Systematic documentation of CIMMYT’s work in the food production-climate change nexus

Report June 2021

Bia Carneiro (Principal investigator)
Giuliano Resce
Giosue Ruscica
Daniel Mazer
Summary

List of Tables and Figures 3

Executive Summary 5

1. Introduction 8
   1.1 Digital methods: an approach to assess discourse and diffusion of information 11
   1.1.1 Analytical Framework 11

2. Data and methods 13
   2.1 Data Collection 14
      2.1.1 CIMMYT Website 14
      2.1.2 CIMMYT Repository 15
      2.1.3 Scopus database 15
      2.1.4 Twitter 16
   2.2 Assessing institutional, scientific and public discourses through text mining 18
   2.3 Mapping information dissemination through hyperlink analysis 20
   2.4 Exploring scientific cooperation and institutional dialogues 22

3. Main results 23
   3.1 CIMMYT’s major work areas within the food production-climate change nexus 24
      3.1.1 Evidence from strategic institutional communications 32
      3.1.2 Evidence from CIMMYT’s knowledge products 34
      3.1.3 Evidence from scientific production 50
      3.1.4 Evidence from social media 51
      3.1.5 Summary 61
   3.2 Dissemination of CIMMYT’s climate-focused knowledge products 62
   3.3 How and with whom is CIMMYT engaging in the network of climate action? 67
      3.3.1 Network of climate action 67
      3.3.2 Network of scientific cooperation 71

4. Conclusions and recommendations 77

References 78

Appendix 1 – Repository Descriptives 80

Appendix 2 – Scopus descriptives 82
### List of Tables and Figures

<table>
<thead>
<tr>
<th>Table/Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Taxonomy (terms collected manually and matched to AGROVOC)</td>
</tr>
<tr>
<td>Table 2</td>
<td>Examples of knowledge products considered for scraping</td>
</tr>
<tr>
<td>Figure 1</td>
<td>Analytical framework for digital research to generate insights about strategic topics</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Description of data sources</td>
</tr>
<tr>
<td>Figure 3</td>
<td>A typical tweet</td>
</tr>
<tr>
<td>Figure 4</td>
<td>Word cloud for tags identified in posts from CIMMYT’s climate change news webpage</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Distribution of topics identified in the “Climate Change” page of CIMMYT’s website. More frequent terms represented by wider wedges in the pie chart.</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Timelines for the prevalence of climate-focused topics identified in the “Climate Change” page of CIMMYT’s website. Topics normalized on scale from 0.0-1.0</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Correlations between ‘Climate change’ and cross-cutting topics identified in the “Climate Change” page of CIMMYT’s website.</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Timelines for the prevalence of climate-focused topics identified in the “Climate Change” page of CIMMYT’s website. Topics normalized on scale from 0.0-1.0</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Correlations between cross-cutting topics and climate topics identified in “Climate Change” page of CIMMYT’s website.</td>
</tr>
<tr>
<td>Figure 10</td>
<td>Correlations between climate topics identified in “Climate Change” page of CIMMYT’s website.</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Distribution of countries identified in posts from CIMMYT’s climate change news webpage.</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Distribution of countries identified in climate-related knowledge products available in the CIMMYT Repository.</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Distribution of topics identified in climate-related knowledge products available in the CIMMYT Repository. More frequent terms represented by wider wedges in the pie chart.</td>
</tr>
<tr>
<td>Figure 14</td>
<td>Timelines for the prevalence of cross-cutting topics identified in climate-related knowledge products available in the CIMMYT Repository. Topics normalized on scale from 0.0-1.0</td>
</tr>
<tr>
<td>Figure 15</td>
<td>Correlations between ‘Climate change’ and cross-cutting topics identified in climate-related knowledge products available in the CIMMYT Repository.</td>
</tr>
<tr>
<td>Figure 16</td>
<td>Timelines for the prevalence of climate topics identified in climate-related knowledge products available in the CIMMYT Repository. Topics normalized on scale from 0.0-1.0</td>
</tr>
<tr>
<td>Figure 17</td>
<td>Correlations between climate topics and cross-cutting topics identified in climate-related knowledge products available in the CIMMYT Repository.</td>
</tr>
<tr>
<td>Figure 18</td>
<td>Correlations between climate topics identified in climate-related knowledge products available in the CIMMYT Repository.</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Word cloud from abstracts of CIMMYT-affiliated climate-focused publications indexed by Scopus.</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Figure 25 Distribution of countries mentioned in abstracts of CIMMYT-affiliated climate-focused publications indexed by Scopus.</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Distribution of topics identified in CIMMYT-affiliated climate-focused publications indexed by Scopus. More frequent terms represented by wider wedges in the pie chart.</td>
</tr>
<tr>
<td>Figure 22</td>
<td>Figure 29 Timelines for the prevalence of cross-cutting topics identified in CIMMYT-affiliated publications indexed by Scopus (topics normalized on scale from 0.0-1.0).</td>
</tr>
<tr>
<td>Figure 23</td>
<td>Correlations between ‘Climate change’ and cross-cutting topics identified in CIMMYT-affiliated publications indexed by Scopus.</td>
</tr>
<tr>
<td>Figure 24</td>
<td>Timelines for the prevalence of climate topics identified in CIMMYT-affiliated publications indexed by Scopus. Topics normalized on scale from 0.0-1.0.</td>
</tr>
<tr>
<td>Figure 25</td>
<td>Correlations between climate topics and cross-cutting topics identified in CIMMYT-affiliated publications indexed by Scopus.</td>
</tr>
<tr>
<td>Figure 26</td>
<td>Correlations between climate topics identified in CIMMYT-affiliated publications indexed by Scopus.</td>
</tr>
<tr>
<td>Figure 27</td>
<td>Distribution of countries identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 28</td>
<td>Word cloud of hashtags identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 29</td>
<td>Figure 36 Distribution of themes identified in Tweets containing either @CIMMYT or #CIMMYT. More frequent terms represented by wider wedges in the pie chart.</td>
</tr>
<tr>
<td>Figure 30</td>
<td>Timelines for the prevalence of cross-cutting topics identified in Tweets containing either @CIMMYT or #CIMMYT. Topics normalized on scale from 0.0-1.0.</td>
</tr>
<tr>
<td>Figure 31</td>
<td>Correlations between ‘Climate change’ and cross-cutting topics identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 32</td>
<td>Timelines for the prevalence of climate topics identified in Tweets containing either @CIMMYT or #CIMMYT. Topics normalized on scale from 0.0-1.0.</td>
</tr>
<tr>
<td>Figure 33</td>
<td>Correlations between climate topics and cross-cutting topics identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 34</td>
<td>Correlations between climate topics identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 35</td>
<td>Figure 40 Frequency of URLs by publication type.</td>
</tr>
<tr>
<td>Figure 36</td>
<td>Frequency of top 25 domain names (CIMMYT website and repository excluded).</td>
</tr>
<tr>
<td>Figure 37</td>
<td>Figure 45 Network visualization of Top Level Domain distribution.</td>
</tr>
<tr>
<td>Figure 38</td>
<td>Figure 44 Top 75 influential nodes in Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 39</td>
<td>Figure 43 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 40</td>
<td>Figure 42 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class.Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 41</td>
<td>Figure 41 Word cloud of account mentions identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 42</td>
<td>Figure 40 Geographical distribution based on country-code TLD.</td>
</tr>
<tr>
<td>Figure 43</td>
<td>Figure 43 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 44</td>
<td>Figure 42 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 45</td>
<td>Figure 41 Word cloud of account mentions identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 46</td>
<td>Figure 40 Geographical distribution based on country-code TLD.</td>
</tr>
<tr>
<td>Figure 47</td>
<td>Figure 43 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 48</td>
<td>Figure 42 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 49</td>
<td>Figure 41 Word cloud of account mentions identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 50</td>
<td>Figure 40 Geographical distribution based on country-code TLD.</td>
</tr>
<tr>
<td>Figure 51</td>
<td>Figure 43 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 52</td>
<td>Figure 42 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
<tr>
<td>Figure 53</td>
<td>Figure 41 Word cloud of account mentions identified in Tweets containing either @CIMMYT or #CIMMYT.</td>
</tr>
<tr>
<td>Figure 54</td>
<td>Figure 40 Geographical distribution based on country-code TLD.</td>
</tr>
<tr>
<td>Figure 55</td>
<td>Figure 43 Overall Twitter mentions network. Parameters: force-directed graph, with node size partitioned as Weighted in-Degree, colored by modularity class. Nodes restricted to those mentioned at least 5 times. Nodes = 2,797 accounts, edges = 7,976 connections.</td>
</tr>
</tbody>
</table>
This report is the output of the external study “Systematic documentation of CIMMYT’s work in the food production-climate change nexus”, carried out by a team of consultants led by Bia Carneiro and coordinated by Tek Sapkota, Agricultural Systems and Climate Change Scientist at CIMMYT. We would like to thank CIMMYT staff for their support, namely Bram Govaerts, Bruno Gérard, Victor Kommerell and Michelle Guertin for the constructive feedback on the analyses, and Marcia MacNeil for providing us the dataset from the CIMMYT website and for communications support. Finally, we are very grateful to Tek Sapkota for his coordination.
Executive Summary

The past two decades have witnessed an evolution in comprehending the nexus between climate change and food production. As debates around the theme have shifted from initial awareness raising to actual programmatic interventions, international research organizations have made significant contributions by generating knowledge and engaging stakeholders.

Nevertheless, while the value of agriculture-focused climate science is widely recognized, for organizations that develop and manage large-scale research programs, time and resource constrains often prevent the systematic analysis of outputs, and the dynamics of climate-sensitive knowledge dissemination are not fully understood.

As metrics traditionally employed to evaluate scientific outputs present limitations to synthesize accumulated institutional documentation and to retrofit indicators to historical outputs, this study proposes an integrated, data-driven approach to systematically analyze the production and diffusion of agriculture-focused climate science within a large-scale agricultural research institution, using the International Maize and Wheat Improvement Center (CIMMYT) as a case study.

This review aimed to locate the institution’s climate change-food production nexus research portfolio within an integrated programmatic framework, seeking to uncover not only what CIMMYT has been doing in relation to the issue, but also how climate sensitivity is articulated with other strategic themes.

By bringing together a Digital Methods perspective and machine learning techniques, we systematically analyzed CIMMYT’s climate research portfolio at various levels and assessed CIMMYT’s engagement with the broader network of climate action. Based on existing literature, the proposed analytical framework aimed to assess the discourse and the knowledge dissemination processes of CIMMYT’s climate research outputs by considering online narratives and relationships as evidence of broader strategic engagement. We propose a mixed methods approach that employs text mining, network analysis and hyperlink analysis to an unstructured mass of existing, publicly available data (research outputs, online content and social media) to uncover narratives and relationships.

Data was extracted from several sources: CIMMYT’s Climate change news page on its website; CIMMYT’s Repository; an index of peer reviewed articles by CIMMYT-affiliated authors, and Twitter content mentioning CIMMYT. A taxonomy was developed in consultation with CIMMYT experts, and custom algorithms were developed to process and analyze the data.
The geographical focus of CIMMYT’s climate research centered on three main countries: Mexico, India, and Ethiopia. Farming systems, Food security, Technology transfer, and Innovation were the most frequent cross-cutting topics across the four data sources.

Profitability, Productivity, Production, Resource management, Adaptation and Mitigation were the most frequent climate-focused topics in all content assessed.

Farming systems and Food security were the cross-cutting topics with the most significant and steady rise in prevalence over time.

The climate-focused topics that have emerged in the last five years across the datasets are Carbon sequestration, Emission reduction, Energy conservation and Sustainable intensification.

Since climate change mitigation and adaptation processes are often intertwined, there was a strong association between the two in all datasets.

Likewise, as an increase in profitability is an expected outcome of higher productivity, the two concepts co-occurred within the same posts or publications in all datasets.

While institutional communications on CIMMYT’s website and on social media have addressed the effects of climate change on the livelihoods of rural people, this connection was not reflected in the institution’s scientific and non-scientific knowledge products, which focused on research around the development of agricultural technologies and practices to support food production and cope with climate impacts.

There was variability among the correlations between cross-cutting topics and climate-focused topics in the different datasets. While CIMMYT’s website most frequently connected climate change to adaptation and mitigation measures, the strongest co-occurrences in publications from the Repository bring together capacity building with risk management and production, as well as climate change with productivity and profitability. Among indexed, peer reviewed publications, besides the interlinkages between climate change and mitigation and adaptation, the impacts of new approaches such as CSA and precision agriculture on socioeconomic development were also addressed. Lastly, Twitter conversations made the connection between stakeholder participation (specially women) in sustainable practices and increased tolerance to climate stresses.

Among climate-focused topics, besides the common trend noted above, CSA was associated to precision agriculture and nutrient management in scientific and non-scientific publications, as well as on social media conversations.

On Twitter, the majority of cross-cutting topics assessed were already well established within the agricultural development community when the platform was launched in 2009, which is reflected in their presence throughout the period of analysis. In fact, many topics, such as capacity building, food security, resource management and participation had their peak prevalence in the beginning of last decade. This is also the case of ‘Farming systems’, although its prevalence has been steadily on the rise for the last five years, indicting heightened awareness towards integrated approaches for sustainable agriculture.

The hyperlink analysis shows that knowledge generated through CIMMYT’s research has been shared across thousands of websites: scientific research such as articles and books were the most frequently disseminated outputs, followed by outreach materials such as handbooks, reports and presentations.

CIMMYT’s outputs have been distributed to more than 150 countries across both the Global North and South. Domains in North America, Australia, and Europe have frequently disseminated CIMMYT research, but also in South and Southeast Asia, and Latin America.

In assessing the dynamics between tweets by CIMMYT and tweets that mentioned the organization either through @CIMMYT or #CIMMYT, the complete network of mentions derived from this climate-focused tweets contained more than 2,700 unique accounts, connected almost 8,000 times.
When connections are restricted to accounts mentioned at least five times, the number of nodes reduces to 704, with 4,755 connections between them, thus making the network denser, with nodes connected on average to 6.8 other nodes. Four major clusters are identified.

A centrality metric applied to establish the most influential nodes in the network found that CIMMYT and CCAFS (@CGIARclimate) are prominent actors within the climate community. While it must be noted that the network is based on CIMMYT mentions, results show that CIMMYT actively engages with key players within international development and research, supporting the dissemination of climate knowledge to a broad range of actors.

**Recommendations:**

- Strengthen CIMMYT’s science-policy interface
- ‘Translate’ scientific outputs to more accessible products
- Increase the visibility of emerging climate-sensitive approaches and technologies
- Expand the application of Digital Methods and Machine learning to support the systematization of CIMMYT’s research portfolios.

Web-based research offers innovative approaches to generate knowledge about strategic issues that are difficult to observe using traditional methods.
1. Introduction
Several parts of the world increasingly experience the effects of climate change through more frequent extreme weather events, higher average temperatures and increased variability. Agricultural systems are particularly sensitive to these effects due to their dependence on stable, long-term climatic conditions that impact productive capacity, quality and yields (Cradock-Henry et al. 2020). In particular, farmers struggle to cope with increasing climate risks and, consequently, impacts on food production make it harder to meet rising global food demand (CIMMYT, 2019).

The past two decades have witnessed an evolution in comprehending the nexus between climate change and food production. As debates around the theme have shifted from initial awareness raising to actual programmatic interventions, international research organizations have made significant contributions by generating knowledge and engaging stakeholders (Runhaar et al. 2018).

Nevertheless, while the value of agriculture-focused climate science is widely recognized, for organizations that develop and manage large-scale research programs, time and resource constraints often prevent the systematic analysis of outputs, and the dynamics of climate-sensitive knowledge dissemination are not fully understood. In fact, a clear framework that measures progress while accounting for the complexities of research-for-development remains a major challenge, as metrics traditionally employed to evaluate scientific outputs present limitations. Two key difficulties are: synthesizing the accumulated information spread not only across thousands of internal and external documents (Bornmann and Mutz, 2015), but also across various communications channels and platforms; and retrofitting indicators to historical research, as concepts and terminologies tend to evolve over time (Garbero et al, 2021).

The integration of digital research methodologies (Rogers, 2013) with big data analytics and machine learning techniques offers an innovative and comprehensive approach to deal with such challenges. Recent studies (di Bella et al. 2016; Einav and Levin, 2014; Blazquez and Domenech 2018, Carneiro et al, 2020), show the potential for leveraging on underexplored sources of information produced within organizations, such as project documentation, institutional communications or research outputs, as well as on digital artifacts available across the web, to generate knowledge about strategic issues that may be difficult to observe using traditional surveys or manual stock takes.

Hence, this study proposes an integrated, data-driven approach to systematically analyze the production and diffusion of agriculture-focused climate science within a large-scale agricultural research institution, using the International Maize and Wheat Improvement Center (CIMMYT) as a case study. CIMMYT has been conducting extensive research on climate change adaptation and mitigation in agriculture, particularly in maize and wheat-based production systems across the regions1. CIMMYT’s climate focused research aims to help farmers adapt to shocks while producing more food and reducing emissions, where possible.

---

1 Funded by bilateral donors and since 2012 also by CGIAR Research Programs MAIZE, WHEAT and CCAFS.
Innovations in maize and wheat-based systems include new maize and wheat varieties; conservation agriculture; farming methods that increase nutrient use efficiency and thus lower production inputs such as water, fertilizer, chemicals and farm energy; climate information services; and index-based insurance for farmers whose crops are damaged by bad weather. CIMMYT implemented this research-for-development together with partners in national agricultural research systems (NARS) and in many cases with other CGIAR Research Centers.

In particular, some of CIMMYT’s innovations in climate change adaptation and mitigation research include:

- **Sustainable intensification**, which, at the very least, decreases the emissions intensity of food, but may also help “spare land” to avoid conversion of high carbon ecosystems such as forests and grasslandes and supports adaptation to extreme climate events.

- **Advances in agronomic practices**, including conservation agriculture, which increase resources use efficiency (water, nutrient, energy, and labor, resources, recycling), help reduce emissions that would have occurred with less efficient systems in the future, and support adapting to weather variabilities and coping with extreme climatic events.

- **Resilient breeding and seed systems** that increase resource use efficiency and adapt to weather stresses.

CIMMYT’s outreach efforts have also contributed to scale-out several climate-smart technologies and practices in the climatically challenged locations and production systems in Asia, Africa and Latin America regions. Strengthening capacity on adaptation and mitigation research for national researchers, strategic and farmers’ field trials, and engagement and communication with national and sub-national stakeholders helped to reach a large number of farmers. For these endeavors, CIMMYT has engaged with many partner institutions at global, regional, and national/sub-national levels including government bodies, universities and research centers, and think-tank organizations.

The strong climate focus of CIMMYT’s research, particularly in the last two decades, has generated several research outputs relevant to adaptation and mitigation actions in agriculture. However, due to the issues discussed above, the institution lacked a consistent framework from which to monitor progress against its strategic climate agenda.

By bringing together a Digital Methods perspective and machine learning techniques, we systematically analyzed CIMMYT’s climate research portfolio at various levels and assessed CIMMYT’s engagement with the broader network of climate action. The web analytics approach sought to identify and classify the research outputs generated by CIMMYT, explore the process of knowledge dissemination for these outputs, and identify opportunities for future works.

The following sections describe the methodology and present the main findings from the systematic documentation of the center’s research for development work in the food production-climate change nexus.

---

2 https://www.cimmyt.org/news/climate-adaptation-mitigation/
1.1. Digital methods: an approach to assess discourse and diffusion of information

The pervasiveness of the internet in people’s lives is undeniable. According to recent data\(^3\), almost 60 percent of the world’s population has access to the internet. Last year alone, an additional three hundred million people went online for the first time, most of them in developing countries. Moreover, just over half of the world’s population is active on social media, which now counts on 4.2 billion users. On average, people spend close to seven hours per day on the internet, of which 2.5 hours are spent on social media.

As the most prominent part of the internet, the World Wide Web plays a crucial role within the technological infrastructure of society, as the evolution of the web has engendered new social practices and new forms of knowledge exchange (Fuchs et al., 2010; Song, 2010). This continuous transformation of information technologies in the digital era have expanded the reach of communications tools to all aspects of social life through networks that are at the same time global and local, and increasingly integrated into existing offline practices and social relationships (Wellman and Haythornthwaite, 2002).

Such hybridization of online and offline dimensions has effectively established the internet as a space for research on social phenomena. The emergence of the internet as a field of research has engendered several approaches that propose moving beyond assessing how much of society is online and towards investigating cultural and social transformations through the internet (Rogers, 2013). One of them is ‘Digital Methods’ (Rogers 2013, 2015), which proposes to study social and cultural change through online data, with the web as the data set and the development of natively digital techniques for analysis and interpretation.

1.1.1. Analytical Framework

It has been extensively documented in academic literature that web and social media activities can be considered proxies for wider public discourse and engagement (Carneiro et al, 2020; Carneiro and Costa, 2020; Lotan et al 2011; Pearce et al. 2019; Resce and Maynard, 2019; Rogers and Marres 2000; Schäfer 2012; Niekler and Wencker 2019, among many others). Based on this notion, and on the digital context discussed above, the analytical framework of this study

\(^3\) https://datareportal.com/reports/digital-2021-global-overview-report
aimed to assess the discourse and the knowledge dissemination processes of CIMMYT’s climate research outputs by considering online narratives and relationships as evidence of broader strategic engagement.

**Figure 1** presents the analytical framework for the application of digital research to generate insights about strategic topics. We propose a data-driven, mixed methods approach that employs text mining, network analysis and hyperlink analysis to an unstructured mass of existing, publicly available data (research outputs, online content and social media) to uncover narratives and relationships.

This framework was used to investigate CIMMYT’s work in the food production-climate change nexus. Specifically, we aimed to address the following research questions:

1. **What** has been CIMMYT’s research for development focus within the food production-climate change nexus?
2. **How** have CIMMYT’s climate-related knowledge products been disseminated across different communities (scientific, policy, practitioners, etc.) and geographies?
3. **How and with whom** is CIMMYT engaging in the broader network of climate change research and action?
2. Data and methods
Framed within the Digital Methods epistemology discussed above, this study leveraged on publicly available digital traces (Hepp et al, 2018) as comprehensive data sources. The data-driven approach followed the typical machine learning process, where data was gathered, processed, and machine learning models were applied to gather insights (Porciello et al, 2020). The research was computationally intensive and entailed the production of several datasets from which the analysis was performed. Figure 2 summarizes our key data sources, and the data extraction process is detailed below.

2.1 Data Collection

2.1.1 CIMMYT Website

This dataset comprises the text and metadata for 444 news items published between 2006-March 2021 under the “Climate Change” research theme, which provide a first representation of how the institution has framed its climate-focused research from a strategic perspective. It was the starting point for the development of the taxonomy.
2.1.2 CIMMYT Repository

The CIMMYT Repository was selected as a key source for CIMMYT’s knowledge products, after review of other potential data repositories such as the CGIAR Gardian platform and the website of the CGIAR Research Program on Climate Change, Agriculture and Food security (CCAFS). Upon testing a few combinations, the repository was queried for the keywords “climate” and “clima”, which resulted in 2,463 publications, including both scientific and communications products, disseminated between 1960 and 2021. A custom algorithm was developed to scrape the results and extract the available metadata and data was collected on 12 April 2021. The following information was extracted for each record:

- Publication title
- Publication type
- Authors
- Abstract
- Year
- URL
- Publisher

2.1.3 Scopus database

Bibliometric analysis comprises quantitative methods to analyze books, articles and other publications. While traditionally, this method has been used to assess authorship, impact and reach of academic outputs, recent research has expanded the field to include systematic, machine-driven reviews of subject areas and keywords (Aristovnik et al 2020). Though based on bibliometric analysis principles, our study does not include analysis of publication impact metrics.

Rather, we focused on uncovering relevant subjects through text mining techniques. Based on the Scopus database – one of the most comprehensive abstract and citation databases of peer-reviewed literature available – a query was performed to identify literature in which authors declared affiliation to CIMMYT. From an initial (AF-ID("CIMMYT")) query, the possible institutional affiliations indexed in Scopus were identified and the following complete query was performed:

(AF-ID("Centro Internacional de Mejoramiento de Maiz y Trigo" 60015079) OR AF-ID("International Maize and Wheat Improvement Center Harare" 60051398) OR AF-ID("International Maize and Wheat Improvement Center CIMMYT" 60106318) OR AF-ID("International Maize and Wheat Improvement Center Nepal" 60071799) OR AF-ID("International Maize and Wheat Improvement Center" 60120614))

Data was collected on 25 March 2021. The query resulted in 3612 items published between 1974-2021. Full citation data and metadata were exported, including:

- Publication title
- Journal title, volume, issue
- Authors
- Affiliations
- Keywords (author and indexed)
- Funding details
- Abstract
- Year
- DOI and URL
- References
2.1.4 Twitter

Recent academic research has relied extensively on social media platforms as proxies for wider public discourse and engagement, with the application of content analysis techniques to identify trends in political agendas, over time and across geographies (Carneiro et al., 2020). But while social media is often discussed in general terms, it is not a homogeneous entity. In fact, social media contains many different “platform cultures” arising from a combination of technical affordances and user behaviors.

Specifically, Twitter has become an important venue for institutional communications; news media increasingly rely on the platform as a primary source for official statements and position-taking. The platform has been described as a digital forum, and in the literature that has assessed it in connection to online climate dialogues and action, it is considered an important “source for climate change information-exchanges” (Pearce et al. 2019).

The analysis focused on Twitter’s potential as a real-time, topic-driven platform that enables rapid detection of trends, in order to explore discourse dynamics (McDonald, 2013), as narratives around climate are considered important elements of “shifting rationalities” (Käkönen et al, 2014), as well as engagement in dialogue with various stakeholders. To uncover the conversations CIMMYT has participated in, Twitter was scraped for all tweets that either mentioned the institution’s handle (@cimmyt) or that contained the hashtag #CIMMYT.

Yet, while Twitter is still the most searchable social media platform and is used widely for academic research, accessing data through its API has two important limitations: rate limit on requests and a temporal restriction on only gathering back data from eight days in the past. As such, a custom program was developed in Python language to enable extracting all the tweets from 2009 to 2021, and, by means of anti-block algorithms, in a reasonable amount of time.

Data was collected on 13 April 2021. In total, 69,027 tweets were gathered in the period between 2009-02-06 to 2021-04-12. Figure 3 shows a typical tweet with all its affordances. For every tweet, the following data was taken:

- Time of publication
- Text of tweet
- Hashtags
- Mentions
- Number of favorites
- Number of replies
- Number of retweets
- URL of tweet

The following sections describe the methods and processes to address each research question in detail.

---

An API (Application Programming Interface) is a programming interface between otherwise separate applications that allows each application to connect, communicate and share information with the other.
2.2. Assessing institutional, scientific and public discourses through text mining

Within the social sciences, content analysis methods allow researchers to identify patterns and changes in political agenda over time and across geographies (Brandt, 2019). Among digitally based content analysis approaches, Text Mining is broadly defined as an Artificial Intelligence (AI) technique that uses Natural Language Processing (NLP) to transform unstructured text of documents/databases such as web pages, newspaper articles, e-mails, press, posts/comments on social media, in structured and normalized data (Resce and Maynard, 2018). Words, the carriers of meaning, are identified and transformed into a processable data structure. To enable the repurposing of existing, unstructured data, and to efficiently extract meaningful information from large datasets, “the combination of interpretative appraisal and statistical techniques has the potential to generate novel insights, ultimately contributing to evidence-based policy-making” (Niederle and Wencker 2019:3).

Taking guidance from CIMMYT experts, we constructed a framework comprising two overarching themes.

Text mining was employed to derive insights into what CIMMYT’s research outputs say about climate, and to classify and quantify the relationship between this and other strategic topics in the data. A crucial step for the analysis was the development of a custom taxonomy to identify key terminology against which we could map text from the data sources. Taking guidance from CIMMYT experts, we constructed a framework comprising two overarching themes. First, cross-cutting topics cover the strategic themes mainstreamed into CIMMYT’s research as part of its integrated approach to food systems; and secondly, climate-focused topics were identified to reflect specific techniques and technologies researched. The framework recognizes the interconnection between climate and other research agendas within the organization and aims to unpack those relationships as represented in textual data.

Within each theme, the terminology used by CIMMYT was matched to AGROVOC⁵, the Food and Agriculture Organization’s (FAO) comprehensive, open-source, multilingual vocabulary. AGROVOC consists of more than 37,000 concepts covering FAO’s areas of interest, such as food, nutrition, agriculture, fisheries, forestry, environment etc. Vocabulary is available in up to 37 languages.

⁵ https://agrovoc.fao.org/
### Table 1: Taxonomy (terms collected manually and matched to AGROVOC)

#### CROSS-CUTTING TOPICS
- Capacity development
- Climate change
- Diversification
- Business development
- System approach
- Food security
- Gender
- Health
- Innovation
- Livelihood
- Legume integration
- Soil-borne diseases (nematodes)
- Nitrification
- Nutrition
- Social inclusion
- Policy influence
- Service delivery
- Socioeconomic development
- Abiotic/Biotic stress
- Technology
- Scaling
- Yield

#### CLIMATE-FOCUSED TOPICS
- Agroforestry
- Carbon sequestration
- Adaptation
- Mitigation
- Climate smart agriculture (CSA)
- Conservation agriculture
- Crop residue management
- Greenhouse gas emission reduction
- Energy efficiency
- Mulching/soil cover
- Nutrient management
- Nutrient use efficiency
- Precision agriculture
- Increased production
- Increased productivity
- Economic profitability
- Resistance breeding
- Resource management
- Climate risk mitigation
- Sustainable intensification
- Tolerant varieties
- Water management
- Zero tillage/no tillage

#### AGROVOC MATCH
- Capacity development
- Climate change
- Diversification
- Enterprises
- Farming systems
- Food security
- Gender equity
- Health
- Innovation
- Livelihoods
- Mixed cropping
- Nematoda
- Nitrification
- Nutrition
- Participation
- Policies
- Services
- Socioeconomic development
- Stress
- Technology
- Technology transfer
- Yields
Table 1 shows the taxonomy. For each topic, the corresponding AGROVOC definition was extracted in a JSON file, and a custom algorithm was developed to detect and classify the related terminology within the text of the various data sources. For each document \( j \) (abstracts, tweets, etc), we quantified the presence of a term \( i \) defined in AGROVOC as follows:

\[
\frac{\sum_{j} (\text{Words}_{\text{Document}_j} \in \text{Words}_{\text{AGROVOC}})}{\sum \text{Words}_{\text{Document}_j}}
\]

Max

\[
\frac{\sum_i (\text{Words}_{\text{Document}_j} \in \text{Words}_{\text{AGROVOC}})}{\sum \text{Words}_{\text{Document}_j}}
\]

In addition, a country detection algorithm was developed to identify countries mentioned in the text.
2.3. Mapping information dissemination through hyperlink analysis

An approximation of how CIMMYT’s climate-related knowledge products have been disseminated globally was explored through a hyperlink analysis, which used the list of publications extracted from CIMMYT’s repository to explore how they are spread through the web. This led to the identification, by means of web scraping algorithms, of the web pages hyperlinking to every item of the considered list.

The dataset can be thought of as sets of pairs composed of a CIMMYT knowledge product and a web page, either hyperlinking or mentioning that product. The starting point for the creation of this dataset was the list of almost 2,500 climate-related knowledge products identified in CIMMYT’s repository (of 5,400 in total), and which contained the following information:

- **Repository URL**: the URL for the repository publication
- **Publication title**: the title or reference of the publication

<table>
<thead>
<tr>
<th>Repository URL</th>
<th>Publication title</th>
</tr>
</thead>
<tbody>
<tr>
<td>repository.cimmyt.org/handle/10883/20129</td>
<td>Out scaling climate-smart technologies to smallholder farmers in Malawi, Zambia &amp; Zimbabwe: vulnerability assessment report</td>
</tr>
<tr>
<td>repository.cimmyt.org/handle/10883/19273</td>
<td>Prioritizing climate-smart agricultural land use options at a regional scale</td>
</tr>
</tbody>
</table>

Table 2 Examples of knowledge products considered for scraping

Table 2 shows some examples of deliverables used for the dataset. For every unique publication, a search for all the web pages hyperlinking to it was carried out. The task of finding hyperlinks through the whole web is very ambitious; in fact, dedicated projects exist with the goal of building up updated and comprehensive hyperlink graphs, such as the Web Data Commons project.
In this study, results are promising, but a key challenge to consider is the coverage of the entire web. The most complete collection of web pages is provided by private companies such as Google, Yahoo, and Microsoft, which monetize their services. Public web archives exist (e.g., WayBackMachine or CommonCrawl), and they are suitable for many research projects but might miss a good number of web pages, resulting in limited results (see for example Stolz and Hepp, 2015). A potential solution could be the development of a dedicated search engine, though this is something that takes time to be written, especially when dealing with billions of records and petabytes of data. Ordinary technology cannot be used, requiring Big Data infrastructures instead.

As a good balance between coverage and resource limitations, commercial search engines have been used for this task, and in particular Google, Bing and DuckDuckGo. The advantages of this strategy are the great web coverage of the chosen platforms, and the ability to use their existing search engines. Yet, some limitations include that they do not allow searching directly for hyperlinks, and they have anti-scraping policies that need to be overcome.

The algorithm developed for this task searches for the presence of all words in the title of the research output and a string of the disseminated URL (not a link). For instance, using the first example in table 2, the corresponding query would be:

[out scaling climatesmart technologies to smallholder farmers in Malawi zambia zimbabwe vulnerability assessment report] OR [allafrica.com/stories/201712140408.html]

This approach works very well when publications have specific titles, such as the example above, but broad titles can lead to ambiguous results. To tackle this, as a part of the scraping program, an algorithm was developed to spot such ambiguous results. Due to strict anti-scraping policies of some search engines, the use of automated browsers such as Selenium was necessary, resulting in an elapsing time of almost two days on a single machine. Every extracted record is composed of the following fields:

- Title of publication
- URL of web page containing a reference to the publication
- Domain of the found URL (e.g., repository.cimmyt.org from https://repository.cimmyt.org/handle/10883/20129)
- Top Level Domain (TLD) of the found URL (e.g. .org or .com)
- Metadata: publication type, year, authors and publisher
2.4. Exploring scientific cooperation and institutional dialogues

Using the data collected from Twitter, it was possible to assess CIMMYT’s place within a broader network of actors engaged in the food production-climate nexus by analyzing the relationships between accounts mentioned on tweets about climate through a network analysis. This approach enables the visualization of relational data organized as matrices, where entities are the nodes – in this case, @mentions within the tweets – and their relations are the lines connecting pairs of nodes. This means accounts are connected if they mention one another. The strength (or weight) of this connection is based on the times mentioned by the same account, which captures both the extensive and the intensive margins of connections – that is, not just the presence of a connection, but also the strength of the connection as a measure of significance.

A matrix containing the accounts scraped and the accounts mentioned by them in the corpus of tweets was constructed. The open-source software Gephi (Bastian et al 2009) was used to import the matrix and construct the network graph. The force-directed algorithm “Force Atlas 2” was applied to show the spatialization of nodes by mapping the proximity and the authority of categories in relation to each other (Jacomy et al. 2014). This means that linked nodes are drawn closer while unrelated nodes are pushed farther apart, thus allowing for a visual interpretation of the dynamics between actors in the network. The modularity algorithm developed by Blondel et al. (2008) is then applied to identify “communities”, or clusters – as represented by nodes that are more densely connected together than to the rest of the network, and which were colored accordingly.

Similarly, the dynamics of scientific contributions was explored from the Scopus dataset. In this case, VOSviewer (van Eck and Waltman, 2010), an open-source bibliometric analysis software, was used to import the Scopus data and create network visualizations based on the declared affiliations of publication authors and the countries of their respective institutions.
3. Main results
3.1. CIMMYT’s major work areas within the food production-climate change nexus

From the four data sources collected, it was possible to systematically review CIMMYT’s research priorities in relation to climate change, as well as understand how the resulting knowledge products have been communicated to the public. Through text mining, we uncover trends and insights to answer the above question from three perspectives:

1. **At the institutional communications level**, how have climate issues been communicated in relation to CIMMYT’s strategic vision of resilient food systems and improved livelihoods?

2. **At the scientific production level**, what aspects of climate science has CIMMYT engaged with in its research, and how have these been integrated into its areas of expertise?

3. **At the broader public outreach level**, what climate conversations is CIMMYT part of?

The next sections present the main findings from each data source. A comparative summary is presented in the last section.
3.1.1. Evidence from strategic institutional communications

At the time of data retrieval, CIMMYT’s website had a dedicated news section for “Climate Change”, which contained updates, press releases and other communications materials.

Figure 4 shows a word cloud for the tags identified in the metadata of the publications. Word clouds visually represent the occurrence of words in a corpus, by sizing the terms according to their frequency and are an effective way to draw attention to high occurring terms as a first approximation of topic prevalence.

In this case, it is possible to see that the institutional communications of CIMMYT regarding climate change has given prominence to issues around resource and water management, as represented by the high frequency of ‘drought’, as well as to innovations such as conservation agriculture.

The custom algorithm to classify the two broad topics and their respective sub-topics was applied to the corpus containing the text of the website posts. Figure 5 shows their overall distribution, as frequency counts. The cross-cutting topics ‘Farming systems’, ‘Technology transfer’, ‘Food security’ and ‘Innovation’ are the most frequently covered subjects in CIMMYT’s communications of climate-related research. Regarding climate-focused topics, ‘Profitability’, ‘Productivity’, ‘Production’ and ‘Resource management’ appear most frequently in the text.
While the overall distribution of topics can uncover the cumulative prominence of themes, a temporal distribution provides a more nuanced perception of topic prevalence over time. Beyond the presence or absence of a topic, the algorithm developed to detect the selected themes also quantified their presence. This prevalence was then normalized on a scale from 0-1, where 1 represents the maximum value detected over the period of analysis. Figure 6 presents the time trends for the 22 cross-cutting topics. This visualization enables understanding when certain topics entered CIMMYT’s research programs and when they were most in focus. There seems to be a gap in updates regarding climate change research between 2008 and 2012, when the frequency picked up considerably. This may represent a shift in the organization’s climate agenda, which from 2012 onwards has been well integrated with other strategic issues.
Figure 6 Timelines for the prevalence of cross-cutting topics identified in the “Climate Change” page of CIMMYT’s website (topics normalized on scale from 0.0-1.0)
To further unpack the interlinkages between different concepts within the communications disseminated through CIMMYT’s website, a measure of correlation was established, which aims to identify when terms are present within the same body of text, in this case, the text of the update. A strong correlation indicates that the terms consistently occur within the same document. In figure 7, it is possible to see the correlation between climate change and the other cross-cutting topics. In CIMMYT’s institutional communications, climate change is most often discussed in relation to ‘Farming systems’, but is also associated to ‘Food security’, ‘Livelihoods’ and service delivery (‘Services’).

The normalized time trends for the presence of climate-specific topics are shown in Figure 8. While the majority of topics are consistently represented in the text, it is possible to see that ‘Carbon sequestration’, ‘Emission reduction’, and ‘Energy conservation’ have entered CIMMYT’s research agenda in the last five years.
Figure 8 Timelines for the prevalence of climate-focused topics identified in the "Climate Change" page of CIMMYT’s website (topics normalized on scale from 0.0-1.0)
Figure 9 presents the association between cross-cutting topics and climate-focused topics. Expectedly, climate change is strongly associated to mitigation and adaptation. Other significant correlations include between ‘Farming systems’ and ‘Conservation agriculture’, ‘Mixed cropping’ and ‘Agroforestry’, as well as between ‘Innovation’ and ‘Technology transfer’ with ‘Productivity’. These associations indicate the particular agricultural practices and technologies that are researched in relation to broader strategic themes of the institution.

Finally, the correlations between climate topics themselves were determined and are presented in figure 10.
Lastly, the map in figure 11 presents the distribution of countries identified in the location tags of the news items. India was the most frequently mentioned country, followed by Zimbabwe, Ethiopia and Bangladesh. At the regional level, the focus of climate change communications in CIMMYT’s website has been Africa (287 updates) and Asia (221 updates).

![Figure 10 Correlations between climate topics identified in “Climate Change” page of CIMMYT’s website](image)

![Figure 11 Distribution of countries identified in posts from CIMMYT’s climate change news webpage](image)
3.1.2 Evidence from CIMMYT’s knowledge products

The almost 2,500 knowledge products containing the term “climate” (or “clima” in Spanish) in CIMMYT’s Repository comprise 46% of the entire database of documents dating back to 1960, which enables broad coverage for text mining of the corpus of abstracts for relevant topics. A descriptive analysis of the metadata shows that the vast majority of items available are scientific outputs, such as articles and books, but there are also hundreds of reports, presentations, handbooks, brochures and other communications products. Besides their availability in the repository, the main publishers for the research include scientific journals by Elsevier, Springer and MDPI, but also funders and partners such as the Australian Centre for International Agricultural Research (ACIAR), USAID and CGIAR. Charts summarizing these descriptives are available in Appendix 1.

Figure 12 presents the distribution of countries identified by our country detection algorithm. In this case, CIMMYT’s climate research has focused on Mexico and India, as the most frequently mentioned countries, followed by Bangladesh, Ethiopia, Kenya, Pakistan, and Nepal.

In this case, CIMMYT’s climate research has focused on Mexico and India, as the most frequently mentioned countries, followed by Bangladesh, Ethiopia, Kenya, Pakistan, and Nepal.
The classification algorithm was applied to the corpus containing the abstracts of the repository items. Figure 13 shows their overall distribution, as frequency counts. As this dataset was already constrained by the presence of the term “climate”, the cross-cutting topic climate change was excluded from the figure.

The most frequently detected cross-cutting topic was ‘Farming systems’, followed by ‘Food security’ and ‘Technology transfer’. With regards to climate-specific topics, ‘Profitability’ and ‘Productivity’ were detected in over 700 publications each, followed by climate change mitigation and adaptation.

To understand the historical evolution of CIMMYT’s research outputs, Figure 14 presents the time trends for the 22 cross-cutting topics. Overall, we note an increase in prevalence among all topics, most notably, ‘Food security’, ‘Enterprises’, ‘Health’, ‘Stress’, and ‘Yields’. Regarding climate change, while a presence was detected across the outputs, the timeline clearly shows two peaks in 2006 and 2016.
Figure 14 Timelines for the prevalence of cross-cutting topics identified in climate-related knowledge products available in the CIMMYT Repository (topics normalized on scale from 0.0-1.0)
To determine the linkages between climate-focused research and other strategic issues in CIMMYT’s research agenda, Figure 15 presents the correlation between climate change and the remaining cross-cutting topics. It shows that climate change is most closely associated to ‘Capacity development’ and ‘Technology’, but also strongly correlated to ‘Farming systems’, ‘Technology transfer’, and ‘Innovation’.

Lastly, we assess the presence of climate-specific topics. The normalized time trends are shown in Figure 16. While an increase is also noticeable across all topics, indicating increased attention to the specific technologies and practices that impact agricultural production’s climate sensitivity, adaptation and mitigation demonstrate consistent growth in prevalence since the beginning of this century. Sustainable intensification shows a sharp increase in the last five years, demonstrating a heightened interest in integrated approaches to build farmer resilience.
Figure 16 Timelines for the prevalence of climate topics identified in climate-related knowledge products available in the CIMMYT Repository (topics normalized on scale from 0.0-1.0)
Figure 14 presents the association between cross-cutting topics and climate-focused topics. Capacity development is strongly associated to risk management and to production. Climate change regularly appears in the same documents as productivity and profitability. Likewise, risk management is often addressed together with innovation and technology transfer.

![Correlations between climate topics and cross-cutting topics identified in climate-related knowledge products available in the CIMMYT Repository](image)

The correlations between climate topics themselves are shown in Figure 18. Similar to the results shown for CIMMYT’s website, the expected association between climate change adaptation and mitigation is also visible, as is the strong correlation between ‘Productivity’ and ‘Profitability’. Yet, another notable association identified in this particular corpus is among ‘Climate smart agriculture’ (CSA), ‘Precision agriculture’, and ‘Nutrient management’ – the high co-occurrence of these topics in the repository documentation suggests the integration of climate-sensitive technologies.
3.1.3 Evidence from scientific production

The Scopus dataset provided a clean source from which to explore peer-reviewed, scientific publications by CIMMYT-affiliated researchers. From the 3,523 results extracted until the end of 2020, just over a third of all indexed publications since 1974 – a total of 1,266 – were found to cover topics related to climate change and were used for this analysis. The word cloud in figure 19 shows the most frequent terms detected in the abstracts of climate-related publications. Besides wheat and maize, the core of CIMMYT’s work, research has focused mostly on agricultural production, represented by the term yield, and seed improvement, represented by terms such as genetic, resistance, and breeding.
Almost all countries worldwide have been mentioned in some capacity.

Graphs providing the descriptives of CIMMYT’s overall outputs indexed by Scopus are available in Appendix 2. Among other things, they show that while the institution has published extensively in journals from the fields of agricultural and biological sciences, there have also been many papers featured in biochemistry, genetics and molecular biology, as well as in environmental and social sciences.

Figure 20 maps the countries detected in the corpus of abstracts and shows that Mexico has been mentioned in more than 200 publications. India, Zimbabwe and Ethiopia have been identified in 80, 45 and 43 abstracts, respectively. However, the geographical coverage of CIMMYT publications is remarkable, as almost all countries worldwide have been mentioned in some capacity.
The aggregated frequency distribution of topics, as detected by the custom algorithm, is shown in figure 21. Once again, as the corpus is restricted to abstracts containing the term ‘Climate change’, this topic was excluded. Contrary to the CIMMYT Repository, ‘Food security’ was the most frequent cross-cutting concept found among the publications, with almost twice the mentions compared to the second topic, ‘Farming systems’, at 835 and 479, respectively, suggesting that peer reviewed publications have focused on the relationship between climate and food security as a major concern. Another variation from the Repository is that, while ‘Profitability’ and ‘Productivity’ are also the most frequent topics detected in the abstracts, they are followed by ‘Resource management’, ‘Water management’, ‘Risk management’ as the top climate-focused topics.

The incorporation of cross-cutting topics over time is presented in the timelines in figure 23. Following the trends identified in the repository documents, climate change shows a significant increase in prevalence among indexed publications. Other consistent upward trends include for ‘Farming systems’, ‘Food Security’, ‘Mixed cropping’, ‘Stress’, ‘Technology’ and ‘Technology Transfer’, as well as ‘Yields’. ‘Gender equity’ has gained space in CIMMYT’s research agenda in the last five years.
Figure 22 Timeline for the prevalence of cross-cutting topics identified in CIMMYT-affiliated publications indexed by Scopus (topics normalized on scale from 0.0-1.0)
The correlation between cross-cutting topics and climate change is shown in figure 24. Climate change is associated to almost all topics, except health, whereas the strongest co-occurrence is with ‘Farming systems’, followed by ‘Stress’ and ‘Yields’.

Figure 23 Correlations between ‘Climate change’ and cross-cutting topics identified in CIMMYT-affiliated publications indexed by Scopus

The time trends for climate topics are presented in figure 25 show the evolution of traditional crop science, such as resistant plant varieties or resource and water management, as well as broader concepts like productivity and profitability, which have been researched for many decades prior to the association to the effects of climate change. The figure also evidences the scientific advance of new processes, such as climate change adaptation and mitigation, and of novel approaches like conservation agriculture, carbon sequestration, sustainable intensification and zero tillage.
Figure 24 Timelines for the prevalence of climate topics identified in CIMMYT-affiliated publications indexed by Scopus (topics normalized on scale from 0.0-1.0)
Figure 25 presents the associations between the climate and cross-cutting topics. The strongest correlations are between climate change and adaptation and mitigation. Strong co-occurrence is also noticeable between ‘Socioeconomic development’ and CSA, as well as ‘Socioeconomic development’ and ‘Precision agriculture’.

The correlations within climate topics in figure 26 support findings from the other data sources, identifying strong relationships between adaptation and mitigation, between ‘Productivity’ and ‘Profitability’, and between CSA, ‘Precision agriculture’, and ‘Nutrient management’. ‘Resource management’ and ‘Water management’ also co-occur frequently in the same publications.
3.1.4 Evidence from social media

Twitter data provides a broader perspective of public, non-scientific perceptions of CIMMYT’s work and how its expertise is recognized by the wider community. First, we identify the geographical coverage of conversations that CIMMYT has engaged with. The presence of countries is even more widespread in this dataset, with Mexico and India still the most frequent, followed by Kenya, Ethiopia and Pakistan.
A first approximation of prevalent topics that CIMMYT has engaged with on the social networking platform is through hashtag analysis. Hashtag use is a typical affordance of Twitter, with topics and issues frequently denoted by specific hashtags (Resce and Maynard, 2018). Besides the expected wheat and maize hashtags, the word cloud in figure 28 identifies a strong presence of food security and climate change in the conversations.

However, as word clouds reflect absolute frequencies, which are influenced by the level of activity of particular accounts on Twitter, we employed the more in-depth measure of topic prevalence in the corpus of tweets. In this analysis, biased data is tackled by normalizing the prevalence over a period of time, thus keeping any potential influence constant throughout the period of analysis. Figure 29 shows the frequency distribution of topics. As noted, there is a
generally even distribution among the two higher level themes and their sub-topics, though the most frequent are in line with the other data sources: ‘Farming systems’ and ‘Food security’ for cross-cutting topics, and ‘Profitability’ and ‘Productivity’ for climate-focused topics. However, as the Twitter dataset does not contain specific criteria to select only climate focused content, it is important to note that ‘Climate change’ is among the top mentions.

The quantification of topics is particularly important in this case – when the majority of topics are present to a similar extent – in order to unpack the variability between them and any emerging trends. Hence, figure 30 shows the timelines for the normalized prevalence of the cross-cutting topics. As the time span for the Twitter dataset is shorter than for the other sources, the majority of topics were already widely discussed within the agricultural development community, and as such, are mostly present since 2009, with the exception of ‘Diversification’, ‘Gender equity’, ‘Livelihoods’, ‘Nitrification’ and soil-borne diseases (‘Nematoda’). In fact, many topics, such as capacity building, food security, resource management and participation had their peak prevalence in the beginning of last decade. This is also the case of ‘Farming systems’, although its prevalence has been steadily on the rise for the last five years, showing the strong interest in conversations about holistic perspectives to agricultural development. With regards to climate change, the topic has maintained a consistent presence over the last ten years.
Figure 30 Timelines for the prevalence of cross-cutting topics identified in Tweets containing either @CIMMYT or #CIMMYT (topics normalized on scale from 0.0-1.0)
Figure 31 presents the correlation of climate change with the cross-cutting topics. The strongest association is with ‘Enterprises’, followed by ‘Capacity development’.

Lastly, we assess Twitter dialogue concerning climate-focused topics and in this case, in figure 32 we see some emerging technologies and practices, such as ‘Carbon sequestration’, CSA, ‘Energy conservation’, ‘Nutrient use efficiency’ and ‘Zero tillage’, all of which have entered the conversations in recent years, following the trends identified in the other datasets.
Figure 32 Timelines for the prevalence of climate topics identified in Tweets containing either @CIMMYT or #CIMMYT (topics normalized on scale from 0.0-1.0)
When the interlinkages between climate topics and cross-cutting topics are established through the correlation chart in figure 33, the results differ from the other data sources: the strongest co-occurrences in the Twitter conversations put connect ‘Tolerance’ to ‘Participation’ and to ‘Services’, and ‘Mixed cropping’ to ‘Agroforestry’.

Figure 33 Correlations between climate topics and cross-cutting topics identified in Tweets containing either @CIMMYT or #CIMMYT.
The final correlation visualization in figure 34 shows the associations within climate topics and support the results from the other data sources that show CSA’s positive co-occurrence with ‘Precision agriculture’ and ‘Nutrient management’, as well as the strong correlations between climate change adaptation and mitigation, and between ‘Productivity’ and ‘Profitability’.

Figure 34 Correlations between climate topics identified in Tweets containing either @CIMMYT or #CIMMYT
The geographical focus of CIMMYT’s climate research centered on three main countries: Mexico, India, and Ethiopia.

Farming systems, Food security, Technology transfer, and Innovation were the most frequent cross-cutting topics across the four data sources.

Profitability, Productivity, Production, Resource management, Adaptation and Mitigation were the most frequent climate-focused topics in all content assessed.

Farming systems and Food security were the cross-cutting topics with the most significant and steady rise in prevalence over time.

The climate-focused topics that have emerged in the last five years across the datasets are Carbon sequestration, Emission reduction, Energy conservation and Sustainable intensification.

Since climate change mitigation and adaptation processes are often intertwined, there was a strong association between the two in all datasets.

Likewise, as an increase in profitability is an expected outcome of higher productivity, the two concepts co-occurred within the same posts or publications in all datasets.

While institutional communications on CIMMYT’s website and on social media have addressed the effects of climate change on the livelihoods of rural people, this connection was not reflected in the institution’s scientific and non-scientific knowledge products, which focused on research around the development of agricultural technologies and practices to support food production and cope with climate impacts.

There was variability among the correlations between cross-cutting topics and climate-focused topics in the different datasets. While CIMMYT’s website most frequently connected climate change to adaptation and mitigation measures, the strongest co-occurrences in publications from the Repository bring together capacity building with risk management and production, as well as climate change with productivity and profitability. Among indexed, peer reviewed publications, besides the interlinkages between climate change and mitigation and adaptation, the impacts of new approaches such as CSA and precision agriculture on socioeconomic development were also addressed. Lastly, Twitter conversations made the connection between stakeholder participation (specially women) in sustainable practices and increased tolerance to climate stresses.

Among climate-focused topics, besides the common trend noted above, CSA was associated to precision agriculture and nutrient management in scientific and non-scientific publications, as well as on social media conversations.

On Twitter, the majority of cross-cutting topics assessed were already well established within the agricultural development community when the platform was launched in 2009, which is reflected in their presence throughout the period of analysis. In fact, many topics, such as capacity building, food security, resource management and participation had their peak prevalence in the beginning of last decade. This is also the case of ‘Farming systems’, although its prevalence has been steadily on the rise for the last five years, indicting heightened awareness towards integrated approaches for sustainable agriculture.
3.2. Dissemination of CIMMYT’s climate-focused knowledge products

The hyperlink analysis conducted in this study was based on issue mapping approaches (Rogers, 2010), which aim to identify actors and geographies of particular issues as represented by hyperlink connections, among other techniques. In this case, we have adapted the approach to explore knowledge diffusion and uncover information flows. Of the 2,463 climate-related research outputs collected in CIMMYT’s Repository, our algorithm generated non-ambiguous results for 2,263 items, for a total of 55,151 web pages pointing to them. Out of those, more than 10 thousand were unique domains; additionally, over 150 countries were identified. Figure 35 breaks down the frequency of URLs by type of publication, as identified by Repository metadata. It shows that scientific research such as articles and books were by far the most frequently disseminated outputs, followed by outreach materials such as handbooks, reports and presentations.

Scientific research, such as articles and books were by far the most frequently disseminated outputs, followed by outreach materials such as handbooks, reports and presentations.

![Figure 35 Frequency of URLs by publication type](image)
Figure 36 shows the 25 most frequent domain names detected in the analysis. Books stored in Google Books contain more than 3,000 references to CIMMYT climate-related outputs, and the majority of frequent domains pertain to academic publishers such as Springer, or ScienceDirect, which reflects both the scientific activity of CIMMYT as well as citations in external papers. However, government partners such as the Agricultural Knowledge Resources and Information System Hub for Innovations (KRISHI) of the Indian Council of Agricultural Research (ICAR), and the Food and Agriculture Organisation (FAO) are also featured among key multipliers of CIMMYT’s research.

**Figure 36 Frequency of top 25 domain names (CIMMYT website and repository excluded)**

---

**TOP 25 DOMAINS**

1. books.google.com
2. krishi.icar.gov.in
3. semanticscholar.org
4. researchgate.net
5. expertfinder.cgiar.org
6. guardian.bigtoday.cgiar.org
7. scholar.google.com
8. links.springer.com
9. mdpi.com
10. sciencedirect.com
11. agris.fao.org
12. repositorioacceso.ernet.mx
13. orcid.org
15. onlinelibrary.wiley.com
16. cambridge.org
17. x-mol.com
18. frontiersin.org
19. academia.edu
20. nature.com
21. oui.dnth.gov.ua
22. docplayer.net
23. journals.plos.org
25. t PeerJOnline.com

---

Figure 37 presents a network visualization of the top-level domains (TLD) extracted from the URlS and their respective top domains. TLDs are split into two types: generic TLDs that do not correspond to a particular country or region, such as .com and .org; and country-code TLDs, which represent a geographical location, such as .uk or .mx. In this network, we see that the majority of sites sharing CIMMYT’s knowledge products are “.com”, a TLD generally used by commercial institutions, followed by “.org”, which are often used by non-profit organizations. Three other TLDs are well represented in the graph: “.gov” (used by government bodies), “.edu” (reserved for educational institutions) and “.net” (generally used by online service providers).

The TLD distribution is an approximation of the types of entities picking up and disseminating CIMMYT’s outputs. It shows that CIMMYT reaches a variety of institutions, from commercial publishers and private companies to civil society organizations, universities, and governments.
In figure 38, a further breakdown of the most frequent domains with " .com " TLD shows that despite a prevalence of commercial publishers, CIMMYT content is also disseminated across social networking platforms such as LinkedIn, Twitter and Facebook. We detected CIMMYT’s presence in 13,630 URLs from 2,543 unique domains. A similar breakdown of " .org " domains is shown in figure 39. In total, 12,467 URLs were detected from 1,334 domains. Here we see again many academic repositories, but also partners and funders such as the CGIAR, FAO and the World Bank.
**Figure 38** Frequency of top 25 .com domain names

**TOP 25 .COM DOMAINS**

- scholar.google.com
- link.springer.com
- mdpi.com
- onlineibrary.wiley.com
- sciencedirect.com
- x-mol.com
- nature.com
- linkedin.com
- books.google.com
- tandfonline.com
- pubfacts.com
- eurekamag.com
- scrbd.com
- biomedcentral.com
- issuu.com
- academic.oup.com
- cdnsciencemag.org
- twitter.com
- academic-accelerator.com
- businessdocbox.com
- facebook.com
- academic.naver.com
- journals.sagepub.com
- yumpu.com
- readcube.com

**Figure 39** Frequency of top 25 .org domain names

**TOP 25 .ORG DOMAINS**

- figshare.org
- semantic scholar.org
- fan.org
- orcid.org
- frontiersin.org
- cambridge.org
- journals.plos.org
- academicjournals.org
- scielo.org
- repec.org
- europepmc.org
- worldcat.org
- booksc.org
- bioRxiv.org
- worldbank.org
- meta.org
- dx.doi.org
- apstnet.org
- cyberleninka.org
- diggfor.org.tr
- sciorg.org
- aminer.org
- wikipedia.org
- worldwidescience.org
- researchgate.org

(CIMMYT website and repository excluded from analysis)
Lastly, figure 40 presents the geographical distribution of the country-code TLDs. The countries colored in red contain TLDs with more than 180 URLs; the orange shades are between 60 and 180; and yellow are between 1 and 60. A key limitation must be noted in this particular analysis: in general U.S.-based institutions do not use their country specific TLD (.us), normally opting for the general TLDs, which has resulted in an under-representation of U.S.-based domains.

CIMMYT’s outputs have been distributed to more than 150 countries across both the Global North and South. Domains in North America, Australia, and Europe have frequently disseminated CIMMYT research, but also in South and Southeast Asia, and Latin America.
3.3. How and with whom is CIMMYT engaging in the network of climate action?

Network analysis has proven to be a key tool for understanding the structure and operation of a particular network (Normann, 2017; Soomai et al., 2013). Issue networks can be defined as meso-level social structures consisting of a configuration of social relations among interdependent actors, which form around particular issues or policy problems (Bolleyer and Börzel, 2010). Insights emerging from network analysis have driven recommendations regarding the facilitation of contact, collaboration and relationships as central drivers of climate science dissemination within the scientific community, to the public and to policymakers (Oliver et al., 2014).

3.3.1 Network of climate action

One of the key affordances of Twitter is the ability to interact with other users through the @mention, when an account is tagged and notified, giving it the ability to engage in direct dialogue. Figure 41 shows a word cloud for the accounts mentioned in the Twitter corpus (CIMMYT’s main English account excluded), in which the main CGIAR account is the most prevalent, followed by USAID (@usaid), CIMMYT’s Spanish sustainable agriculture account (@accimmyt), the institution’s Chief Operating Officer (@bramaccimmyt), and the Bill & Melinda Gates Foundation (@gatesfoundation).

Figure 41 Word cloud of account mentions identified in Tweets containing either @CIMMYT or #CIMMYT
Taking the data collected from Twitter and focusing on a sub-set of approximately 5,000 tweets about climate, we assessed the dynamics between tweets by CIMMYT and tweets that mentioned the organization either through @CIMMYT or #CIMMYT. The complete network of mentions derived from this subset contained more than 2,700 unique accounts, connected almost 8,000 times, as shown in Figure 42.

In order to uncover the key actors within this broad network, the analysis considered only accounts that were mentioned at least five times (i.e. in-degree distribution). This criterion reduced the number of nodes to 704, with 4,755 connections between them, thus making the network denser, with nodes connected on average to 6.8 other nodes. The sizes of the labels have been set according to the node’s in-degree centrality. The resulting visualization in Figure 43 shows four major clusters. The largest is the blue cluster, which represents 42% of the network and contains CIMMYT as the key actor. Within this cluster we also find funding organisations such as USAID, DFID, IFAD and the Gates Foundation close to CIMMYT, which indicates interaction between these accounts. The red cluster represents 25% of the network and contains CCAFS’s account (@CGIARclimate) as a central figure. Other CG Centers are also prominent, such as IITA, CIAT, and ILRI and ICRAF. The pink cluster contains CIMMYT’s Spanish language communications.
Figure 44 shows another analysis, in which a different centrality metric was applied to establish the most influential nodes in the network. The sizes of the nodes correspond to their eigenvector centrality, a common metric for social media analysis that measures the influence of actors within a network by considering not only how many connections a node has, but also the centrality of the nodes that it is connected to (Hansen et al, 2011). Essentially, the importance of a node is measured by how much it is connected to other important nodes in the network, and typically on social media, actors with high eigenvector centrality are important centers of attention. Expectedly, based on this metric the most influential nodes are CIMMYT and CCAFS (@CGIARclimate). The main CGIAR profile and IITA are also prominent actors.
The networks display a visual representation of the dynamics of information exchanges related to climate change, in which CIMMYT occupies a central place connecting actors and influencing information flows. As the organization dialogues with many accounts within the network, it holds a prominent role through both centrality measures applied here. While it must be noted that this network is based on CIMMYT mentions, and hence may be biased towards the institution, the visualization demonstrates that CIMMYT actively engages with key players within international development and research, supporting the dissemination of climate knowledge to a broad range of actors. Also, while social media “echo chambers” are commonly portrayed as a negative phenomenon, a connected and cohesive group of likeminded entities can strengthen the call for climate action, as they are able to exchange knowledge, build up narratives, and organize concerted actions (Pearce et al. 2019), which reflects the figures presented.

While it must be noted that this network is based on CIMMYT mentions, and hence may be biased towards the institution, the visualization demonstrates that CIMMYT actively engages with key players within international development and research, supporting the dissemination of climate knowledge to a broad range of actors.
3.3.2. Network of scientific cooperation

The Scopus dataset contained author affiliation metadata that could be displayed as networks of scientific collaboration. The network analysis was performed on 1,289 climate-focused publications that identified more than 500 unique institutions. First, figures 45 to 47 present the most prominent authors, affiliations (besides CIMMYT) and funders identified in this subset.

Figure 45. The 25 most frequent authors identified in CIMMYT-affiliated climate-focused publications indexed by Scopus.
Figure 46 The 25 most frequent institutional affiliations identified in CIMMYT-affiliated climate-focused publications indexed by Scopus.

Figure 47 The 15 most frequent funders identified in CIMMYT-affiliated climate-focused publications indexed by Scopus.
As institutional affiliations are not standardized in the Scopus index, the dataset was manually checked and cleaned for duplicates – different departments or regional offices were merged into their top-level institution. Figure 48 presents a network visualization for the top collaborations between CIMMYT and research organizations. In this case, the colors are coded by organization type: Universities (in blue, representing 58% of the network), CGIAR Research Centers (in red, comprising 20% of the network), public research institutions (in green, representing 20%) and the private sector (in teal green, representing 1.5%). Notably, the most frequent scientific collaborations occurred with the University of California, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), the Indian Council of Agricultural Research (ICAR), the U.S. Department of Agriculture’s Agricultural Research Service (USDA-ARS) and the International Rice Research Institute (IRRI).

A network analysis performed on 1,289 climate-focused publications identified more than 500 institutions.
A different functionality offered by the bibliometric analysis software was employed to map the geographical distribution of these institutional collaborations, by their attributed countries. Figure 54 presents this visualization, with the color scheme indicating the average publication year for those countries’ institutions, which enables determining the timeframes of collaborations and consequently uncovering the evolution of research agendas. The most frequent countries identified in the institutional affiliation of authors were Mexico, the U.S., and India. Countries in yellow such as Germany, Canada, Tanzania or Zambia represent more recent co-authorships, as the average publication date is 2020, while purple nodes comprise older collaborations, namely with the Philippines, Peru and Austria, among others.
4. Conclusions and recommendations
CIMMYT’s mission of improving livelihoods through wheat and maize science inevitably recognizes the interdependence between the many dimensions of food systems. This systematic review aimed to locate the institution’s climate change-food production nexus research portfolio within this integrated framework, seeking to uncover not only what CIMMYT has been doing in relation to the topic, but also how climate sensitivity is articulated with other strategic themes.

Text mining techniques enabled a comprehensive assessment of CIMMYT’s climate-related outputs at various levels. The diversification of CIMMYT’s climate portfolio is clearly visible through the increased prevalence of research on various climate-sensitive approaches. Whereas only ten or twenty years ago, the climate discourse was virtually non-existent and food production was still focused on large scale, intensive agriculture – specially with maize and wheat as the main crops in the green revolution – the data shows a significant shift towards sustainability and the development of a suite of innovative technologies and practices. The climate change-food production nexus is also represented by the consistent increase in food security as a prevalent topic within climate-focused scientific and non-scientific knowledge products. Additionally, CIMMYT’s climate-related research is strongly incorporated into a holistic systems approach, as seen by the significant association between climate change and farming systems across the various datasets. Hence, the data supports CIMMYT’s integrated research approaches and reinforces its promotion of technical innovations to strengthen farming systems and increase food production.

Regarding other emerging issues in the climate research agenda, carbon sequestration, emission reduction and energy conservation in relation to agri-food systems have recently started to gain traction both in the literature and on social media conversations that CIMMYT has engaged with.

Taking CIMMYT’s outputs as a starting point, the hyperlink analysis set off to discover their information pathways across the web. It found that knowledge generated through CIMMYT’s research has been shared across thousands of websites in over 150 countries, confirming its strongest presence on academic and research platforms, but also on social media, governments, and international organization websites from both the Global North and South. As CIMMYT’s strategic plan commits to generating “expertise, innovation and knowledge at the local and global scales” (CIMMYT, 2016), we see a reflection of this aim reproduced online.

Finally, the network analyses revealed that CIMMYT is actively engaged in knowledge exchanges with thousands of key players from the scientific, development and public policy communities. While CIMMYT’s expertise in traditional maize and wheat research is well-established, the increased Twitter activity around climate mitigation and adaptation demonstrates the institution is well positioned to inform and influence conversations that improve the capacity of stakeholders to better understand, articulate, and promote climate-sensitive agriculture.
The network of scientific cooperation identified through Scopus data also supports CIMMYT’s pursuit of co-production and inter-institution relationship building, as the organization has collaborated with hundreds of respectable institutions to enhance climate research. Overall, results support the notion that CIMMYT is effectively engaging in the uptake and diffusion of climate science, at both the global and the local level.

Lastly, we have the following recommendations arising from this study:

1. To strengthen CIMMYT’s science-policy interface: while policy was present in the text mining analysis, it was not a prominent topic, both within the institutional data sources and in the broader Twitter conversations. If CIMMYT intends to increase its policy-informing role to create enabling policy environments (CIMMYT, 2016), the development of a targeted advocacy strategy that maps policy actors and proposes specific outreach is essential.

2. This could also include the ‘translation’ of scientific outputs to more accessible products: as seen by the repository publication types, scientific articles and books are not only the most produced research outputs; they are also the most disseminated despite the existence of many other non-scientific resources. This presents a potential for CIMMYT to create opportunities for exchanges beyond the scientific community and to have a bigger impact on policymaking.

3. To increase the visibility of emerging climate-sensitive approaches: insights from the data show us there is room for CIMMYT to diversify its narrative beyond its traditional and well-established research agenda, to create more spaces for emerging technologies and innovations. Specifically drawing back to Twitter topic prevalence, the absence of conversations around the aforementioned emerging issues does not mean CIMMYT is not working on them, but that perhaps the broader community is not yet aware of CIMMYT’s contributions.

4. To expand the application of digital methods and machine learning to support the systematization of CIMMYT’s research portfolios. As seen in this study, innovative approaches can leverage on existing research and programmatic data to uncover insights, by employing data-driven methods that enable data repurposing to answer new questions. The ability to repurpose and analyze large historical datasets supports knowledge management and can inform future strategic planning. Such an approach not only enables gaining further insight into data, but also delivers added value.
References


Carneiro, B., Resce, G., Ruscica, G., Ma, Y., Pacillo, G., & Läderach, P. (2020). A web analytics approach to map the influence and reach of CCAFS A web analytics approach to map the influence and reach of CCAFS (No. 326)


Appendix 1 – Repository Descriptives

Figure 50 Distribution of publication types for climate-related knowledge products available in the CIMMYT Repository

Figure 51 The 25 most frequent authors identified in climate-related knowledge products available in the CIMMYT Repository
Appendix 2 – Scopus descriptives

**Figure 53** Frequency of subject areas identified in CIMMYT-affiliated publications indexed by Scopus

**Figure 54** The 25 most frequent source titles identified in CIMMYT-affiliated publications indexed by Scopus
Figure 55: Distribution of author keywords identified in CIMMYT-affiliated publications indexed by Scopus.