidence*), provided they are allowed accommodation space.
WGII 12.ES	Inclusive water regimes that overcome social inequalities and approaches including nature-based solutions, such as wetland restoration and water storage and infiltration infrastructure, with synergies for ecosystem conservation and disaster risk reduction, have been found to be more successful for adaptation and sustainable development (*high confidence*).
WGII 3.4	Without careful management of freshwater inputs, sediment augmentation and/or the restoration of shorelines to more natural states, transformation and loss of intertidal areas and wetland vegetation will increase with SLR (*high confidence*) (Doughty et al., 2019; Leuven et al., 2019; Yu et al., 2019; Raw et al., 2020; Shih, 2020; Stein et al., 2020), with small, shallow microtidal estuaries being more vulnerable to impacts than deeper estuaries with well-developed sediments (*medium confidence*) (Leuven et al., 2019; Williamson and Guinder, 2021).
WGIII 7.4	There is *medium confidence* that coastal wetland restoration has a technical potential of 0.3 (0.04–0.84) gtco2-eq yr –1 of which 0.1 (0.05–0.2) gtco2-eq yr –1 is available up to usd100 tco2–1.
WGIII 7.4	There is *high confidence* that coastal wetlands, especially mangroves, contain large carbon stocks relative to other ecosystems and *medium confidence* that restoration will reinstate pre-disturbance carbon sequestration rates.
WGIII 7.4	There is *low confidence* on the response of coastal wetlands to climate change; however, there is *high confidence* that coastal wetland restoration will provide a suite of valuable co-benefits.

Table 3: Statements that contain both "wetland" and "restoration" key terms.

lighted and categorized into nine distinct groups using a term-clustering scheme we developed with GPT-4, described in detail in Appendix A.3.

Measuring Similarity We posit that related statements should exhibit topical similarity. There appear to be two primary approaches: analyzing semantic similarity and filtering and intersecting by glossary terms. In this section, we first apply the semantic similarity analysis, followed by the glossary term method. To measure the pair-wise similarity, we first embed statements $d_i \in D$ and $n_j \in N$ into 1,536-dimensional vectors using the `text-embedding-3-small` model of OpenAI embedding API. We then calculate the cosine similarity between d_i and n_j.

The average similarity score between each d_i and all statements in N approximates 0.5 (detailed distribution is in Figure 6 in Appendix A.2). It suggests a moderate level of relatedness between statements in summary and base chapters.

To establish links between statements based on their similarity scores, we define a threshold $\theta \in [0, 1]$. Statements d_i and n_j are considered linked if their similarity score $Sim(d_i, n_j) > \theta$. We explored threshold values from 0.5 to 0.99 in increments of 0.01 and computed the average difference in similarity scores between linked and unlinked statement pairs. The trend of these differences (illustrated in Figure 7, Appendix A.2) reveals the steepest increase in the gap as θ transitions from 0.77 to 0.78. Consequently, we set $\theta = 0.78$, linking statements only when their similarity scores exceed this value.

Results and Discussion Table 2 shows the result that only one pair of statements from set D (of 12 wetland-related statements) and N are above the threshold, both originating from the same chapter's (WGII Chapter 12) executive summary and body. Upon reading these statements, we confirm that they are essentially the same statement pitched at different levels of detail. This outcome demonstrates the high precision of this semantic similarity-based method. However, the method may miss valid links as the recall is undetermined. This limitation may stem from the complexity in sentence structures and wording, and whole-sentence embedding may not adequately capture these nuances. The key terms in the base chapter statement (i.e., {"wetland"}) is a subset of those in the summary chapter statement (i.e., {"ecosystem", "wetland", "health", "energy"}). This observation prompts further investigation into whether key term overlaps could indicate potential links between statements.

To assess the potential for further matching, we (the authors of this work) examine all six statements that include the glossary term "restoration" (in blue) as well as "wetland" (in green) – two from summary chapters and four from chapter bodies, shown in Table 3. Despite the fact that glossary terms among these statements intersect, we did not identify any additional pairs that could be linked. For example, the first and fourth statements both mention that wetland restoration benefits carbon sequestration, but the first is broader and mentions additional benefits such as supporting livelihoods, while the fourth is more detailed and quantitative.

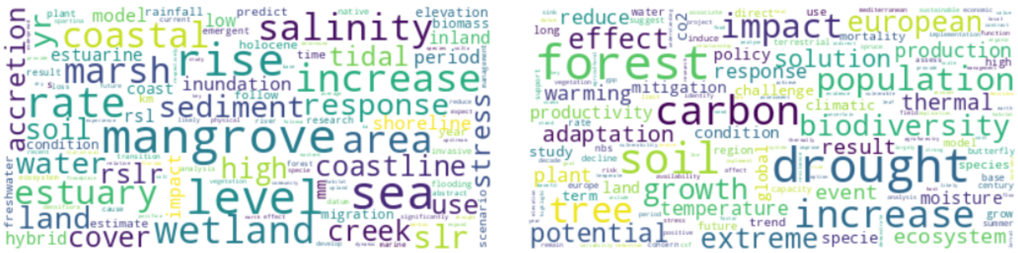

Figure 4: Word clouds generated from the abstracts of papers supporting the statements s_1(left) and s_2 (right).

It encourages us to explore the integration of multi-dimensional features for linking statements in future work, beyond mere semantic similarity or key term matching. For completeness, all 26 statements from Chapter body text with the glossary term "wetland", are listed in Table 8 in Appendix A.3, and we denote this set N'.

6 Case Study 2: Supporting References

In this section, we attempt to identify the scientific research supporting a statement by extracting its cited references (named **supporting references**). While this may be a trivial task for statements containing local citations, it becomes more complex in general because many statements summarize several paragraphs or reference entire sections that include numerous irrelevant citations.

Method We select one statement s_1 from the set N' as a case study: "Otherwise, wetland ecosystems must migrate either inland or upstream, or face gradual submergence in deeper, increasingly saline water (very high confidence) (Section 3.4.2.4; Andres et al., 2019; Jones et al., 2019b; Cohen et al., 2020; Mafi-Gholami et al., 2020; Magolan and Halls, 2020; Sklar et al., 2021)." This statement contains both a reference to WGII Section 3.4.2.4 and six local citations. From the section content, we aim to find the *evidence sentences* that contain the citation information that the statement may refer to. In particular, we select sentences that contain at least one local citation as candidate evidence sentences and compute similarities between the statement and each candidate evidence sentence using the same methodology as in Section 5.

Section 3.4.2.4 contains a total of 21 candidate evidence sentences and their resulting similarity scores range from 0.39 to 0.69. We select the three sentences with the highest scores (details in Table 6, Appendix A.5) together with the six local citations for a total of 17 supporting references to s_1.

Results and Discussion By obtaining and concatenating the abstracts of the 17 supporting references via OpenAlex[7], we built a word cloud, as shown on the left in Figure 4. Major keywords specific to s_1 that are present in its word cloud, including "mangrove" (i.e., one kind of wetland), "salinity", "increase" and "wetland", also occur in s_1, which indicates the relativity between the supporting references and s_1.

As a comparison, we perform the same analysis on another statement s_2: "Appropriately implemented ecosystem-based mitigation, such as reforestation with climate-resilient native species (Section 13.3.1.4), peatland and wetland restoration, and agroecology (Section 13.5.2), can enhance carbon sequestration or storage (medium confidence) (Seddon et al., 2020)." yielding a total of 14 supporting references, of which 13 are used to construct the word cloud on the right in Figure 4, as one reference was not found on OpenAlex. Contrasting with s_1, the major keywords for s_2 are "carbon", "forest", "mitigation", "tree", and "ecosystem". The difference in major keywords between the two word clouds supports our supporting reference detection methodology: Although both s_1 and s_2 mention "wetland", their context and emphasis differ significantly, aligning with their respective statements. This variation potentially validates the precision of our approach in using text-based analysis to extract and link supporting references to statements. However, the recall of this method still needs to be evaluated in future work.

7 Case Study 3: A Comparison with GPT Extracted Statements

We further conduct a case study to evaluate the quality of our generated statements by comparing them with those generated by large language models (LLMs).

Method We focus on the question "What are the main scientific statements on *wetland restoration* in IPCC reports?" The specific LLM used was the GPT-4 model (OpenAI, 2023) and the details

[7] https://docs.openalex.org/

of the prompt given to GPT-4 are detailed in Appendix A.6. Inspired by the pioneering work of ChatClimate (Vaghefi et al., 2023), which builds a retrieval-augmented-generation (RAG)-based conversational LLM using IPCC reports, we explored two methodologies: (1) pure zero-shot learning with GPT, where we provided the prompt directly to the model for statement extraction; and (2) RAG-based GPT, which involved enhancing the GPT-4 model's performance by providing the top five retrieved IPCC paragraphs relevant to the query. These paragraphs were selected based on the cosine similarity between each paragraph in our Elasticsearch database and the query. Additionally, we extracted statements from our database that contained the key terms "wetland" and "restoration" (as shown in Table 3) for comparison.

Results and Discussion The full results generated by the three methods are presented in Table 7 in Appendix A.6. Unlike our method, the zero-shot GPT model often produces statements that cite inaccurate IPCC sections. For instance, all three generated statements that cite "WGII Section 6.5" are incorrect – The term "wetland" does not appear in that section. Furthermore, the RAG-based GPT model, assisted by the top five retrieved paragraphs, exhibits improved accuracy in identifying IPCC sections related to wetland restoration (e.g., IPCC WGIII Section 7.4). However, it still tends to excessively condense content and generate hallucinations, similar to the zero-shot GPT model. This decreases the quality of the generated statements. For instance, consider the sentence "Their restoration and rewetting is crucial to meet 1.5°C–2°C pathways by 2050" from the second statement generated by the RAG-based GPT model. In the cited section (i.e., IPCC WGIII Section 7.4), we find sentences such as "... both peatland protection and peatland restoration (Section 7.4.2.7) are needed to achieve a 2°C mitigation ..." and "... peatlands, coastal wetlands, and forests are particularly important as most carbon lost from these ecosystems is irrecoverable through restoration by the 2050 timeline ...". However, there is insufficient evidence to justify summarizing these specific details into the broader statement provided by the model.

Thus, only our statements contain the scientific publication information that the statement refers to, e.g., "... Doughty et al., 2019; Leuven et al., ..." (cf. the third statement in Table 3). As mentioned in Section 6, such references provide important scientific evidence supporting the statements. Additionally, the GPT-generated statements lack uncertainty assessment information as we do. Confidence and likelihood levels are crucial for evaluating the validity and probability of the statements. On the other hand, our own generated statements also face issues: they are not comprehensive enough because we directly select sentences from the IPCC reports. For example, it is difficult for readers to fully understand the fourth statement "... wetland restoration has a technical potential of 0.3 (0.04–0.84) gtco2-eq yr –1 of which 0.1 (0.05–0.2) gtco2-eq yr –1 is available up to usd100 tco2–1." that we generated, as it stands alone with no in-context information. Providing background information such as explanations of terms (e.g., "gtco2-eq") may potentially enhance comprehension, which urges us to seek engagement with more climate experts in the future.

8 Conclusion

Reading, comprehending, and tracking scientific statements in large-scale literature, especially in the complex climate change domain, is a critical but challenging task. In this paper, we take the first few steps towards profiling and analyzing statements from the IPCC assessment reports. By automating the process, we provide researchers, policymakers, and stakeholders with a more accessible way to navigate the extensive and complex information found in IPCC reports. We aim to enable more informed decision-making and foster a deeper understanding of climate change dynamics.

Discussion We reflect on several limitations in the current dataset and methods, which could guide our future work. Previous IPCC assessment reports such as AR5 and AR4 are only available in PDF format, necessitating the exploration of advanced PDF parsing tools. Once extracted, evolution of statements across the different assessment reports over the last few decades could be explored. The results of linking statements (cf. Section 5) underscore the current challenges in understanding complex climate-related statements using matching-based and data-driven methods. Additionally, as mentioned in Section 7, a systematic evaluation is required to assess the coverage and validity of the supporting references. We believe that further engaging in cross-disciplinary collaborations involving climate scientists and linguists can enhance our interpretation of statements and help pave the way for designing tools that can ultimately help scientists, policy-makers and other stakeholders.

References

Titipat Achakulvisut, Chandra Bhagavatula, Daniel Acuna, and Konrad Kording. 2019. Claim extraction in biomedical publications using deep discourse model and transfer learning. *arXiv preprint arXiv:1907.00962*.

C. Adler, P. Wester, I. Bhatt, C. Huggel, G.E. Insarov, M.D. Morecroft, V. Muccione, and A. Prakash. 2022. *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Paola Arias, Nicolas Bellouin, Erika Coppola, Christopher Jones, Gerhard Krinner, Jochem Marotzke, Vaishali Naik, Gian-Kasper Plattner, Maisa Rojas, Jana Sillmann, Trude Storelvmo, Peter Thorne, Blair Trewin, Krishna Achutarao, Bhupesh Adhikary, Kyle Armour, Govindasamy Bala, Rondrotiana Barimalala, Sophie Berger, and Kirsten Zickfeld. 2021. Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change; technical summary.

Millennium Ecosystem Assessment. 2005. *Millennium ecosystem assessment*. Island Press, Washington, DC.

Hannah Bast and Claudius Korzen. 2017. A benchmark and evaluation for text extraction from pdf. In *2017 ACM/IEEE joint conference on digital libraries (JCDL)*, pages 1–10. IEEE.

Max Callaghan, Carl-Friedrich Schleussner, Shruti Nath, Quentin Lejeune, Thomas R Knutson, Markus Reichstein, Gerrit Hansen, Emily Theokritoff, Marina Andrijevic, Robert J Brecha, et al. 2021. Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nature climate change*, 11(11):966–972.

Thomas Diggelmann, Jordan Boyd-Graber, Jannis Bulian, Massimiliano Ciaramita, and Markus Leippold. 2020. Climate-fever: A dataset for verification of real-world climate claims. *arXiv preprint arXiv:2012.00614*.

Global Energy Assessment. 2012. *Global Energy Assessment —- Toward a Sustainable Future*. Cambridge Univ. Press, Cambridge UK.

IPCC. 2024a. How does the IPCC select its authors? https://www.ipcc.ch/site/assets/uploads/2024/04/IPCCFactSheet_SelectAuthors.pdf. Accessed May 2024.

IPCC. 2024b. What is the IPCC? https://www.ipcc.ch/site/assets/uploads/2024/04/IPCCFactSheet_WhatisIPCC.pdf. Accessed May 2024.

Roger E Kasperson and Paul C Stern. 2010. *Facilitating Climate Change Responses: A Report of Two Workshops on Knowledge from the Social and Behavioral Sciences*. National Academies Press.

Nicole Klenk and Katie Meehan. 2015. Climate change and transdisciplinary science: Problematizing the integration imperative. *Environmental science & policy*, 54:160–167.

Romain Lacombe, Kerrie Wu, and Eddie Dilworth. 2023. Climatex: Do llms accurately assess human expert confidence in climate statements? *arXiv preprint arXiv:2311.17107*.

Manling Li, Revanth Gangi Reddy, Ziqi Wang, Yishyuan Chiang, Tuan Lai, Pengfei Yu, Zixuan Zhang, and Heng Ji. 2022. COVID-19 claim radar: A structured claim extraction and tracking system. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 135–144, Dublin, Ireland. Association for Computational Linguistics.

Xiangci Li, Gully Burns, and Nanyun Peng. 2021. Scientific discourse tagging for evidence extraction. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2550–2562.

Katharine J Mach and Christopher B Field. 2017. Toward the next generation of assessment. *Annual Review of Environment and Resources*, 42:569–597.

Michael D. Mastrandrea, Christopher B. Field, Thomas F. Stocker, Ottmar Edenhofer, Kristie L. Ebi, David J. Frame, Hermann Held, Elmar Kriegler, Katharine J. Mach, Patrick R. Matschoss, Gian-Kasper Plattner, Gary W. Yohe, and Francis W. Zwiers. 2010. Guidance note for lead authors of the ipcc fifth assessment report on consistent treatment of uncertainties. https://www.ipcc.ch/site/assets/uploads/2017/08/AR5_Uncertainty_Guidance_Note.pdf.

Norman Meuschke, Apurva Jagdale, Timo Spinde, Jelena Mitrović, and Bela Gipp. 2023. A benchmark of pdf information extraction tools using a multi-task and multi-domain evaluation framework for academic documents. In *International Conference on Information*, pages 383–405. Springer.

David L Miller. 1947. The nature of scientific statements. *Philosophy of Science*, 14(3):219–223.

OpenAI. 2023. Gpt-4: Generative Pre-trained Transformer 4. https://openai.com/research/gpt-4.

P.R. Shukla, J. Skea, R. Slade, A. Al Khourdajie, R. van Diemen, D. McCollum, M. Pathak, S. Some, P. Vyas, R. Fradera, M. Belkacemi, A. Hasija, G. Lisboa, S. Luz, and J. Malley (eds.). 2022. *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK.

Manfred Stede and Ronny Patz. 2021. The climate change debate and natural language processing. In *Proceedings of the 1st Workshop on NLP for Positive Impact*, pages 8–18.

Mihai Surdeanu, Ramesh Nallapati, and Christopher Manning. 2010. Legal claim identification: Information extraction with hierarchically labeled data. In *LREC Workshop*, page 22. Citeseer.

UN Advisory Body on Artificial Intelligence. 2023. Interim report: Governing AI for humanity. https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_governing_ai_for_humanity_interim_report.pdf. Accessed: May 2024.

Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Mathias Kraus, Simon Allen, Chiara Colesanti-Senni, Tobias Wekhof, Tobias Schimanski, et al. 2023. Chatclimate: Grounding conversational ai in climate science. *Communications Earth & Environment*, 4(1):480.

Stefanie Walter, Janne Görlach, and Michael Brüggemann. 2020. Climate feedback: Science comments on journalism and develops multi-system competence. *Publizistik*, 65:567–589.

Bärbel Winkler, John Cook, Timo Lubitz, and Ken Rice. 2021. Skeptical science. In *World Scientific Encyclopedia of Climate Change: Case Studies of Climate Risk, Action, and Opportunity Volume 1*, pages 301–314. World Scientific.

Amelie Wuehrl, Lara Grimminger, and Roman Klinger. 2023. An entity-based claim extraction pipeline for real-world biomedical fact-checking. In *Proceedings of the Sixth Fact Extraction and VERification Workshop (FEVER)*, pages 29–37, Dubrovnik, Croatia. Association for Computational Linguistics.

A Appendix

A.1

IPCC conducted guidelines for determining the degree of certainty of statements, which is mapped into the confidence and likelihood levels in our statement profile. The guideline framework is shown in Figure 5.

Besides the distribution of the statements with confidence levels (C) shown in the main context (cf. Figure 2), the distribution of the statements according to the likelihood levels (L) is shown in Table 4. Similar to C, the majority of statements in L also located in chapter bodies. And the likelihood labels are quite imbalanced, over 80% of L are with a *likely* label or a *very likely* label.

A.2 Semantic Similarity between Statements

As described in Section 5, we calculate the cosine similarity between each statement (d_i) in the 12-statement set D and all the statements in N. Figure 6 illustrates the distribution of similarity scores for each d_i, with each box representing these scores. To determine the threshold θ for defining similar statements, we compared the average similarity score between linked and unlinked statements under various candidate thresholds. The results are shown in Figure 7. Notably, the steepest increase in the gap occurs between 0.77 and 0.78, leading us to set $\theta = 0.78$.

A.3 Statements with the Key Term ‘wetland’

The 38 statements that contain the key term “Wetland” are shown in Table 3.

And the categories of the key terms that are highlighted in different colors in the Table 3 are generated by asking ChatGPT to categorize the key terms three times and the ensemble result of the three-time categorization. The mapping dictionary is defined in Table 5, each cluster of key terms is named as a category based on their meanings in climate change domain.

A.4 Merging Likelihood Level Variations

We (the authors of this paper) found 12 variations of the likelihood levels and manually combined them based on domain knowledge. The dictionary provided below illustrates the mapping relationships, where the variation is the key and the corresponding original likelihood level is the value.

{ virtual certainty: virtually certain, very likely to be virtually certain: ’virtually certain’,
high certainty: virtually certain, highly likely: very likely, more or less likely: likely, more likely: likely, likely than not: more likely than not, as likely as not: about as likely as not,
less likely: unlikely,
not likely: unlikely,
large uncertainty: unlikely, deep uncertainty: extremely unlikely }

Out of the 1,508 statements in set L, 29 of them have variant levels of likelihood and have been matched to the original 10 scales using the dictionary above.

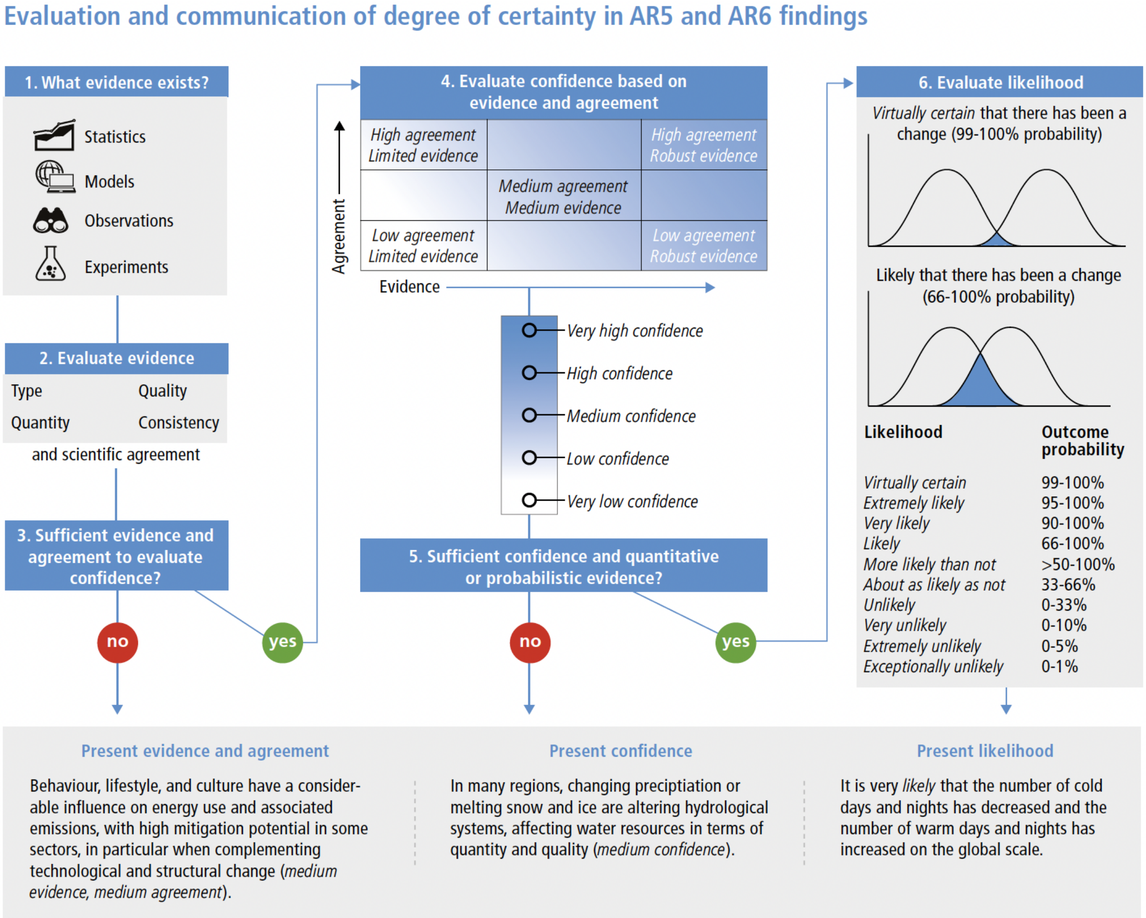

Figure 5: The IPCC AR6 framework for applying expert judgment in the evaluation of degrees such as confidence and likelihood of statements (cf. IPCC AR6 WGII Figure TS.1).

	Total	**SummPol**	**TechSumm**	**Chapters**	
				ChapSumm	**ChapBody**
Virtually certain	184 (12.20%)	3 (7.69%)	38 (24.52%)	22 (14.10%)	121 (10.45%)
Extremely likely	30 (1.99%)	0	4 (2.58%)	1 (0.64%)	25 (2.16%)
Very likely	450 (29.84%)	18 (46.15%)	38 (24.52%)	38 (24.36%)	356 (30.74%)
Likely	772 (51.19%)	17 (43.59%)	73 (47.10%)	90 (57.69%)	592 (51.12%)
More likely than not	11 (0.73%)	0	1 (0.65%)	1 (0.64%)	9 (0.78%)
About as likely as not	4 (0.27%)	0	0	1 (0.64%)	4 (0.35%)
Unlikely	26 (1.72%)	0	0	0	25 (2.16%)
Very unlikely	11 (0.73%)	0	0	2 (1.28%)	9 (0.78%)
Extremely unlikely	20 (1.33%)	1 (2.56%)	1 (0.65%)	1 (0.64%)	17 (1.47%)
Total	1,508	39	155	156	1,158

Table 4: Distribution of likelihood levels in Statements, including the number of statements by likelihood label: Summary for Policymakers (SummPol), Technical Summary (TechSumm), Executive Summary of Chapters (ChapSumm) and the remaining Chapter contents (ChapCont).

A.5 Evidence Sentences for Reference Checking

Given the 21 candidate evidence sentences and the statement s_1, the top three evidence sentences with the highest similarity scores are shown in Table 6.

Based on the abstracts of the 17 cited papers in the three evidence sentences, we build a word cloud of the concatenated abstracts. Specifically, we apply tokenization and lemmatization using SpaCy, and remove English stopwords (along with 'cli-

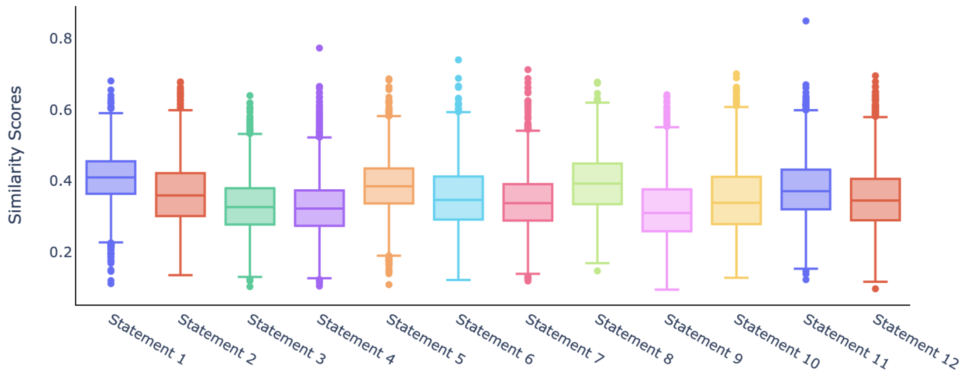

Figure 6: Comparative distribution of semantic similarity scores between each statement in D (Statement 1–12) and all statements in N.

Category	**Key Terms**
Climate Processes	Climate variability, Climate change, Global warming, Green infrastructure
Climate Impact	Sea level rise (SLR), Flood, Drought, Heat island, Impacts
Ecosystem Services	Biodiversity, Carbon sequestration, Ecosystem, Cultural services, Wetland
Climate Response	Adaptation, Mitigation, Resilience, Sustainable development, Restoration, Reforestation
Risk and Vulnerability	Risk, Disaster risk, Vulnerability, Uncertainty, Trade-off
Energy and Resources	Fossil fuels, Bioenergy, Energy, Biomass
Socio-economic Factors	Inequality, Livelihood, Access (to food), Health, Opportunities
Management and Policy	Adaptive capacity, Capacity, Forest management, Land use, Emissions, Agreement
Miscellaneous	Assets, Stock, Driver, Technical potential, Region, Confidence, Urban, Potential, Infrastructure, Measures, Transformation, Drivers, Soil erosion, Land, Scenario, Emission scenario, Peatlands, Evidence, Cost, Settlements, Sequestration, Carbon stock, Forest, Extinction, Agroecology, Co-benefits

Table 5: The ChatGPT categorized key term categories.

Similarity score	**Evidence sentence**
0.69	Elevated water levels also alter submergence patterns for intertidal habitat (high confidence) (Andres et al., 2019), moving high-water levels inland (high confidence) (Peteet et al., 2018; Appeaning Addo et al., 2020; Liu et al., 2020e) and increasing the salinity of coastal water tables and soils (high confidence) (Eswar et al., 2021).
0.65	These processes favour inland and/or upstream migration of intertidal habitat, where it is unconstrained by infrastructure, topography or other environmental features (high confidence) (Kirwan and Gedan, 2019; Parker and Boyer, 2019; Langston et al., 2020; Magolan and Halls, 2020; Saintilan et al., 2020).
0.63	Along estuarine shorelines, changing submergence patterns and upstream penetration of saline waters interact synergistically to stress intertidal plants, changing species composition and reducing above-ground biomass, in some cases favouring invasive species (Xue et al., 2018; Buffington et al., 2020; Gallego-Tévar et al., 2020).

Table 6: The three most similar evidence sentences to s_1 and their corresponding similarity scores.

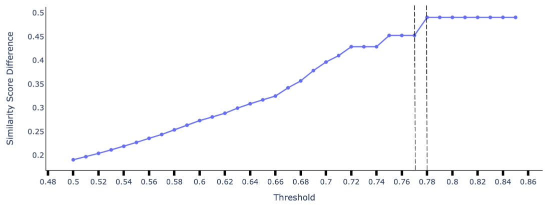

Figure 7: Setting different threshold θ, the corresponding difference in average similarity score between linked and not linked statement pairs. The vertical dashed lines indicate the most steep gap between two possible thresholds, we chose $\theta = 0.78$.

mate", 'change' and any single-character words).

A.6 Comparison with ChatGPT generated Statement of 'wetland restoration'

To compare our statement extraction results (cr. Table 3) with large language models (LLMs), we further employ GPT-4[8] model to extract statements on 'wetland restoration'. The full prompt is:

> "You are a chatbot with knowledge in climate change and IPCC report.
>
> Using the prior knowledge in climate question, answer the user's question, citing references back to the report whenever possible.
>
> Provide examples whenever possible.
>
> Use clear, simple and concise language.
>
> When citing references to the IPCC report, return a link in Markdown format. E.g. if the citing Working Group (WG2) Chapter 3, Section 3.5.2, then return [IPCC WG2 3.5.2](https://www.ipcc.ch/report/ar6/wg2/chapter/chapter-3/#3.5.2)"

Given the prompt and the query as "What are the main scientific statements on wetland restoration?", the full responses generated by vanilla GPT and RAG-based GPT, together with our retrieved statements, are listed in Table 7.

[8]https://chatgpt.com/?model=gpt-4

Vanilla GPT-4	RAG-based GPT-4	Ours.
Wetland restoration is a critical aspect of climate change mitigation and adaptation, as highlighted in the IPCC reports. Here are some main scientific statements on wetland restoration based on the IPCC reports: **1. Carbon Sequestration:** Wetlands, including peatlands, mangroves, and salt marshes, are highly effective carbon sinks, storing significant amounts of carbon dioxide (CO2) and other greenhouse gases. Restoring degraded wetlands can enhance their capacity to sequester carbon, helping to mitigate climate change (IPCC WG1 5.4). **2.Biodiversity Conservation:** Wetland restoration contributes to biodiversity conservation by providing habitat for numerous plant and animal species. This enhances ecosystem resilience and adaptive capacity in the face of climate change impacts (IPCC WG2 6.5). **3. Flood Regulation:** Healthy wetlands play a crucial role in flood regulation by absorbing and storing excess water during heavy rainfall events. Wetland restoration can help reduce the risk of flooding in vulnerable areas, protecting communities and infrastructure (IPCC WG2 6.5). **4. Coastal Protection:** Coastal wetlands, such as mangroves and salt marshes, act as natural barriers against coastal erosion and storm surges. Restoring degraded coastal wetlands can enhance coastal protection, reducing the vulnerability of coastal communities to sea-level rise and extreme weather events (IPCC WG2 6.5). **5. Water Quality Improvement:** Wetlands play a vital role in purifying water by filtering out pollutants and excess nutrients. Restoring degraded wetlands can improve water quality, benefiting both ecosystems and human populations dependent on clean water sources (IPCC WG3 5.4). Overall, the IPCC reports emphasize the importance of wetland restoration as a nature-based solution for climate change mitigation and adaptation, offering multiple co-benefits for biodiversity, ecosystem services, and human well-being.	There are several main scientific statements on wetland restoration based on the IPCC reports: Coastal wetland restoration has a technical potential of 0.3 GtCO2-eq yr −1 and it offers valuable co-benefits including enhanced water quality, biodiversity, aesthetic values, and protection from rising sea levels. It can also help with fisheries production, contributing to food security IPCC WG3 7.4. Restoration of coastal wetlands would decrease CO2 emissions and net GHG emissions from degraded peatlands as per recent literature with high to medium confidence levels. Their restoration and rewetting is crucial to meet 1.5°C–2°C pathways by 2050 IPCC WG3 7.4. Wetland restoration and preservation restore coastal protection and reduce community vulnerability to flooding. They also maintain coastal ecosystem services. Restoring coral reefs, oyster reefs, and mangroves dissipate wave energy, accrete sediment and elevate shorelines, thereby reducing exposure to waves and storm surges IPCC WG2 3.6. Restoration of peatlands helps in the increased accumulation of carbon in vegetation and soils and avoids ongoing CO2 emissions. They deliver co-benefits for biodiversity, regulate water flow, prevent downstream flooding and reduce the risk of fire IPCC WG3 7.4. It is important to note that there are risks involved in the restoration of wetlands, such as uncertain permanence under future climate scenarios, increased methane and nitrous oxide release, and conflict with other land uses IPCC WG3 7.4. Therefore, restoration projects should involve local communities at all stages and consider both biophysical and socio-political contexts.	• (WGII TS.D) Restoration of wetlands could support livelihoods and help sequester carbon (*medium confidence*), provided they are allowed accommodation space. • (WGII 12.ES) Inclusive water regimes that overcome social inequalities and approaches including nature-based solutions, such as wetland restoration and water storage and infiltration infrastructure, with synergies for ecosystem conservation and disaster risk reduction, have been found to be more successful for adaptation and sustainable development (*high confidence*). • (WGII 3.4) Without careful management of freshwater inputs, sediment augmentation and/or the restoration of shorelines to more natural states, transformation and loss of intertidal areas and wetland vegetation will increase with SLR (*high confidence*) (Doughty et al., 2019; Leuven et al., 2019; Yu et al., 2019; Raw et al., 2020; Shih, 2020; Stein et al., 2020), with small, shallow microtidal estuaries being more vulnerable to impacts than deeper estuaries with well-developed sediments (*medium confidence*) (Leuven et al., 2019; Williamson and Guinder, 2021). • (WGIII 7.4) There is *medium confidence* that coastal wetland restoration has a technical potential of 0.3 (0.04–0.84) gtco2-eq yr −1 of which 0.1 (0.05–0.2) gtco2-eq yr −1 is available up to usd100 tco2−1. • (WGIII 7.4) There is *high confidence* that coastal wetlands, especially mangroves, contain large carbon stocks relative to other ecosystems and *medium confidence* that restoration will reinstate pre-disturbance carbon sequestration rates. • (WGIII 7.4) There is *low confidence* on the response of coastal wetlands to climate change; however, there is *high confidence* that coastal wetland restoration will provide a suite of valuable co-benefits.

Table 7: Zero-shot learning GPT, RAG-based GPT and our retrieved ‘wetland restoration’-related statements.

WG	Chapter	Section	Text
WGI	TechSumm	TechSumm.2	There is *high confidence* that this recent growth is largely driven by emissions from fossil fuel exploitation, livestock, and waste, with ENSO driving multi-annual variability of wetland and biomass burning emissions.
WGII	SummPol	SummPol.C	Coastal wetlands protect against coastal erosion and flooding associated with storms and sea level rise where sufficient space and adequate habitats are available until rates of sea level rise exceed natural adaptive capacity to build sediment (*very high confidence*).
WGII	SummPol	SummPol.C	Natural river systems, wetlands and upstream forest ecosystems reduce flood risk by storing water and slowing water flow, in most circumstances (*high confidence*).
WGII	SummPol	SummPol.C	Enhancing natural water retention such as by restoring wetlands and rivers, land use planning such as no build zones or upstream forest management, can further reduce flood risk (*medium confidence*).
WGII	TechSumm	TechSumm.C	TS.C.5.1 Under all emissions scenarios, coastal wetlands will likely face high risk from sea level rise in the mid-term (*medium confidence*), with substantial losses before 2100.
WGII	TechSumm	TechSumm.D	The options include vulnerability-reducing measures, avoidance (e.g., disincentivising developments in high-risk areas and addressing existing social vulnerabilities), hard and soft protection (e.g., sea walls, coastal wetlands), accommodation (e.g., elevating houses), advance (e.g., building up and out to sea) and staged, managed retreat (e.g., landward movement of people and development) interventions (*very high confidence*).
WGII	TechSumm	TechSumm.D	Nature-based interventions, for example wetlands and salt marshes, can reduce impacts and costs while supporting biodiversity and livelihoods but have limits under high warming levels and rapid sea level rise (*high confidence*).
WGII	TechSumm	TechSumm.D	Restoration of wetlands could support livelihoods and help sequester carbon (*medium confidence*), provided they are allowed accommodation space.
WGII	TechSumm	TechSumm.D	Flood-risk measures that work with nature by allowing flooding within coastal and wetland ecosystems and support sediment accretion can reduce costs and bring substantial co-benefits to ecosystems, liveability and livelihoods (*high confidence*).
WGII	12 Central and South America	12.ES	Disruption in water flows will significantly degrade ecosystems such as high-elevation wetlands and affect farming communities, public health and energy production (*high confidence*).
WGII	12 Central and South America	12.ES	Inclusive water regimes that overcome social inequalities and approaches including nature-based solutions, such as wetland restoration and water storage and infiltration infrastructure, with synergies for ecosystem conservation and disaster risk reduction, have been found to be more successful for adaptation and sustainable development (*high confidence*).
WGIII	SummPol	SummPol.D	D.2.1 Sustainable urban planning and infrastructure design including green roofs and facades, networks of parks and open spaces, management of urban forests and wetlands, urban agriculture, and water-sensitive design can deliver both mitigation and adaptation benefits in settlements (*medium confidence*).

WG	Chapter	Section	Text
WGII	3 Ocean and coastal ecosystems and their services	3.4	Overall, warming will drive range shifts in wetland species (medium to *high confidence*), but SLR poses the greatest risk for mangroves and salt marshes, with significant losses projected under all future scenarios by mid-century (*medium confidence*) and substantially greater losses by 2100 under all scenarios except SSP1-1.9 (*high confidence*).
WGII	3 Ocean and coastal ecosystems and their services	3.4	Under SSP5-8.5, wetlands are very likely at high risk from SLR, with larger impacts manifesting before 2040 (*medium confidence*).
WGII	3 Ocean and coastal ecosystems and their services	3.4	Otherwise, wetland ecosystems must migrate either inland or upstream, or face gradual submergence in deeper, increasingly saline water (*very high confidence*) (section 3.4.2.4; Andres et al., 2019; Jones et al., 2019b; Cohen et al., 2020; Mafi-Gholami et al., 2020; Magolan and Halls, 2020; Sklar et al., 2021).
WGII	3 Ocean and coastal ecosystems and their services	3.4	Nevertheless, previous declines have left wetland ecosystems more vulnerable to impacts from climate-induced drivers and non-climate drivers (*high confidence*) (Friess et al., 2019; Williamson and Guinder, 2021).
WGII	3 Ocean and coastal ecosystems and their services	3.4	Since AR5 and SRCCL, syntheses have emphasised that the vulnerability of rooted wetland ecosystems to climate-induced drivers is exacerbated by non-climate drivers (*high confidence*) (Elliott et al., 2019; Ostrowski et al., 2021; Williamson and Guinder, 2021) and climate variability (*high confidence*) (Day and Rybczyk, 2019; Kendrick et al., 2019; Shields et al., 2019).
WGII	3 Ocean and coastal ecosystems and their services	3.4	Without careful management of freshwater inputs, sediment augmentation and/or the restoration of shorelines to more natural states, transformation and loss of intertidal areas and wetland vegetation will increase with SLR (*high confidence*) (Doughty et al., 2019; Leuven et al., 2019; Yu et al., 2019; Raw et al., 2020; Shih, 2020; Stein et al., 2020), with small, shallow microtidal estuaries being more vulnerable to impacts than deeper estuaries with well-developed sediments (*medium confidence*) (Leuven et al., 2019; Williamson and Guinder, 2021).
WGII	4 Water	4.3	Many wetland-dependent species have seen a long-term decline, with the Living Planet Index showing that 81% of populations of freshwater species are in decline and others being threatened by extinction (Davidson and Finlayson, 2018; Darrah et al., 2019; Diaz et al., 2019) (*high confidence*).
WGII	4 Water	4.3	The loss and degradation of freshwater ecosystems have been widely documented, and SRCCL assessed with *medium confidence* the loss of wetlands since the 1970s (Olsson et al., 2020).

WG	Chapter	Section	Text
WGII	4 Water	4.5	SR1.5 concluded with *high confidence* that limiting global warming to 1.5°C, rather than 2°C, will strongly benefit terrestrial and wetland ecosystems and their services, including the cultural services provided by these ecosystems (Hoegh-Guldberg et al., 2018).
WGII	11 Australasia	11.3	Improved coastal modelling, experiments and in situ studies are reducing uncertainties at a local scale about the impact of future sea level rise (SLR) on coastal freshwater terrestrial wetlands (*medium confidence*) (Shoo et al., 2014; Bayliss et al., 2018; Grieger et al., 2019).
WGII	12 Central and South America	12.3	Drought has affected wetlands (*low confidence*) (Zhao et al., 2016; Domic et al., 2018) and desert ecosystems (*medium confidence*: medium evidence, high agreement) (Acosta-Jamett et al., 2016; Neilson et al., 2017; Díaz et al., 2019).
WGII	12 Central and South America	12.3	The projected impacts of climate change will lead to profound changes in the annual flood dynamics for Pantanal wetlands, altering ecosystem functioning and severely affecting biodiversity (*high confidence*) (Thielen et al., 2020; Marengo et al., 2021).
WGII	12 Central and South America	12.3	Disruptions in water flows will significantly degrade or eliminate high-elevation wetlands (*high confidence*) (Bury et al., 2013; Dangles et al., 2017; Mark et al., 2017; Polk et al., 2017; Cuesta et al., 2019).
WGII	13 Europe	13.3	Appropriately implemented ecosystem-based mitigation, such as reforestation with climate-resilient native species (section 13.3.1.4), peatland and wetland restoration, and agroecology (section 13.5.2), can enhance carbon sequestration or storage (*medium confidence*) (Seddon et al., 2020).
WGII	13 Europe	13.3	Trade-offs between ecosystem protection, their services and human adaptation and mitigation needs can generate challenges, such as loss of habitats, increased emissions from restored wetlands (Günther et al., 2020) and conflicts between carbon capture services, and provisioning of bioenergy, food, timber and water (*medium confidence*) (Lee et al., 2019; Krause et al., 2020).
WGII	13 Europe	13.3	Average wetland area is not projected to change at 1.7°C GWL across Europe, while for >4°C GWL expanding sites in NEU are not sufficient to balance losses in SEU and WCE (*high confidence*) (Xi et al., 2021).
WGII	13 Europe	13.4	While rising sea levels will also directly threaten intertidal and beach ecosystems, coastal wetlands will benefit (*medium confidence*), in case lateral accommodation space and the opportunity for systems to migrate landward and upwards is provided, enhancing their ability to capture and store carbon (Lecocq et al., 2022; Rogers et al., 2019).
WGII	13 Europe	13.10	Ecosystem-based solutions, such as wetlands, can reduce waves' propagation, provide co-benefits for the environment and climate mitigation, and reduce costs for flood defences (*medium confidence*) (section 13.2.2.1).
WGII	13 Europe	13.10	Around 2°C GWL, losses accelerate in marine ecosystem and appear across systems, including habitat losses especially in coastal wetlands (Roebeling et al., 2013; Clark et al., 2020), biodiversity and biomass losses (Bryndum-Buchholz et al., 2019; Lotze et al., 2019) and ecosystem services such as fishing (*high confidence* on the direction of change, but *medium confidence* on the local and regional magnitude) (Raybaud et al., 2017).

WG	Chapter	Section	Text
WGII	14 North America	14.5	Other adaptation responses to reduce temperature effects include modifying structures (roofs, engineered materials) and the urban landscape through green infrastructure (e.g., urban trees, wetlands, green roofs), which increases climate resilience and quality of life by reducing urban heat island effects, while additionally improving air quality, capturing stormwater and delivering other co-benefits to the community (e.g., access to food, connection to nature, social connectivity) (*high confidence*) (see box 14.7; Ballinas and Barradas, 2016; Emilsson and Sang, 2017; Kabisch et al., 2017; Krayenhoff et al., 2018; Petrovic et al., 2019; Schell et al., 2020).
WGII	14 North America	14.5	These environmental conditions also stress natural assets (e.g., urban forests, wetlands, household gardens, green walls) and performance of green infrastructure leading to higher operation and maintenance costs (*high confidence*) (Kabisch et al., 2017; Terton, 2017).
WGII	15 Small Islands	15.3	SLR has been projected to impact the terrestrial biodiversity of low-lying islands and coastal regions via large habitat losses both directly (e.g., submergence) and indirectly (e.g., salinity intrusion, salinisation of coastal wetlands and soil erosion) at even the 1-m scenario (medium to *high confidence*).
WGIII	7 Agriculture, Forestry and Other Land Uses (AFOLU)	7.4	There is *medium confidence* that coastal wetland protection has a technical potential of 0.8 (0.06–5.4) gtco2-eq yr –1 of which 0.17 (0.06–0.27) gtco2-eq yr –1 is available up to usd100 tco2–1.
WGIII	7 Agriculture, Forestry and Other Land Uses (AFOLU)	7.4	There is *high confidence* that coastal wetlands, especially mangroves, contain large carbon stocks relative to other ecosystems and *medium confidence* that restoration will reinstate pre-disturbance carbon sequestration rates.
WGIII	7 Agriculture, Forestry and Other Land Uses (AFOLU)	7.4	There is *low confidence* on the response of coastal wetlands to climate change; however, there is *high confidence* that coastal wetland restoration will provide a suite of valuable co-benefits.
WGIII	7 Agriculture, Forestry and Other Land Uses (AFOLU)	7.4	There is *medium confidence* that coastal wetland restoration has a technical potential of 0.3 (0.04–0.84) gtco2-eq yr –1 of which 0.1 (0.05–0.2) gtco2-eq yr –1 is available up to usd100 tco2–1.

Table 8: Statements that contain the keyword term 'wetland'. The key terms in the statement are highlighted, colors represent the categories they belong to.