**An in situ sharing approach of data in the collaborative environmental modeling**

**Introduction**

**The significance of collaborative modeling**

**Modeling is not only a knowledge elicitation process but also a cooperation and negotiation process (SJBA Hoppenbrouwers et al.,2005).** **With the increasing need of collaboration among domain modelers, to more collaborative engagements gradually emphasized attention (Frost et al.,2007; Michiel Renger et al., 2008;** **Peter Rittgen,2009;). Collaborative modeling gathers dispersed various domain experts, stakeholders and users to participate the communication of problems, ideas and solutions as a team in the geographic modeling effort (Maxwell et al.,1997; Balram et al.,2006;).** **Supported by the feedback form participatory communicative process and effectively collaborative modeling, trans****disciplinary rather than isolated geographic solutions and comprehensive scientific approach such as problem structuring method (Rosenhead, J et al., 1993), group modeling building (Andersen, D.F et al.,2007) and enterprise analysis(Morton, A. et al., 2003,** **Dean, D.L et al.,1994) contribute to the complicated geo-problems solving (Xu, B et al.,2011; Horsburgh et al., 2016;** **Min Chen et al., 2019;). In fact, by leveraging the collaborative modeling, many geo-problems and geo-solutions have been solved and improved effectively, such as decision support in water resources (Langsdale, S et al., 2013), land use (Hewitt et el., 2104), regional-scale brine migration (Scheer, D et al.,2017), spatial information management (Rambaldi, G et al.,2006), sharing irrigation water in watershed (Gurung, T. R. et al.,2006), environment planning (Zellner, M et al., 2012) and marine spatial planning (Martin, K et al.,2008).** **Thus, there is generally agreed that collaborative modeling can establish trust, enhance stakeholders’** **comprehension of modeling experience in active team communication, ensure the stakeholders to participate fully in the construction of models and facilitate the formation of better solutions, which is support to the integrated environmental modeling, collective decision-making, management, scientific planning and other problems finally (Batten, D. F ,2009; Voinov, A. et al., 2010).**

**The need of data sharing in collaborative modeling**

**In fact, one of the most essential needs of collaborative modeling is appropriate datasets, besides model and compute resources, for data as the input drive most of the geo-analysis models such as GeoDetector, SWAT, FVCOM and SWMM to support the complicated geo-problems solving in the end (Shi, T et al., 2018; Chaubey, I et al.,2005; Yue, S et al.,2015; Lai, Z. et al., 2010). Due to the fact that the multiple levels of data needs vary from different modeling tasks and the organic integration of stakeholders, models and data in collaborative modeling increases the complexity of data requirements, it’s crucial that meet the needs of data in targeted level in collaborative modeling (Carver, S et al.,1996;** **Sallis, P et al.,2008). In this context, currently the needs of data in collaborative modeling can be divided into three levels in general, including file, information and presentation level.**

**In the file level, data mainly participated in the form of file in the collaborative modeling.** **Collaborative Building Information Modeling (BIM) process mainly depends on file transfer, for BIM data are stored and exchanged in the form of file with vendor specific formats (Shafiq MT et al., 2013; Zhang, J. et al.,2014; Afsari, K. et al., 2016;).** **High Level Architecture (HLA) as an advanced distributed simulation architecture, which can integrate models of domains to form simulation system, interact with other stakeholders by using binary format RID file (Tang S et al.,2010). Multidimensional space-time data are collaboratively shared by users in the NetCDF data format according to the developed approach in a next generation hydrologic information system (Gan, T et al., 2020). These studies of collaborative modeling require data in the form of file, which is reflected from the side of the file data requirements.**

**In the information level, data always are extracted into information or as data manipulation participating in the collaborative modeling. The data as the information are always key content of them rather than entire file or document, which is also required by many scenarios of collaborative modeling. The environment data for driver-pressure-state-impact (DPSI) models were retrieved from sensors (Leenhardt, P et al., 2017).** **Another part of the information level are geographical processing approaches (e.g., data mapping, refactoring), which can be regarded as the needs of information level in many collaborative scenarios. The modeling shell was developed for data manipulation (Argent, R.M et al., 2003;).**

**In the presentation level, the visualization of data is an important method for the participants to see what the data looks like. With the help of effective data visualization, stakeholders can make better understanding of data, interactive with their partners, conduct analysis and even get more feedback about models (Isenberg, P et al., 2011;** **Donalek, C. et al.,2014).** **From these three levels, the needs of data in collaborative modeling are obvious.**

**Nevertheless, now that the earth environment is a complicated giant open system containing dynamic interdependencies between system elements of various kinds, it’s no debate that collaborative modeling in geology involving interdisciplinary experts requires stakeholders’ sharing of data in file, information and presentation levels to advance the collaborative modeling. A shared version contributes to the successful research team’s collaborative efforts (Bennett, L. M. et al.,2012). The data should be shared more open,** **whether it's in the form of files, information, or visualization (Palmer, R. N. et al.,2013). The HydroShare web system is developed for sharing of hydrologic data to meet the needs of collaborative modeling, in which stakeholders can share their own data with other stakeholders** **(Tarboton, D.G. et al., 2014). Open Geographic Modeling and Simulation (OpenGMS) web platform also aggregates more than 2000 links of data downloading support to the collaborative modeling. The Open Street Map (OSM) collects data in a global scale shared by volunteers, namely stakeholders, guided by the concept of Volunteered Geographical Information (VGI) to support the collaborative communicate with local communities in the modeling (Basco-Carrera et al., 2017; Fonte, C. C et al.,2016). In a word, it’s strongly necessary for the stakeholders to share their data with their partners in the level of file, information and presentation to facilitate collaborative modeling.**



**Figure The requirements of data in three level in collaborative modeling**

**The disadvantages of data sharing method currently**

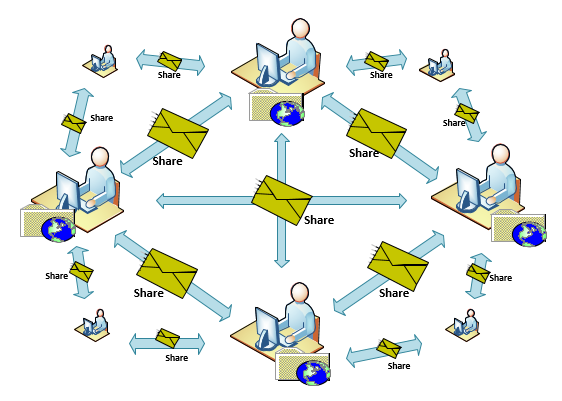
**However, there are many deficiencies for these data sharing methods obviously. The open sharing platform for datasets like HydroShare actually is a centralized collecting means tightly coupling with the various databases substantially (Tarboton, D.G. et al., 2014).** **Actually, this form of data sharing is not sufficient to meet the needs of dynamically operation of data resources collaboratively and has high cost of migration and system running. Similarly, in the aspect of data utilization, datasets of models are always sharing in the way of centralized uploading or downloading in whole, while part of them is needed in some participatory modeling cases. Moreover, in the aspect of security, the multi-source data resources are centralized managed free from data-owners’ respectively safety supervision increasing security risk. Although collaborative modeling based on OSM data is a great application of VGI, the quantity of data contributed from volunteers can’t be guaranteed (Foody, G. M. et al., 2013).**

**This study proposes approach**

**In this paper,** **a distributed in situ sharing method of data resources is proposed to support participatory modeling. Compared to the conventional methods, the approach makes modeling participants to share their own local data resource in file, information and presentation levels with other modelers and obtain other participants' local data resources on demand, which supports to the complicated geo-problems solving in participatory modeling finally.**

**Basic concept**

**In this paper, collaborative modeling for geo-problems solving require stakeholders to share their data resources with their collaborative modelers, which provides available as well as suitable data to the execution of models and interpretation of result. In the collaborative modeling preparation phase, multidisciplinary molders or pure data providers always prepare and maintain their data in the local storage. If the stakeholders keep their local data resources in isolation all the time, other modelers can’t run their models without suitable data.** **Only in the support of data sharing, can the collaborative modelers interact with each other to trigger the modeling work.**



**Figure In situ sharing**

**The in situ sharing of data approach expects that dispersed multidisciplinary modelers share their local data resources with other modeling partners on demand directly and immediately,** **for wider public involvement and the true democratisation of the modeling process. The data resources aren’t collected, prepared and centralized managed conventionally any more. In contrast, the data resources represent the characteristics of distributed, high quality and low overhead, for every computer can be regard as a data saving node in the network nowadays and dispersed data-owners preserve their local data assets in the distributed network structure. Under collaborative modeling requirements, distributed local data resources in the distributed structure network are shared with collaborators by owners from their storages in computers through the Internet after the collaborative communication and ideas exchanging.** **Without complicated data configuration and normalization, multi-source heterogeneous data can be shared and acquired straightly according to the requirements of specific modeling task relying on network communication among distributed nodes, which means modeling the end-to-end connections of the work in the aspect of data.**



**Figure** **data share and reuse through Internet**

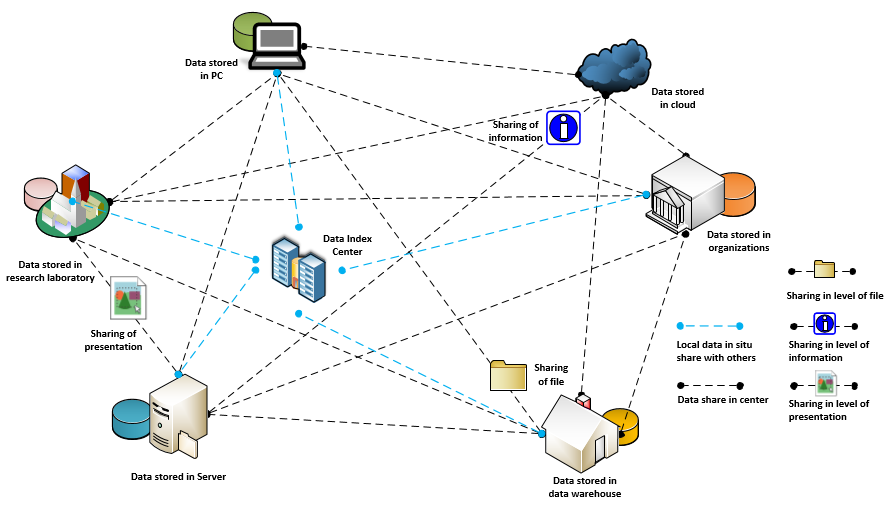
**Moreover, aimed at different levels of data needs in collaborative modeling, this approach supplies to the distributed data sharing in file, information and presentation levels. The file level mainly focuses on the data sharing in the form of entire document with specific formats, such as Shapefile and Tiff data. The information level is the sharing of information containing in the data, which can be metadata or part of the whole data. Besides, the processing approaches of data that should be shared and reused to conduct the preprocessing and postprocessing of data in the collaborative modeling such as resampling and interpolation are involved. For the presentation level, the sharing of data visualization provides an intuitive impression to the participatory modelers who may not understand the structure and organization of data at all, which lets they catch sight of what the data looks like at least.**

 **Figure In situ sharing in file, information and presentation level**

**All in all, all the efforts of the approach are committed to advancing the reusing, expedient accessing, science processing and making effective using of data needed in the collaborative modeling serving the complicated geo-problems solving.**

**Details of the distributed in situ sharing of data**

**Network solutions**



**Figure Network structure**

**In the aspect of the network structure, this study conceived a network solution that on the one hand various distributed local data storage of nodes including person computer (PC), private cloud, research laboratory, server, organization, data warehouse, etc. are able to share local data resources with each other in the levels of file, information and presentation and on the other hand share them to the center server, which is used to publish and show all available data to the public, as the Fig show. Relying on current popular network transmission techniques like HTTP, WebSocket, etc., the demand of sharing can be exchanged and then distributed stored data resources can be transmitted to other modeling tasks on demand directly. Based on those, the data-owners maintain independently their local data in their storages as the nodes of storage distributed in the network structure and data are exchanged end-to-end between the nodes immediately when the demand is made.** **Thus, the direct exchange among the nodes in the network structure results the approach of in situ sharing technically.**

**At the file level**

**File is the common storage medium of data resources and the primary form of sharing as well. In the file level, the in situ sharing approach enables distributed data file resources with specific formats (e.g., shp, tiff, etc.) to be shared with stakeholders on demand directly among collaborative modeling tasks.**

**At the information level**

**The in situ sharing in the information level mainly focus on the sharing of data processing approaches that can be applied or reused to conduct the manipulation (e.g., extracting, clipping, etc.) of data.**

**At the presentation level**

**In the presentation level, the in situ sharing includes the direct sharing of data visualization results on demands among stakeholders in collaborative modeling.**

**Case study**

**A case study about taking advantage of GeoDetector model to solve geo-problems under the background of collaborative modeling is described. Data of various risk factors are in situ shared by stakeholders to support the collaborative work.**

**Conclusion and future work**

**Acknowledgments**

**References**

1. Hoppenbrouwers, S. J., Proper, H. A., & van der Weide, T. P. (2005). Formal modelling as a grounded conversation.
2. Michiel Renger, Gwendolyn L. Kolfschoten, & Gert-Jan de Vreede. (2008). Challenges in Collaborative Modeling: A Literature Review. Advances in Enterprise Engineering I, International Workshop Ciao! & International Workshop Eomas, Held at Caise, Montpellier, France, June. DBLP.
3. Min Chen, Songshan Yue, Guonian Lü, Hui Lin, & Dawei Xiao. (2019). Teamwork-oriented integrated modeling method for geo-problem solving. Environmental Modelling and Software, 119.
4. Horsburgh, Jeffery S., Morsy, Mohamed M., Castronova, Anthony M., Goodall, Jonathan L., Gan, Tian, & Yi, Hong等. (2016). Hydroshare: sharing diverse environmental data types and models as social objects with application to the hydrology domain. Jawra Journal of the American Water Resources Association, 52(4), 873-889.
5. Maxwell, T., & Costanza, R. (1997). An open geographic modeling environment. Simulation, 68(3), 175-185.
6. Xu, B., Lin, H., Chiu, L., Hu, Y., Zhu, J., Hu, M., & Cui, W. (2011). Collaborative virtual geographic environments: a case study of air pollution simulation. Information Sciences, 181(11), 2231-2246.
7. Balram, S. (Ed.). (2006). Collaborative geographic information systems. Igi Global.
8. Langsdale, S., Beall, A., Bourget, E., Hagen, E., Kudlas, S., Palmer, R., ... & Werick, W. (2013). Collaborative modeling for decision support in water resources: Principles and best practices. JAWRA Journal of the American Water Resources Association, 49(3), 629-638.
9. Rittgen, P. (2009, January). Collaborative modeling-a design science approach. In 2009 42nd hawaii international conference on system sciences (pp. 1-10). IEEE.
10. Scheer, D., Konrad, W., Class, H., Kissinger, A., Knopf, S., & Noack, V. (2017). Regional-scale brine migration along vertical pathways due to CO 2 injection–Part 1: The participatory modeling approach. *Hydrology and Earth System Sciences*, *21*(6), 2739-2750.
11. Rambaldi, G., Kyem, P. A. K., McCall, M., & Weiner, D. (2006). Participatory spatial information management and communication in developing countries. The electronic journal of information systems in developing countries, 25(1), 1-9.
12. Zellner, M. L., Lyons, L. B., Hoch, C. J., Weizeorick, J., Kunda, C., & Milz, D. C. (2012). Modeling, Learning, and Planning Together: An Application of Participatory Agent-based Modeling to Environmental Planning. Journal of the Urban & Regional Information Systems Association, 24(1).
13. Martin, K. S., & Hall-Arber, M. (2008). The missing layer: Geo-technologies, communities, and implications for marine spatial planning. Marine Policy, 32(5), 779-786.
14. Rosenhead, J.: Rational analysis for a problematic world: problem structuring methods for

complexity, uncertainty and conflict (1993)

1. Andersen, D.F., Vennix, J.A.M., Richardson, G.P., Rouwette, E.A.J.A.: Group model

building: problem structuring, policy simulation and decision support. Journal of the Operational

Research Society 58, 691–694 (2007)

1. Morton, A., Ackermann, F., Belton, V.: Technology-driven and model-driven approaches

to group decision support: focus, research philosophy, and key concepts. European Journal

of Information Systems 12, 110–126 (2003)

1. Dean, D.L., Lee, J.D., Orwig, R.E., Vogel, D.R.: Technological Support for Group Process

Modeling. Journal of Management Information Systems 11, 43–63 (1994)

1. Frost, Sullivan.: Meetings Around the World: The Impact of Collaboration on Business

Performance. Frost & Sullivan White Papers, 1–19 (2007)

1. Hewitt, R., Van Delden, H., & Escobar, F. (2014). Participatory land use modelling, pathways to an integrated approach. *Environmental Modelling & Software*, *52*, 149-165.
2. Voinov, A., & Bousquet, F. (2010). Modelling with stakeholders. *Environmental Modelling & Software*, *25*(11), 1268-1281.
3. Fonte, C. C., Minghini, M., Antoniou, V., See, L., Patriarca, J., Brovelli, M. A., & Milcinski, G. (2016). An automated methodology for converting OSM data into a land use/cover map. In Proceedings of the 6 th International Conference on Cartography & GIS (pp. 462-473).
4. Shi, T., Hu, Z., Shi, Z., Guo, L., Chen, Y., Li, Q., & Wu, G. (2018). Geo-detection of factors controlling spatial patterns of heavy metals in urban topsoil using multi-source data. *Science of the Total Environment*, *643*, 451-459.
5. Chaubey, I., Cotter, A. S., Costello, T. A., & Soerens, T. S. (2005). Effect of DEM data resolution on SWAT output uncertainty. Hydrological Processes: An International Journal, 19(3), 621-628.
6. Yue, S., Wen, Y., Chen, M., Lu, G., Hu, D., & Zhang, F. (2015). A data description model for reusing, sharing and integrating geo-analysis models. *Environmental Earth Sciences*, *74*(10), 7081-7099.
7. Jang, S., Cho, M., Yoon, J., Yoon, Y., Kim, S., Kim, G., ... & Aksoy, H. (2007). Using SWMM as a tool for hydrologic impact assessment. Desalination, 212(1-3), 344-356.
8. Lai, Z., Chen, C., Cowles, G. W., & Beardsley, R. C. (2010). A nonhydrostatic version of FVCOM: 1. Validation experiments. *Journal of Geophysical Research: Oceans*, *115*(C11).
9. Tarboton, D. G., Idaszak, R., Horsburgh, J. S., Heard, J., Ames, D., Goodall, J. L., ... & Hooper, R. (2014). HydroShare: advancing collaboration through hydrologic data and model sharing.
10. Argent, R.M., Grayson, R.B., 2003. A modelling shell for participatory assessment and management of natural resources. Environmental Modelling and Software 18, 541-551
11. Leenhardt, P., Stelzenmüller, V., Pascal, N., Probst, W. N., Aubanel, A., Bambridge, T., ... & Salvat, B. (2017). Exploring social-ecological dynamics of a coral reef resource system using participatory modeling and empirical data. Marine Policy, 78, 90-97.
12. Basco-Carrera, L., Warren, A., van Beek, E., Jonoski, A., & Giardino, A. (2017). Collaborative modelling or participatory modelling? A framework for water resources management. Environmental Modelling & Software, 91, 95-110.
13. Bennett, L. M., & Gadlin, H. (2012). Collaboration and team science: from theory to practice. *Journal of Investigative Medicine*, *60*(5), 768-775.
14. Batten, D. F. (2009). Fostering industrial symbiosis with agent‐based simulation and participatory modeling. Journal of Industrial Ecology, 13(2), 197-213.
15. Gurung, T. R., Bousquet, F., & Trébuil, G. (2006). Companion modeling, conflict resolution, and institution building: sharing irrigation water in the Lingmuteychu Watershed, Bhutan. Ecology and society, 11(2).
16. Afsari, K., Eastman, C. M., & Shelden, D. R. (2016). Data Transmission Opportunities for Collaborative Cloud-Based Building Information Modeling.
17. Zhang, J., Liu, Q., Yu, F., Hu, Z., and Zhao, W. (2014). A Frameworkof Cloud-computing-based BIM Service for Building Lifecycle.Computing in Civil and Building Engineering. 1514-1521.
18. Shafiq MT, Matthews J, Lockley SR. (2013). A Study of BIM Collaboration Requirements and Available Features in Existing Model Collaboration Systems. Journal of Information Technology in Construction. Vol. 18. 148-16.
19. Tang, S., Xiao, T., & Fan, W. (2010). A collaborative platform for complex product design with an extended HLA integration architecture. *Simulation Modelling Practice and Theory*, *18*(8), 1048-1068.
20. Foody, G. M., See, L., Fritz, S., Van der Velde, M., Perger, C., Schill, C., & Boyd, D. S. (2013). Assessing the accuracy of volunteered geographic information arising from multiple contributors to an internet based collaborative project. Transactions in GIS, 17(6), 847-860.
21. Palmer, R. N., Cardwell, H. E., Lorie, M. A., & Werick, W. (2013). Disciplined planning, structured participation, and collaborative modeling—Applying shared vision planning to water resources. JAWRA Journal of the American Water Resources Association, 49(3), 614-628.
22. Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K. L., & Hagen, H. (2011). Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, *10*(4), 310-326.
23. Donalek, C., Djorgovski, S. G., Cioc, A., Wang, A., Zhang, J., Lawler, E., ... & Davidoff, S. (2014, October). Immersive and collaborative data visualization using virtual reality platforms. In 2014 IEEE International Conference on Big Data (Big Data) (pp. 609-614). IEEE.  Schwind
24. Thum, C., Schwind, M., & Schader, M. (2009, October). SLIM—A lightweight environment for synchronous collaborative modeling. In International Conference on Model Driven Engineering Languages and Systems (pp. 137-151). Springer, Berlin, Heidelberg.
25. Carver, S., Frysinger, S., & Reitsma, R. (1996, January). Environmental modelling and collaborative spatial decision-making: some thoughts and experiences arising from the I-17 meeting. In Proceedings 3rd international conference and workshop on integrating geographical information systems and environmental modelling.
26. Sallis, P., Shanmuganathan, S., Pavesi, L., & Muñoz, M. C. J. (2008, May). A system architecture for collaborative environmental modeling research. In 2008 International Symposium on Collaborative Technologies and Systems (pp. 39-47). IEEE.
27. Gan, T., Tarboton, D. G., Horsburgh, J. S., Dash, P., Idaszak, R., & Yi, H. (2020). Collaborative sharing of multidimensional space-time data in a next generation hydrologic information system. Environmental Modelling & Software, 104706.