



Getting the big picture in cross-domain fusion

A research agenda for scenario compilation and stakeholder interaction

Reinhard von Hanxleden¹ · Arne Biastoch² · Nicola Fohrer³ · Matthias Renz⁴ · Athanasios Vafeidis⁵

Accepted: 16 June 2022 / Published online: 27 July 2022
© The Author(s) 2022

Abstract

A central promise of cross-domain fusion (CDF) is the provision of a “bigger picture” that integrates different disciplines and may span very different levels of detail. We present a number of settings that call for this bigger picture, with a particular focus on how information from several domains can be made easily accessible and visualizable for different stakeholders. We propose harnessing an approach that is now well established in interactive maps, which we refer to as the “Google maps approach” (Google LLC, Mountain View, CA, USA), which combines effective filtering with intuitive user interaction. We expect this approach to be applicable to a range of CDF settings.

Introduction

Socio-environmental systems are highly complex and all components are interconnected. Thus, making sustainable and integrated management decisions is challenging. Classical decision support systems typically suffer from inade-

quate modeling of the socio-environmental systems and in most cases end users are not involved in their design. To address this shortcoming, numerical models have been developed in several disciplines, based on empirical findings or process understanding. The use of these models to analyze the impact of driving forces like climate, land use, or other anthropogenic interventions has been a research focus for many years, as has been scenario analysis for decision support.

One of the main obstacles in assessing potential future developments is the lack of direct interaction or the temporal delay between stakeholder requests. Such “Digital Twins” become even more important if different environmental disciplines are involved. Methods for the analysis of the impact of driving forces applied on single sources of data have been well elaborated and understood in recent decades. However, these data sources usually feature a small, narrow, and isolated snapshot of a specific concept of the real world (view). We are still lacking systematic approaches to integrate influencing features across different disciplines, data sources including stakeholder interactions, and models. Approaches enabling and supporting the integration of stakeholder (domain) knowledge are key elements of the framework of methods that enable cross-domain fusion (CDF).

In this paper, we make the case for adopting an approach that is successfully used by applications such as Google Maps for efficiently presenting complex map data for stakeholder interaction and related applications that call for CDF.

✉ Reinhard von Hanxleden
rvh@informatik.uni-kiel.de

Arne Biastoch
abiastoch@geomar.de

Nicola Fohrer
nfohrer@hydrology.uni-kiel.de

Matthias Renz
mr@informatik.uni-kiel.de

Athanasios Vafeidis
vafeidis@geographie.uni-kiel.de

- ¹ Department of Computer Science, Kiel University, Kiel, Germany
- ² GEOMAR Helmholtz Centre for Ocean Research and Kiel University, Kiel, Germany
- ³ Institute of Natural Resource Conservation, Dep. Hydrology and Water Resources Management, Kiel University, Kiel, Germany
- ⁴ Department of Computer Science, Kiel University, Kiel, Germany
- ⁵ Department of Geography, Kiel University, Kiel, Germany

This approach involves not only the visualization itself, but also the interactive and user-friendly modeling, which we aim to provide for general CDF settings.

In the next section, we set the stage by motivating the need for interactive decision support systems. We then define what we consider to be the “Google Maps approach.” This is followed by discussions on how this approach maps to the interpretation of climate and ocean model data, as well as of location-based decision support systems.

From numerical models to digital twins: the need for decision-support systems

Numerical models are important tools for system understanding and decision making. Earth system models (ESMs) provide insight into the climate system, with several degrees of complexity simulating processes from the ocean, atmosphere, cryosphere, and land surface, and provide projections for the future under global warming [11]. ESMs are complex numerical constructs due to the range of variables and spheres representing physics, chemistry, biology, and geology. The quality of models is typically judged by how closely they match reality, and often, high spatial resolution is key for a realistic representation [1]. The combination of a heterogeneous representation with high grid resolution makes ESMs one of the most demanding numerical representations. It typically requires specialists and long integration times on high-performance computers (HPC), hampering the interactive and flexible use for what-if scenarios.

An example compartment of the Earth system is the ocean. Simulations of the global ocean and the coastal waters are required to evaluate societally relevant changes such as regional sea-level rises. Ocean circulation models simulate the movement and hydrography of the ocean. They use a consistent physical theory based on the Navier–Stokes equations and applied for seawater on a rotating earth [9]. Boundary conditions such as the shape of continents and seafloor as well as driving forces from the atmosphere (e.g., winds, heat fluxes) enable “realistic” representations of ocean currents, water masses, and their dynamic evolutions under natural variability and climate change.

Carrying out oceanic simulations requires planning and programming of a specific configuration, the performance on a dedicated HPC architecture, and the analysis and visualization of high-volume output fields. However, both experiment planning and implementation as well as decision-making require more flexible approaches. Such “digital twins” are needed to enable non-specialists to interactively design, set up, and perform state-of-the-art ocean model configurations in a series of options to perform simulations under what-if scenarios, such as global change.

“Stakeholders” could involve marine scientists, e.g., for the planning or interpretation of observations, or even authorities for using the digital twins for actual disasters (oil spill, ship safety) and planning.

Decision-support systems in hydrology were established in the mid-1990s, when the development of hydrological models allowed a reasonable depiction of hydrological processes and human interactions. With the increase of calculation power and the availability of real-time data, these systems progressed to near real-time predictions of, e.g., floods and allowed a model-based decision on mitigation measures [16]. As geographical information systems became available, integrated modeling systems for river basin management were developed [13]. The complexity of these systems broadened towards water quality and agricultural management and allowed the analysis of different management options and their impact on the protection of water resources, as well as the impacts of land use and climate change.

However, the focus of the work so far has been strictly on natural processes. Stakeholders and decision makers have been excluded from software development and its application. While the developed modeling systems were successful in the academic world, most of them never found their way into day-to-day decision making. Even those that were explicitly developed to support decision making were rejected by the decision makers. The complex model systems needed either too much expertise for their parameterization, calibration, and conducting of model runs, or the model results could be only interpreted by scientists. As Brown and Kytä state for participatory mapping (PM), a key research priority is “the development of more effective methods for analyzing data from PM, especially the meaningful aggregation of spatial data, usability of PM applications, and better understanding of user behavior and responses based on the user experience with the PM technology” [4].

Exploring complex models: the Google Maps approach

One underlying challenge in the above setting is not the lack of information, but rather the suitable filtering and presentation of the relevant information to the decision makers. Domains with long experience in addressing that problem are cartography and Geographic Information Systems (GIS), where for one area the actual level of detail and selection of information depend highly on the scale and type of map. The advent of interactive mapping tools such as Google Maps have newly emphasized the importance (and difficulty) of proper *automatic* information filtering. While the concept appears very natural, its realization is not trivial at all and does not always work as expected. For example, the

name of a city may disappear and re-appear several times when zooming in. However, for the most part the mapping applications work quite well, both with respect to usability and utility [19].

We refer to the approach of interactively inspecting complex models, of geographic nature or not, with a user experience inspired by modern mapping applications, such as the “Google Maps approach.” This approach is clearly not specific to Google Maps, but in our experience that term seems quite effective in conveying what we mean by it. Technically, that approach is characterized by the following:

1. The amount of data/information potentially visualized, i.e. the size of the underlying models, is much larger than can be meaningfully represented at once on a screen.
2. To control which part of the data/the model should actually be shown, users apply very simple gestures, basically just zooming and panning.
3. Based on the user selection of what part of the model should be shown, the tool applies some filtering to the model to reduce the amount of information such that it can be displayed well on the screen. Performing this filtering well is a central feature of the tool.

A general research question that we propose to investigate is how to effectively apply the Google Maps approach to other domains beyond mapping. Our hypothesis here is that the Google Maps approach can facilitate efficient stakeholder interaction across domains, due to its intuitive and familiar user experience that requires little to no domain-specific knowledge.

One challenge when trying to adapt the Google Maps approach for exploring complex systems is that they may not always be clearly associated with spatial data. One such example is the schematic process descriptions mentioned in the following section on visual interpretations of climate

data. There is already a large body of algorithm engineering research on this done by the graph drawing community in computer science [2]. This automatic graph drawing is already used effectively for example for software system design and exploration [10, 18]. However, combining the utility of the Google maps approach with automatic graph drawing is still a largely open problem that we aim to jointly investigate.

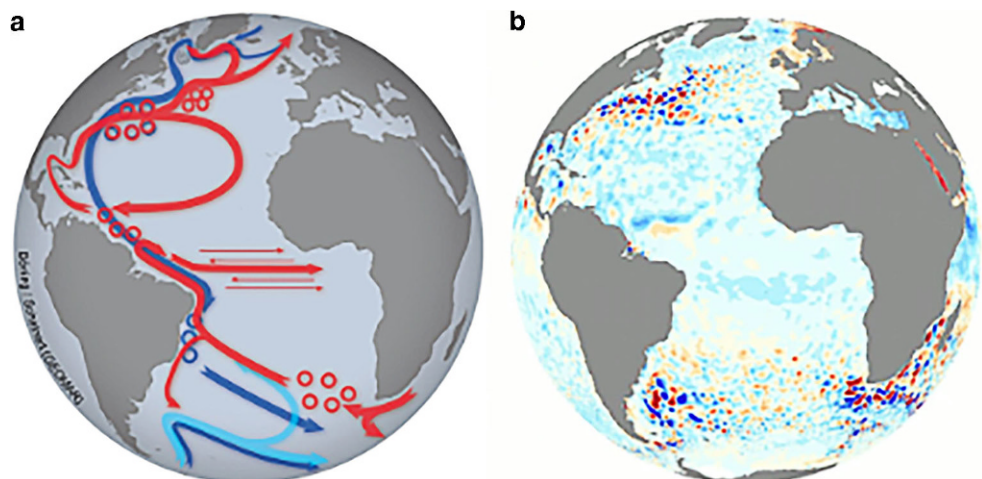
Another, related research question is how to perform the filtering. Part of that question, which must be answered in an interdisciplinary way, is what type of information needs to be represented, what form of knowledge is required [17]. In Google Maps, it is clear that a map of Europe should not show individual buildings, but should focus, e.g., on capital cities. For other domains, as proposed here, it may be less obvious how to filter information automatically, and how the underlying data should be structured. CDF techniques, which may integrate data across very different granularities, may be a key enabler for viewing systems on vastly different scales.

Another challenge is that the results to be presented are often subject to risk or uncertainty. Thus, one research question is how to ensure that this is suitably clarified to the user in an intuitive manner. It should be noted here that many people find it difficult to grasp risk or uncertainty even in significantly less complex representations.

An example: visual interpretations of ocean model data

The exploration of climate predictions and ocean model experiments is a specific task for visualization. Changes in ocean currents and hydrography form the basis for implications for marine ecosystems and protected areas, as well

Fig. 1 Schematics of warm surface (*red*) and cold deep (*blue* colors) flows in the Atlantic Ocean (**a**), in contrast to a snapshot of sea level elevation (*red/blue* colors for positive/negative elevations) as measured from a satellite (**b**)



as consequences for marine spatial planning and coastal protection.

Climate and ocean model experiments exist in large numbers, often precomputed with HPC. However, these are usually too large (Terabytes of data) and too complex to be explored and visualized by non-experts. Visual interpretations are required to summarize the content such as ocean currents and their changes over time. To overcome simple averaging filters, the Google Maps approach as outlined above would allow exploration of climate data in an interactive and intuitive sense. Ideally, a final abstraction level would make it possible to draw schematic, network-like representations of processes and their changes to enable decision-making.

Figure 1 gives an example: The climate scientist is interested in the general functioning of the physical ocean circulation—in this case the interplay between warm surface and cold deep currents in the Atlantic Ocean that regulates climate and is expected to decline under global warming [11]. In contrast, snapshots of the ocean currents (in this example, sea level measured from a satellite or simulated by an ocean model) show a rather noisy pattern, the “oceanic weather” masking the mean field. The Google Maps approach should be able to seamlessly slide between the snapshot and the mean schematics. In an optimal sense, this visual interpretation would also include uncertainties.

Location-based decision support and recommendation

The flood of publicly available geo-referenced data fuels an advanced analysis of big data applications beyond the current research frontiers and provides tremendous potential for discovering new and useful knowledge for spatial decision support and location-based recommendation [6]. Nowadays, many user activities generate data through social media that are annotated with location and contextual information, building a valuable source of data to discover new and useful knowledge. For example, recommendation systems in the social context use geo-social co-location mining in order to find social groups that are frequently found at the same location, as proposed in [21], giving insights into interactions between social groups that are very useful for social-link prediction and recommendation.

With the advent of smart devices and online social networks, a multi-billion dollar industry has emerged that utilizes the fusion of geo-referenced data and social data across domains for advertising and marketing [5]. Geotagged social-media posts, global positioning system (GPS) traces, as well as data from cellular antennas and WiFi access points are used widely to directly access people for advertising, recommendations, marketing, and group purchases.

Furthermore, in recent years, crowdsourced data has become a very important additional source of data. In addition to common quantitative metrics, this source of data enables recommendation and decision support systems to take qualitative features into account, for instance the scenery or the touristic attractiveness of a route provided in travel guide systems [12, 20].

The incorporation of multiple attributes and diversity of user preferences requires multi-preference recommendation approaches. In the data management community, such approaches have been enabled by the introduction of the skyline operator [3] that computes the pareto-optimal solution in multi-dimensional preference feature spaces. In recent decades, skyline computation became a hot research topic due to its importance for recommender systems and high computational requirements. Several scalable algorithms [14], extensions of the topic [8], and adaptations dedicated to transportation recommendation [15] have been proposed.

The interlocking of the approaches mentioned above for the exploitation of geo-referenced data with CDF techniques provides tremendous potential to recommendation and decision support systems in the context of socio-environmental domains.

As socio-hydrological mechanisms become better known, it is now necessary to bring both natural sciences and stakeholder interaction under the roof of one modeling system. It must be assured that both are represented adequately and inclusively, so that complex socio-environmental processes can be analyzed and communicated with modern visualization techniques in space and time. A first step would be to include stakeholders in the scenario development process, carrying out interviews to analyze stakeholder perception and use their knowledge for the joint development of relevant scenarios by a consortium of stakeholders and model developers [7]. More research is needed to allow for an integrative stakeholder approach to unlock the full potential of science-driven models for sound socio-environmental decisions. Modeling results need to be aggregated to a level useful to support the communication process with the stakeholders. Thus here, we advocate a zoom-in zoom-out approach filtering the model results to an adequate complexity.

Conclusions

In this paper, we have made the case for applying the “Google Maps approach” that is successfully used for user-friendly visualizations of very large geographic models to domains outside classical cartography, such as decision support systems, climate data exploration, location-based decision support or developing what-if scenarios. Effectively,

this is a socio-environmental-*technical* setting, where efficient tooling is a key component. Preliminary results, e.g., from exploring complex software models, indicate that such a transfer might be quite rewarding.

The underlying research questions, which reach across domains, include how to perform effective filtering, how to represent uncertainties, and how to use the approach in domains not linked to spatial data. This involves a broad range of computer science areas and communities, including algorithm engineering, data science, graph drawing, and human–computer interaction.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Biastoch A, Schwarzkopf FU, Getzlaff K, Rühls S, Martin T, Scheinert M, Schulzki T, Handmann P, Hummels R, Böning CW (2021) Regional imprints of changes in the atlantic meridional overturning circulation in the eddy-rich ocean model VIKING20X. *Ocean Sci* 17:1177–1211
2. Binucci C, Brandes U, Dwyer T, Groneman M, von Hanxleden R, van Kreveld M, Mutzel P, Schaefer M, Schreiber F, Speckmann B (2019) 10 reasons to get interested in graph drawing. In: Steffen B, Woeginger G (eds) *Computing and software science*. Springer LNCS, vol 10000, pp 85–104
3. Borzsony S, Kossmann D, Stocker K (2001) The skyline operator. In: *Proc. 17th Int. Conference on Data Engineering*. IEEE, pp 421–430
4. Brown G, Kytta M (2018) Key issues and priorities in participatory mapping: toward integration or increased specialization? *Appl Geogr* 95:1–8
5. Bouros P, Magdy A, Renz M, Sarwat M, Züfle A (2020) *Proc. of the Sixth Int. ACM SIGMOD Workshop on Managing and Mining Enriched Geo-Spatial Data*. GeoRich@SIGMOD 2020, Portland, Oregon, USA, 14 June 2020 ACM
6. Bouros P, Dasu T, Kanza Y, Renz M, Sacharidis D (2021) *The Fourth ACM SIGSPATIAL Workshop on Location-Based Recommendations, Geosocial Networks and Geoadvertising: online event*, Nov. 3, 2020: LocalRec 2020 workshop report. SIGSPATIAL Special 12(3):7–10
7. Cairns G, Iftekhar A, Mullett J, Wright G (2013) Scenario method and stakeholder engagement: critical reflections on a climate change scenarios case study. *Technol Forecast Soc Change* 80(1): 1–10
8. Chan CY, Jagadish HV, Tan KL, Tung AK, Zhang Z (2006) June. Finding k-dominant skylines in high dimensional space. In: *Proc. of the 2006 ACM SIGMOD Int. Conference on Management of Data*, pp 503–514
9. Griffies SM (2004) *Fundamentals of ocean climate models*. Princeton University Press
10. von Hanxleden R, Lee EA, Motika C, Fuhrmann H (2012) Multi-view modeling and pragmatics in 2020—position paper on designing complex cyber-physical systems. In: *Proc. of the 17th Int. Monterey Workshop, revised selected papers*. LNCS, vol 7539, pp 209–223
11. IPCC (2021) *Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change* (in press)
12. Jossé G, Schmid KA, Züfle A, Skoumas G, Schubert M, Renz M, Pfoser D, Nascimento MA (2017) Knowledge extraction from crowdsourced data for the enrichment of road networks. *Geoinformatica* 21(4):763–795
13. de Kok JL, Kofalk S, Berlekamp JB, Hahn HG, Wind HG (2009) From design to application of a decision-support system for integrated river-basin management. *Water Resour Manag* 23:1781–1811
14. Kossmann D, Ramsak F, Rost S (2002) Shooting stars in the sky: an online algorithm for skyline queries. In: *VLDB'02: Proc. of the 28th Int. Conference on Very Large Databases*. Morgan Kaufmann, pp 275–286
15. Kriegel HP, Renz M, Schubert M (2010) Route skyline queries: a multi-preference path planning approach. *2010 IEEE 26th Int. Conference on Data Engineering (ICDE 2010)*. IEEE, pp 261–272
16. Nayak PC, Sudheer KP, Ramasastri KS (2004) Fuzzy computing based rainfall–runoff model for real time flood forecasting. *Hydrol Process* 19(4):955–968
17. Pohl C, Krütli P, Stauffacher M (2017) Ten reflective steps for rendering research societally relevant. *GAIA* 26(1):43–51
18. Rentz N, Smyth S, Andersen L, von Hanxleden R (2021) Extracting interactive actor-based dataflow models from legacy C code. In: *DIAGRAMS '21*. LNCS, vol 12909, pp 361–377
19. Roth RE (2013) Interactive maps: what we know and what we need to know. *J Spatial Inf Sci* 6:59–115
20. Skoumas G, Schmid KA, Jossé G, Schubert M, Nascimento MA, Züfle A, Renz M, Pfoser D (2015) Knowledge-enriched route computation. *Int. Symposium on Spatial and Temporal Databases*. Springer, Cham, pp 157–176
21. Weiler M, Schmid KA, Mamoulis N, Renz M (2015) Geo-social co-location mining. *Second Int. ACM Workshop on Managing and Mining Enriched Geo-Spatial Data*, pp 19–24

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.