

Climate Data and Analytical Tools — the Role of Innovation

*Elena Triebskorn**

Abstract

Climate and sustainability indicators are increasingly vital for monetary policy, financial stability, supervision, regulation, reserve management, and payment oversight due to the risks posed by climate change. However, current climate-related data is often incomplete, inconsistent, and difficult to compare, with notable gaps in regional, sectoral, household, real estate, and non-bank financial institution information, as well as insufficient forward-looking and nature-related indicators and incomplete corporate disclosures.

No single data source can address all these shortcomings, so institutions must rely on multiple sources, which improve robustness but increases costs and complexity. Innovative technologies like machine learning and artificial intelligence can help fill data gaps by extracting information, reducing delays, and generating new indicators, but they require human oversight to ensure transparency, quality, and interpretability, as they may introduce new risks such as model errors and greenwashing.

Effective use of these tools demands clear objectives, high-quality datasets, standardization, validation, and collaboration between technical and subject-matter experts. Cooperation among institutions, including international networks, is essential to harmonize standards, share expertise, and expand public access to climate data, which enhances transparency and supports research.

Central banks should transition from fragmented, reactive data collection to coordinated, standardized, and forward-looking climate statistics. The optimal approach combines international collaboration, diverse data sources, centralized management, public dissemination, and validated analytical technologies, ensuring innovative tools are used transparently and with strong human oversight.

Keywords: data gaps, sustainable finance data, climate data, sustainability reports, natural language processing, large language models.

1. INTRODUCTION

As sustainability issues become more important at the national and global levels, demand has grown for reliable data and measurement frameworks. Central banks' interest in sustainability

* Deutsche Bundesbank. The views expressed in this report are personal views of the author and do not necessarily reflect the views of Deutsche Bundesbank or the Eurosystem.

indicators has grown because they support a range of tasks, including monetary policy, financial stability analysis, statistics, supervision, regulation, reserve management and payment oversight. As data users, central banks often use available information, particularly environmental, social and governance (ESG) metrics, to guide their policies. As data producers, they increasingly develop new or experimental sustainability indicators. Central banks have focused on sustainable finance — financial activities that improve environmental or social conditions or governance practices — and green finance, which focuses specifically on environmental aspects.

This article discusses three avenues to help bridge climate-related data gaps. First, international dialogue among stakeholders can foster the exchange of knowledge and experience. The Banco de España conference on “Climate Risk and Sustainable Finance: Bridging Policy, Research and Practice” established a forum for such an exchange with its session on climate data and analytical tools. Secondly, a growing range of innovative analytical approaches can provide more timely data and fill data gaps. Thirdly, the paper examines how the Deutsche Bundesbank has responded to shortages in climate-related data and identifies lessons for developing climate statistics.

2. TAKEAWAYS FROM THE BANCO DE ESPAÑA CONFERENCE ON “CLIMATE RISK AND SUSTAINABLE FINANCE: BRIDGING POLICY, RESEARCH AND PRACTICE”

Despite strong policy demand for better and more detailed insights into sustainability, important information gaps persist. The first session, “Climate Data and Analytical Tools,” showed why reliable, decision-useful climate data are important for financial analysis, risk assessment and policy development. The session featured presentations from representatives of the Network for Greening the Financial System (NGFS), the TPI Global Climate Transition Centre at the London School of Economics and the Banco de España. Participants then discussed climate scenarios, transition assessments, statistical production and data gaps.

Agnieszka Trzcinska presented the NGFS climate scenarios, which help financial institutions and policymakers assess the potential economic and financial effects of climate change. She explained that the NGFS provides both long-term and short-term scenarios. The long-term scenarios extend to 2050 and explore different combinations of transition and physical risks. The recently introduced short-term scenarios focus on effects up to 2030, including rapid or delayed transitions, severe physical hazards and combined transition and physical risks. These scenarios combine climate policy, extreme weather events, macroeconomic trends, sectoral dynamics and emerging transmission channels, such as critical mineral supply bottlenecks.

A key message was that climate risks are not just a distant “tragedy of the horizon.” Severe physical-risk scenarios suggest that economic losses could approach the scale of the global

financial crisis. Under particularly adverse conditions, European gross domestic product could decline by as much as 5% by 2030. Future NGFS work will explore more plausible acute physical-risk shocks, adaptation, geopolitical developments, nature-related risks and interactions with broader macroeconomic shocks. Important challenges remain, particularly the lack of sufficiently granular subnational data and the limited ability to model effects on households and real estate markets.

Carmen Nuzzo introduced the TPI Global Transition Centre, an independent academic organisation that assesses the progress of companies, banks and sovereigns towards a low-carbon economy. The Centre develops transparent, sector-specific methodologies based mainly on publicly available information. Investors use the data for risk assessment, portfolio alignment, screening, product development, reporting, active ownership and proxy voting. The Centre is increasingly moving beyond assessing climate ambition to evaluating actual implementation, capital spending, sectoral conditions and regional policy environments. Nuzzo emphasised that climate data must be rigorous, comparable, independent and transparent. She also stressed the need for a system-wide perspective across companies, banks and sovereigns.

Luis Ángel Maza discussed the Banco de España's experience of including climate considerations in official economic statistics. The Bank has developed databases on corporate environmental, social and governance indicators, economic activity locations, financial exposures, carbon footprints and physical risks. It uses sustainability reports, administrative data, commercial sources, machine learning and data science techniques to create granular information for internal analysis and external researchers. Collaboration with the European Central Bank, national statistical institutes, international organisations and accounting bodies has been essential. Future priorities include adding climate variables to national accounts, developing biodiversity and water-stress indicators, improving forward-looking measures and using artificial intelligence to create more efficient "climate intelligence".

The discussion highlighted the need to assess interactions between climate shocks and financial crises, improve consistency between short- and long-term scenarios, account for adaptation and analyse overlapping portfolios in the non-bank financial sector. Participants also stressed the importance of data sharing, harmonised standards, confidentiality solutions and using climate data to inform regulation and investment decisions.

3. FILLING CLIMATE-RELATED DATA GAPS: LEVERAGING INNOVATIVE APPROACHES

Central banks and supervisory authorities have started to integrate climate-related information into instruments such as stress testing (EBA, 2024) and collateral frameworks (ECB, 2026).

However, developing and implementing these instruments requires high-quality data. A lack of structured and complete issuer-level datasets can create significant data gaps and require analysts to spend considerable time collecting data. Generative artificial intelligence can help address these shortcomings, particularly by improving the availability and quality of issuer-level data. While modern artificial intelligence methods can extract and process un-structured data at scale, substantial challenges remain in validating, standardising and post-processing the data.

A growing range of innovative analytical approaches, including machine learning, can provide more timely data and fill data gaps. Like other statistics, climate risk and sustainability statistics are often published with a delay. This limits the ability to monitor risks in real time. Innovative analytical methods can reduce reporting delays and identify early signals. Effective monitoring of climate performance, risks and opportunities across financial asset classes, portfolios and institutions require credible metrics and robust data. This information is critical for informing investment decisions and designing effective policies, including policies to achieve the goals of the Paris Agreement. However, current disclosure initiatives are often too qualitative and incomplete. Access to high-quality publicly available climate data also remains fragmented, limiting the comparability and scalability of assessments. Innovative data solutions that use new technologies and techniques are emerging to help fill these gaps. While promising, they may create new transparency challenges and risks of greenwashing that need to be better understood. Addressing these trade-offs is critical to building trust and ensuring that these solutions are taken up credibly.

To examine the potential of innovative analytical tools and draw on experiences from various central banks, the Network for Greening the Financial System (NGFS) set up an exchange to share knowledge and experience on this topic through its Expert Network on Data. Central banks worldwide face similar difficulties in obtaining climate-related information. Cooperation can help them avoid duplicating work, pool expertise and use their comparative strengths. However, many initiatives are currently being developed independently, so stronger coordination is needed. The findings of the Expert Network on Data are summarised below.

3.1. The use of natural language processing for climate data

Climate reporting is mainly based on text. It covers documents such as scientific reports, policy papers, corporate disclosures and news articles. Central banks and supervisors need to process and analyse large volumes of this information to assess climate-related risks and opportunities. Natural language processing (NLP) can help address this challenge by extracting relevant information from unstructured text. NLP can automate the extraction and analysis of data from different sources. This can help institutions process large volumes of information, identify potential risks and make informed decisions.

NLP can support a wide range of tasks. Its most immediate use is to analyse companies' or financial institutions' disclosures of climate-related risks and impacts. It can also support more specific tasks, such as analysing how central banks and other supervisory authorities communicate about climate change. Different NLP techniques can significantly improve the analysis of climate-related text data.

A key challenge is the need for high-quality datasets with labelled data. Adapting models to specific tasks can be complex and resource-intensive. Integrating climate-specific data into existing models is also challenging. Performance and reliability can improve through comprehensive datasets, better model adaptation and a wider range of applications. Reducing costs and complexity can also make these tools more effective.

The first step in using NLP is to define clear objectives that can be put into practice. These objectives determine the overall design, data organisation and analytical methods needed. Clear objectives also ensure that work reflects institutional priorities and available resources. Common objectives include searching texts by meaning, retrieving climate-related information, analysing sentiment and classifying topics. These objectives provide a framework for using NLP or large language models (LLMs) to monitor climate-related discussions, analyse compliance with regulations and assess climate-related risks.

Setting realistic objectives requires an understanding of what NLP and LLMs can and cannot achieve, and where human expertise is needed. Institutions should therefore define objectives and develop a modelling approach in an iterative process that draws on subject matter expertise and technical knowledge.

3.2. Opportunities and challenges of machine learning applications for nature-related data

Nature-related data vary significantly in quality, coverage and consistency. Collecting these data is particularly constrained by geography because it often requires physical sampling. Assessing cascading effects requires more detailed geospatial data and data on interdependencies for many ecosystem services. Financial risk assessments are also hampered by the lack of standardised reporting frameworks. Machine learning offers opportunities but also poses challenges. It can identify statistical relationships between input and output data without knowing the underlying physical processes. However, it may overlook complex interactions and produce misleading assessments.

Machine learning (ML) models can generate data by classifying and predicting environmental parameters. This can provide useful insights for developing metrics and indicators. However, using ML to generate nature-related data may oversimplify environmental processes. It may

overlook complex interactions between parameters or undetected spatial patterns, leading to incomplete or misleading assessments. Furthermore, artificial intelligence (AI) can support the development and analysis of biodiversity data. Biodiversity data are essential for AI models and innovation. Companies whose long-term innovation strategies depend on AI and biodiversity may have an incentive to support biodiversity conservation. This could preserve existing biodiversity data and enable future discoveries that generate further data. AI can also support effective and efficient conservation by enabling complex, targeted responses to ecosystem degradation or biodiversity loss. It can provide accurate and reliable data for monitoring, reporting and verification by investors and financial institutions. Making this information available should ultimately help direct more funding towards biodiversity protection across industry sectors as links with biodiversity become more visible.

4. CENTRAL BANK STEPS — THE CASE OF THE DEUTSCHE BUNDESBANK

The Deutsche Bundesbank has responded to shortages in climate-related data and identified lessons learned for developing climate statistics. Climate change has become increasingly relevant to central banks because it creates both physical and transition risks. Physical risks result from events such as floods, heatwaves and other extreme weather events, while transition risks arise from political, technological and economic changes associated with the move towards a low-carbon economy. These risks affect central banks' responsibilities, including financial stability, monetary policy, banking supervision, market analysis and risk management. As a result, several divisions within the Bundesbank require reliable climate-related information.

A major challenge is that climate-related statistics and sustainable-finance data are relatively new. When Bundesbank analysts first needed these data, the institution lacked established internal sources and comprehensive statistical systems. Individual users therefore had to search for relevant information at public institutions or private providers and arrange procurement independently. Although this approach could meet immediate needs, it was inefficient, costly and likely to result in duplicated work. It also made it difficult to coordinate data requirements, compare information across departments and ensure consistent quality.

To address these problems, the Bundesbank established the Sustainable Finance Data Hub within its statistics department in early 2020. The hub centrally coordinates the supply of and demand for climate-related data. It identifies common and ongoing data needs through regular communication with internal users, examines available products on the market and procures data from external providers. It also serves as a central point of contact for questions about data sources, definitions and methodology, connecting Bundesbank users with data providers.

The hub is involved in nearly every stage of the data provision process. Its responsibilities include surveying the market, contacting providers, obtaining trial datasets, supporting procurement staff during contract negotiations and cleaning and processing the data once acquired. Data preparation can be technically difficult and time-consuming because providers often use different formats and structures. By centralising these tasks, the hub reduces duplication, creates economies of scale and allows individual departments to focus on their analytical responsibilities. The hub also contributes its experience to international statistical and sustainable-finance groups. This allows the Bundesbank to share its practical knowledge and learn from other institutions.

4.1. A multi-source approach to climate data

Experience has shown that a multi-source approach is necessary because climate-related data are subject to significant uncertainty and there is no universally accepted framework. Third-party providers often use different definitions, methodologies and estimation techniques. They may also cover different regions, markets, sectors or types of financial instruments. Consequently, no single provider can satisfy all potential users or provide a complete and fully reliable picture.

The Bundesbank obtains data from several providers for three main reasons. First, different departments have different analytical needs. Second, providers differ in the scope and geographical coverage of their datasets. Third, climate-related measurements may be uncertain, particularly when information is estimated or based on incomplete corporate disclosures. Comparing data from multiple providers allows analysts to perform robustness checks and identify differences in methodologies or results. A Bundesbank survey confirmed the value of this strategy: more than 70% of respondents who used climate-related data obtained information from two providers. Using multiple sources therefore improves confidence in analyses, even though it can increase procurement and data-management costs.

4.2. Improving public access to climate data

Another important element of the Bundesbank's response is improving public access to climate-related information. Public data remains scarce and is often dispersed among different institutions. By collecting, processing and presenting these data in a single location, the Bundesbank can support both its own analysts and the wider public. Greater accessibility improves transparency and helps raise awareness of climate change and its financial and economic consequences, including those of the transition to a more sustainable economy.

For this purpose, the Bundesbank created a public dashboard containing a selected group of climate-related macroeconomic and financial indicators. The dashboard presents the

information in charts and offers an accessible overview of developments in Germany and the European Union. Its indicators cover three broad areas. Financial indicators include green bonds, environmental tax revenues and fossil-fuel subsidies. Real-economy indicators include greenhouse-gas emissions and energy consumption. Climate indicators include measures such as average air temperature and the number of days with heavy rainfall. Together, these data illustrate both physical risks and transition-related developments.

The Bundesbank publishes the underlying data and sources alongside the dashboard. This strengthens transparency and allows users to understand how the indicators were compiled. The dashboard is updated annually and relies mainly on publicly available, free-of-charge data. However, information on certain sustainable financial products, such as green bonds, is not generally available from public sources. To address this gap, the Bundesbank reached an agreement with a private data provider that permits the publication of aggregated information on sustainable bonds. Although such arrangements are useful in the short term, the long-term objective should be to create publicly available datasets that comply with recognised statistical standards. Cooperation with private providers can temporarily bridge gaps while more permanent public systems are developed.

5. CONCLUSION

Climate and sustainability indicators help inform monetary policy, financial stability assessments, supervision, regulation, reserve management and payment oversight. Their importance is growing because climate change creates physical and transition risks for the economy and financial system. However, significant gaps in climate-related data remain. Available data are often incomplete, delayed, fragmented, qualitative or difficult to compare.

Important shortcomings include:

- insufficiently detailed regional and sectoral data;
- limited information on households, real estate and non-bank financial institutions;
- inconsistent definitions and methods among data providers;
- insufficient forward-looking indicators;
- limited data on biodiversity, water stress and nature-related risks;
- incomplete corporate disclosures.

No single climate data source can address these shortcomings. Differences in coverage, quality, definitions and estimation methods mean that institutions need to use several sources. Comparing data from different providers makes assessments more robust and helps identify methodological inconsistencies. However, it also increases procurement and data management costs.

Innovative technologies can also help close data gaps. Machine learning, natural language processing, generative artificial intelligence and other data science methods can:

- extract information from unstructured reports and news;
- reduce reporting delays;
- identify emerging risks;
- generate environmental and biodiversity indicators;
- improve monitoring, reporting and verification.

However, these tools do not eliminate the need for human expertise. Their outputs may be difficult to validate, reproduce or interpret. Machine-learning models may also overlook complex ecological or economic relationships. Therefore, transparency and quality control are essential when using artificial intelligence. Data generated or extracted by artificial intelligence can create new risks, including model errors, limited explainability and greenwashing. To use these tools effectively, institutions need clear objectives, high-quality labelled datasets, standardization, validation and an iterative process involving technical specialists and subject-matter experts.

Cooperation among institutions is also necessary. Central banks and other authorities face similar data problems. International dialogue and networks such as the Network for Greening the Financial System can reduce duplicated work, share expertise, harmonize standards and improve coordination. Public access to climate data should also be expanded. Dashboards and publicly available datasets improve transparency, support research and help the public understand the economic and financial consequences of climate change.

Central banks need to move from collecting fragmented climate information in response to immediate needs towards coordinated, standardized and forward-looking climate statistics. The most effective strategy combines international cooperation, multiple data sources, centralized data management, public dissemination and carefully validated analytical technologies. Although innovative tools can significantly improve the timeliness and coverage of sustainability data, they must be used transparently with strong human oversight and recognized statistical standards.

REFERENCES

ALONSO-ROBISCO, A., CARBO, J. M., KORMANYOS, E., & TRIEBSKORN, E. (2024). Houston, we have a problem: Can satellite information bridge the climate-related data gap? *Deutsche Bundes-bank Technical Report*, 2024-02.

ARAUJO, D., BRUNO, G., CAP, A., MARCUCCI, J., SCHMIDT, R., SIRELLO, O., and TISSOT, B. (2026). Generative artificial intelligence in central banking. *IFC Bulletin*, no 67.

ARSLANALP, S., KOSTIAL, K., and QUIRÓS-ROMERO, G. (eds.). (2023). *Data for a greener world: a guide for practitioners and policymakers*. IMF.

BARNAVA, A., DUIJM, P., NOELS, J., SCHMIEDER, C., SEVEN, Ü, SEYHUN, Ö, TISSOT, B., TRIEBSKORN, E. (2026). The quest for policy-relevant sustainability data in central banking. *IFC Bulletin*, no 68.

BOLTON, P., DESPRÉS, M., PEREIRA DA SILVA, L., SAMAMA, F., SVARTZMAN, R. (2020). *The green swan – central banking and financial stability in the age of climate change*. BIS.

BUNDESBANK. (2023). *Climate-related disclosures by the deutsche bundesbank*. DeutscheBundes bank.

CAMPIGLIO, E., DEYRIS, J., ROMELLI, D., & SCALISI, G. (2025). Warning words in a warming world: Central bank communication and climate change. *Grantham Research Institute on Climate Change and the Environment Working Paper*, 418

DOLL, H., KORMANYOS, E., WALTER, S., and ALVES WERB, G. (2026). The climate data iceberg – a depth of information to integrate. *IFC Bulletin*, no 66.

EUROPEAN BANKING AUTHORITY. (2024). Annual Report.

EUROPEAN CENTRAL BANK (ECB). (2025). *Climate-related financial disclosures of the ECB's non-monetary portfolios*.

EUROPEAN CENTRAL BANK (ECB). (2026). ECB to extend use of climate factors in Eurosystem collateral framework to non-financial corporate credit claims. ECB Press Release. July 24th.

FEHR, M., TRIEBSKORN, E., and MEHRHOFF, J. (2022). Constructing forward-looking climate relat-ed physical risk indicators. 11th Biennial IFC Conference on Post-pandemic landscape for central bank statistics, BIS Basel, August 2022.

GONZALEZ, C. I., and TRIEBKORN, E. (2023). The road to net zero: the role of transition plans and forward-looking indicators in portfolio management. *Economic Bulletin, Banco de España*, issue 2023/Q3.

GONZALEZ, C. I., and TRIEBKORN, E., AND FÉLEZ DE TORRES, I. (2025). The Puzzle of Forward-Looking Climate Transition Risk Metrics (2025). *Banco de España Occasional Paper No.*, 2515.

NEFZI, D., NOELS, J., PERONACI, R., SCHMIEDER, C., SEVEN, Ü, SEYHUN, Ö, TISSOT, B., and TRIEBKORN, E. (2025a). Addressing climate change data needs: the global debate and central banks' contribution. *IFC Bulletin*, no 63.

NETWORK FOR GREENING THE FINANCIAL SYSTEM (NGFS). (2022). Final report on bridging data gaps. *NGFS technical document*.

NETWORK FOR GREENING THE FINANCIAL SYSTEM (NGFS). (2026b). Navigating nature-related data: metrics, sources and uses. *NGFS Technical document*.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). (2022). ESG ratings and climate transition: an assessment of the alignment of E pillar scores and metrics. *OECD business and finance policy papers*, no 6.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). (2025). *Behind ESG ratings: unpacking sustainability metrics*.

PARK, T.-J. (2025). *Quantifying Material Risks from Textual Disclosures in Financial State-ments using LLM Agents – A Case Study on Physical Risks*. IFC.

TASKFORCE ON NATURE-RELATED FINANCIAL DISCLOSURES (TNFD). (2023). Recommendations of the Taskforce on Nature-related Financial Disclosures.