Design of a Data-driven Intervention Dashboard for SDG Localization

Pooja Bassin, **Abraham G K**, **Srinath Srinivasa**
International Institute of Information Technology Bangalore, India
{pooja.bassin, abraham.gk, sri}@iiitb.ac.in,

Abstract

The localization problem of the United Nations Sustainable Development Goals (SDGs) involves adopting strategies that are in tune with local conditions, to achieve a given SDG target. However, even within a given region, localized conditions may vary drastically. With increasing amounts of Open Government Data (OGD) being available, there is an opportunity to systematically address the localization problem by using predictive and prescriptive modeling techniques. This work presents a predictive and prescriptive modeling dashboard for the SDG indicator *maternal deaths* (MD) for the Indian state of Karnataka. The dashboard was created by examining a vast set of data points to focus on four factors that showed high correlations with that of MD. We then construct a multivariate linear regression model to showcase the differential impact that a given factor has on the indicator and identify prescribed values for different factors to achieve a given target value of the indicator. Finally, a budget allocation dashboard is also provided that helps policymakers allocate budgets to specific schemes to help operationalize these changes. This dashboard was built by combining data coming from five different OGD sources.

1 Introduction

Sustainable Development Goals (SDGs) are designed by taking national frameworks into account but the success of SDGs strongly depends on their implementation at local and regional levels [Oosterhof, 2018]. The 17 SDGs with 169 targets and 231 unique indicators[1] may not be feasible or applicable uniformly worldwide. Every community has its unique challenges regarding socio-economic issues, cultural practices, environmental concerns, and government structures. The local stakeholders have a better understanding of the issues of their territory. Thus, the process of SDG localization promotes an understanding of the local context for the implementation of the SDGs at sub-national levels by encouraging national agencies to work in cooperation with the local and regional authorities. The provision of training, expertise, and resources to strengthen the local communities and institutions is part of the SDG localization process. Thus, local and regional governments need effective collaboration within the national framework such that transformational changes can be implemented considering local priorities and concerns.

To address the SDG localization problem in the Indian state of Karnataka, we use Open Government Data (OGD) for our intervention dashboard. In our present study, we focus on *maternal mortality/deaths* (MD) indicator from SDG 3: Good Health and Well-Being. Our intervention dashboard is built by combining health-related data from five different OGD sources - Karnataka At a Glance[2] (KAG), National Family Health Survey[3] (NFHS-4), Karnataka Health Profile[4], Men and Women in Karnataka[5] and eJanMa - Births, Deaths and Stillbirths Registration[6]. The intervention dashboard demonstrates predictive and prescriptive modeling to understand the impact of policy interventions [Bassin *et al.*, 2021; Bassin, 2022] on maternal health and also provides recommendations on budgetary allocations for various related schemes. Section 2 explains the methodology for performing prescriptive analysis. Section 3 presents a case study on reducing MD by executing the best course of action - 'what to do'.

Related Literature. Data visualization tools have been extensively developed for monitoring and reporting the progress of SDG indicators. [Garwood *et al.*, 2020] cater a dynamic collaborative dashboard for business schools and universities for SDG information sharing. [Bissio and Watch, 2016] report on diverse and independent dimensions such as gender equity, basics capabilities, etc. Dashboards on equity monitoring in vaccination coverage [Arsenault *et al.*, 2017], urban water security [Zainuddin *et al.*, 2023], health coverage [Azcuna, 2016] and, climate change [Canepa *et al.*, 2022] are primarily engaged as reporting tools for SDGs. Sustainable Development Report [Sachs *et al.*, 2023], SDG Urban Index and

[1]https://unstats.un.org/sdgs/indicators/indicators-list/

[2]https://kgis.ksrsac.in/kag/

[3]https://rchiips.org/nfhs/NFHS-4reports/Karnataka.pdf

[4]https://nhm.karnataka.gov.in/info-2/Demography+and+Evaluation+Cell/Publications/en

[5]https://des.karnataka.gov.in/info-2/Publication+Training+and+Co-ordination+(PTC)/Reports/en

[6]https://ejanma.karnataka.gov.in/Sitemap.aspx

Dashboard[7], India SDG Dashboard[8] and, NIUA SDG Dashboard[9] also act as monitoring and reporting tools to evaluate the status and achievement of various SDGs. We present an intervention dashboard[10] to study the differential impact on different regions of the state based on interventions at multiple levels. Also, a prescriptive modeling approach is followed to prescribe the magnitude of interventions required to lower MD annually. Consequently, the impact and stability of recommendations are also calculated. The dashboard goes beyond reporting and monitoring SDG indicators. It assists policymakers with listing factors deserving the most attention, regions with the highest proportion of beneficiaries, and thus a suitable distribution of the funds. Hence, our work provides actionable insights to policymakers by going beyond predictions, monitoring, and reporting tools.

2 Prescriptive Modeling with Stability Mapping

We divide the methodology into prescriptive modeling with a 'what-to-do' analysis and stability mapping as a subsequent scenario.

2.1 Prescriptive Modeling

To optimize SDG localization strategies, this study takes a prescriptive modeling approach. For a given geographical unit say a, multiple factors $(F_1^a, F_2^a, ..., F_k^a)$ associated with the outcome variable, O^a are evaluated based on the multivariate linear regression model as shown in equation 1:

$$\hat{O}^a = \hat{\beta}_0 + \hat{\beta}_1 F_1^a + \hat{\beta}_2 F_2^a + ... + \hat{\beta}_k F_k^a \tag{1}$$

For SDG localization, it is crucial to identify factors that influence MD and regulate them to mitigate the problem. To approach a given target for MD, equation 2 determines the new value for a factor F_j for district a, (denoted as ${}^{new}F_j^a$), based on the given state target, O_{tar}, and the current value of the indicator, O_{curr}.

$${}^{new}F_j^a = {}^{old}F_j^a + \alpha_j * \left[\frac{(O_{tar} - O_{curr})}{\hat{\beta}_j}\right] \tag{2}$$

where, ${}^{old}F_j^a$ is the original value of a factor. α_j denotes the sensitivity of the factor F_j where $\alpha_j \in [0, 1]$. α_j is external to the model and can be calibrated by the policymaker to adjust the impact of a factor on the outcome variable.

2.2 Stability Mapping

Our previous work [Rachuri *et al.*, 2023] on SDG 2 indicators, demonstrated how a single intervention impacts multiple crops in different Agro-climatic Zones. We proposed a novel approach of *Stress Modeling* to understand the behavior of a system before and after an intervention. We use a modified approach here to infer the impact and stability of interventions.

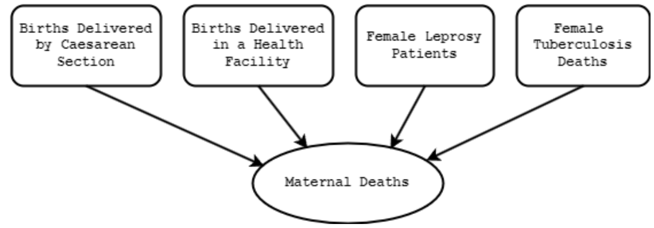

Figure 1: The network showcases the four factors associated with the SDG indicator, Maternal Deaths

The *stress* of a district is the disparity between its impact and the impact of its neighboring districts. Interventions even with a high impact on a given region may not be sustained if the disparity in outcomes is too high. Stable interventions are those that result in minimal stress.

For stability modeling, we prepare an undirected graph $G = (V, E)$, where the nodes V represent the set of administrative regions under consideration. Currently, each node is a district of the state of Karnataka. Edges, (u, v) represent adjacency between two districts. Let the set of neighbors of a given node v be represented as $\Gamma(v)$. In a generic sense, if m indicators were to be impacted with an intervention, each node $v \in V$ is associated with a vector $c(v)$ of m dimensions. Given this, the stability of a node v is given as,

$$stability(v) = 1 - \underbrace{\sum_{\forall u \in \Gamma(v)} L(c(v), c(u))}_{stress} \tag{3}$$

where, $L(.,.)$ is a suitable norm function to calculate the distance between two vectors. Here, we have just one indicator, [MD] associated with each v.

3 Case Study: Intervention Modeling for Maternal Deaths

For MD, we analyzed 115 attributes across all the five datasets mentioned in section 1. Of these, based on consultations with domain experts, we narrowed down to four significantly correlated factors, namely: Births Delivered by a Caesarean Section (CS), Births Delivered in a Health Facility (HF), Female Leprosy Patients (LP), and Female Tuberculosis Deaths (TB). Hence, equation 1 takes the form:

$$\hat{MD}^a = \hat{\beta}_0 + \hat{\beta}_{CS} F_{CS}^a + \hat{\beta}_{HF} F_{HF}^a + \hat{\beta}_{LP} F_{LP}^a + \hat{\beta}_{TB} F_{TB}^a \tag{4}$$

Figure 1 illustrates these associations.

As explained in the demo video[11], we divide our intervention dashboard into four parts to illustrate an efficient and lucid approach to policymakers for the given SDG indicator, Maternal Deaths.

3.1 Differential Impact Analysis

We analyze how the predicted change in a target variable for a district compares with the neighboring districts' change. Thus, by changing factor F_j by $x\%$, the predicted change in the target variable O^a is calculated as,

[7] https://sdgindiaindex.niti.gov.in/urban/
[8] http://www.sdgindia2030.mospi.gov.in/dashboard/
[9] https://niua.in/niua-sdg-dashboard
[10] https://wsl.iiitb.ac.in/intervention-dashboard-maternal-deaths/
[11] https://youtu.be/0Slg39gf1og

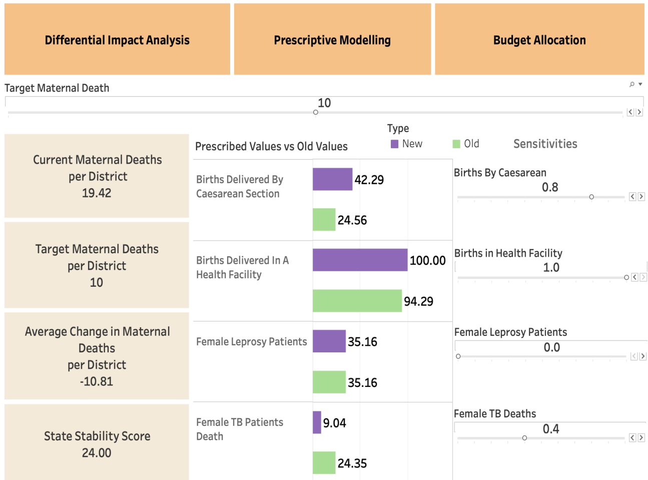

Figure 2: Intervention Dashboard integrated with Differential Impact Analysis, Prescriptive Modeling with Stability Mapping, and Budget Allocation for the given SDG indicator, Maternal Deaths.

$$\Delta O^a = (\frac{x}{100}) \times \hat{\beta}_j \times F_j^a \tag{5}$$

3.2 Prescriptive Analysis

To reduce maternal deaths from the current state average to a chosen target value, the policymaker may choose to set sensitivities for each parameter based on its influence on the indicator. In figure 2, α_{HF} = 1.0, α_{CS} = 0.8, α_{LP} = 0.0 and α_{TB} = 0.4. The resultant bar chart shows the new prescribed values vs. the old values for the factors. We see the biggest drop in female tuberculosis deaths. Hence, TB deaths should decline drastically to lower the overall MD in the state. We also notice that institutional deliveries need to be carried out throughout the state uniformly[12] to increase survival prospects for mother and baby. Thus, with a combination of prescribed interventions, the state average for maternal deaths can be lowered to 10 from 19.42.

3.3 Stability vs. Impact Analysis

Corresponding to the new prescribed values, districts are distributed into high/low impact and stability zones. Figure 3 shows 5 districts that fall into the most desirable zone with high impact and high stability for the new set of interventions. Dharwad (DWD) district shows high impact but relatively low stability. This means DWD may not be sustained in this state for the long term. Thus, policymakers need to engage more with the districts with low stability to carry out sustained improvements toward maternal health and well-being.

3.4 Budget Allocation

The prescribed values from the dashboard aid in efficient budget allocation for the various schemes that will promote

[12]https://hfw.delhi.gov.in/fw/janani-shishu-suraksha-karyakram

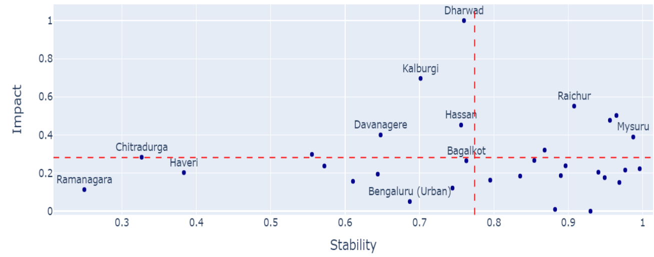

Figure 3: A scatter plot between impact and stability of prescribed interventions for the districts of Karnataka. The red dotted lines indicate average impact (0.28) and average stability (0.77) along the axes.

awareness toward maternity benefit programs. Hence, the amount (A) allocated for a scheme will be,

$$A_{F_j} = \frac{\alpha_j * \frac{1}{|\hat{\beta}_j|}}{\sum_{i=1}^{k} \alpha_i * \frac{1}{|\hat{\beta}_i|}} \tag{6}$$

where k denotes the total number of factors and β_j denotes the coefficient of a factor. From figure 2, schemes related to HF will get the largest share of the budget followed by CS and then TB.

Schemes and Beneficiaries

For n districts, the amount allocated to a district a for a scheme corresponding to factor F_j is calculated as,

$$A^a_{F_j} = A_{F_j} \times \frac{F_j^a}{\sum_{i=1}^{n} F_j^i} \tag{7}$$

where F_j^a is the value of the factor F_j for district a.

The total number of beneficiaries, B, for a district under the scheme corresponding to factor F_j is calculated as,

$$B^a_{F_j} = \frac{A^a_{F_j}}{u_{F_j}} \tag{8}$$

where u_{F_j} is the unit cost of the scheme.

4 Conclusions, Limitations and Future Work

Policymakers need data-driven recommendations to reach expected targets. The intervention dashboard illustrates data stories for strategic decision-making by prescribing the magnitude of interventions for mitigating maternal mortality and thus distributing funds appropriately in relevant schemes. This aids in optimizing region-specific SDG localization strategies where some regions may have disparate responses towards the same set of interventions.

As a public-facing demonstration, we have limited ourselves to using datasets available only at the district level. Previously, multiple dashboards in collaboration with the Planning Department of the Government of Karnataka were developed using proprietary datasets at a more granular level.

Hereafter, we aim to enhance the dashboard with indicators from other domains and pursue a deeper understanding of lateral impact and internal stress among the indicators based on some common interventions.

Ethical Statement

There are no ethical issues. All the datasets used for our study are available in the public domain and free to download from government websites.

Acknowledgments

The authors thank the Planning and Statistics Department, Government of Karnataka, India, and the Machine Intelligence and Robotics (MINRO) Center funded by the Government of Karnataka, India for supporting this work. The authors also like to thank Dr. Mukund Raj from the United Nations Development Programme (UNDP) for his expertise and insightful discussions, and Manish Reddy Koppula for his assistance with modeling and analysis.

References

[Arsenault *et al.*, 2017] Catherine Arsenault, Sam Harper, Arijit Nandi, José M Mendoza Rodríguez, Peter M Hansen, and Mira Johri. An equity dashboard to monitor vaccination coverage. *Bulletin of the World Health Organization*, 95(2):128, 2017.

[Azcuna, 2016] Rosebelle May Azcuna. Monitoring universal health coverage in the western pacific: Framework, indicators, and dashboard. 2016.

[Bassin *et al.*, 2021] Pooja Bassin, Niharika Sri Parasa, Srinath Srinivasa, and Sridhar Mandyam. Big data management for policy support in sustainable development. In *International Conference on Big Data Analytics*, pages 3–15. Springer, 2021.

[Bassin, 2022] Pooja Bassin. Network learning on open data to aid policy making. In Sebastian Link, Iris Reinhartz-Berger, Jelena Zdravkovic, Dominik Bork, and Srinath Srinivasa, editors, *Proceedings of the ER Forum and PhD Symposium 2022 co-located with 41st International Conference on Conceptual Modeling (ER 2022), Virtual Event, Hyderabad, India, October 17, 2022*, volume 3211 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2022.

[Bissio and Watch, 2016] Roberto Bissio and Social Watch. Towards a 2030 agenda dashboard. *Spotlight on Sustainable Development*, page 148, 2016.

[Canepa *et al.*, 2022] Maria Canepa, Adriano Magliocco, and Nicola Pisani. Data visualization and web-based mapping for sgds and adaptation to climate change in the urban environment. In *International Conference on Technological Imagination in the Green and Digital Transition*, pages 715–724. Springer, 2022.

[Garwood *et al.*, 2020] Kathleen Campbell Garwood, David Steingard, and Marcello Balduccini. Dynamic collaborative visualization of the united nations sustainable development goals (sdgs): Creating an sdg dashboard for reporting and best practice sharing. In *VISIGRAPP (3: IVAPP)*, pages 294–300, 2020.

[Oosterhof, 2018] Pytrik Dieuwke Oosterhof. Localizing the sustainable development goals to accelerate implementation of the 2030 agenda for sustainable development: The current state of sustainable development goal localization in asia and the pacific. 2018.

[Rachuri *et al.*, 2023] Sowmith Nandan Rachuri, Arpitha Malavalli, Niharika Sri Parasa, Pooja Bassin, and Srinath Srinivasa. Modeling the impact of policy interventions for sustainable development. In *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, pages 7167–7170, 2023.

[Sachs *et al.*, 2023] Jeffrey D Sachs, Guillaume Lafortune, Grayson Fuller, and Eamon Drumm. Sustainable development report 2023: implementing the sdg stimulus. 2023.

[Zainuddin *et al.*, 2023] Zada Qusyairin Mohd Zainuddin, Farashazillah Yahya, Ervin Gubin Moung, Bashirah Mohd Fazli, and Mohammad Fikry Abdullah. Effective dashboards for urban water security monitoring and evaluation. *International Journal of Electrical & Computer Engineering (2088-8708)*, 13(4), 2023.