

Supplementary Information for

Precision mapping child undernutrition for nearly 600,000 inhabited census villages in India

Rockli Kim, Avleen Bijral, Yun Xu, Xiuyuan Zhang, Jeffrey C. Blossom, Akshay Swaminathan, Gary King, Alok Kumar, Rakesh Sarwal, Juan M. Lavista Ferres, S.V. Subramanian

Correspondence to S.V. Subramanian Email: svsubram@hsph.harvard.edu

This PDF file includes:

Figures S1 to S4 Tables S1 Legend for Dataset S1 to S2

Other supplementary materials for this manuscript include the following:

Datasets S1 to S2

Fig S1. The initial values for villages are obtained by taking the average of the estimates corresponding to the clusters these villages are mapped to

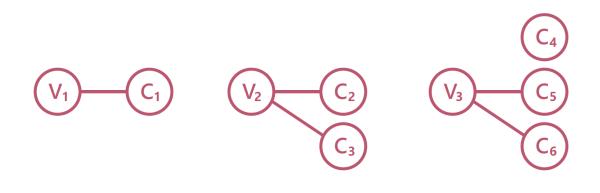


Fig S2. 1-dimensional illustration of how unlabeled data can help in a simple classification problem

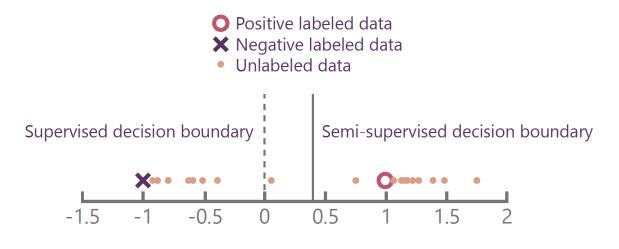
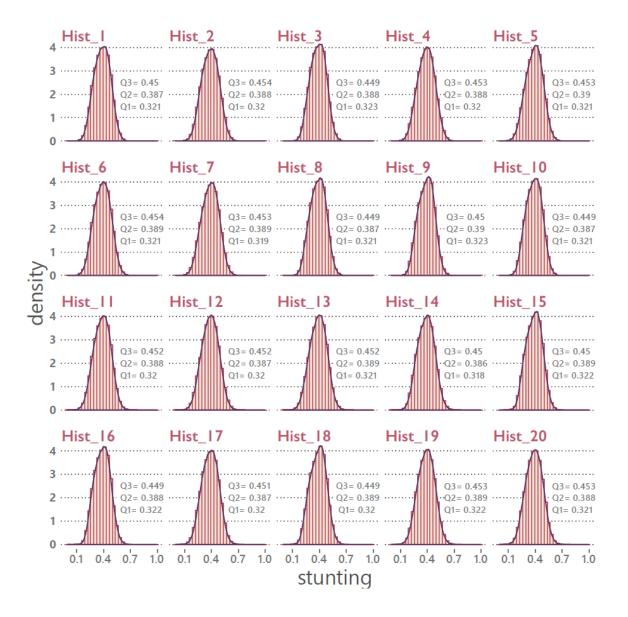
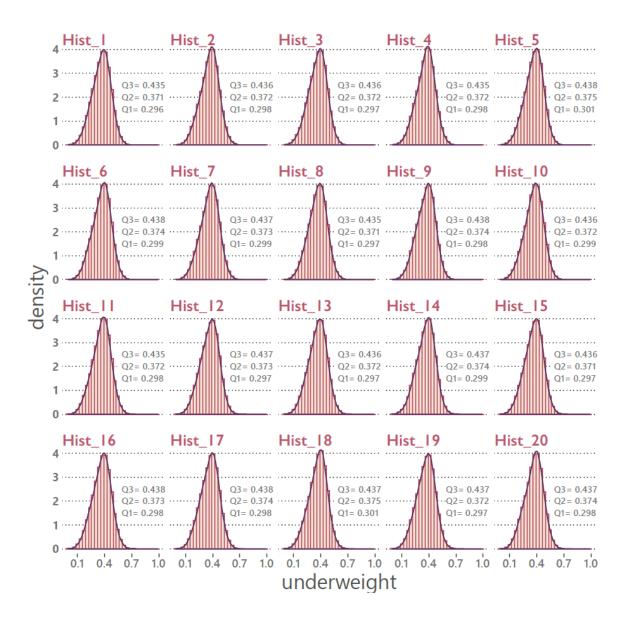


Fig S3. Prediction histograms for 20 randomized initial label assignments for (a) stunting, (b) underweight, and (c) wasting

(a) Stunting



(b) Underweight



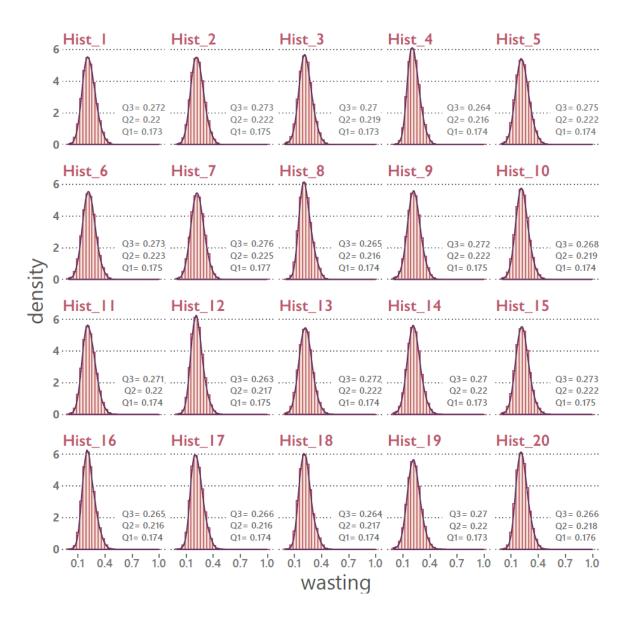
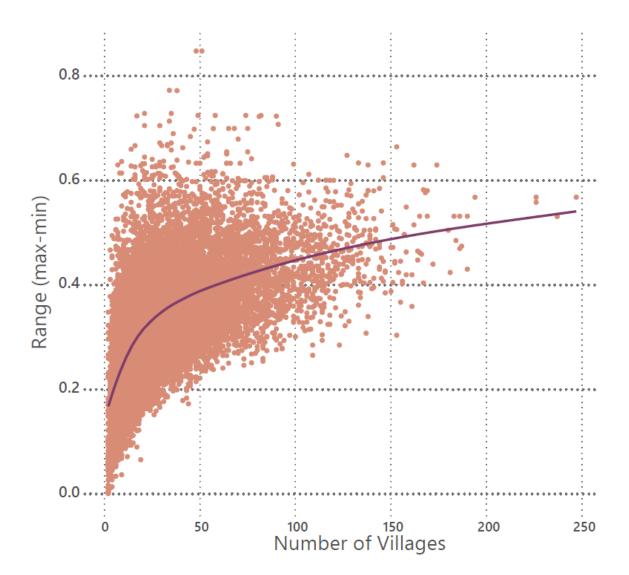


Fig S4. Range variation in (a) stunting, (b) underweight, and (c) wasting as the number of mapped villages increases.

(a) Stunting

Scatterplot

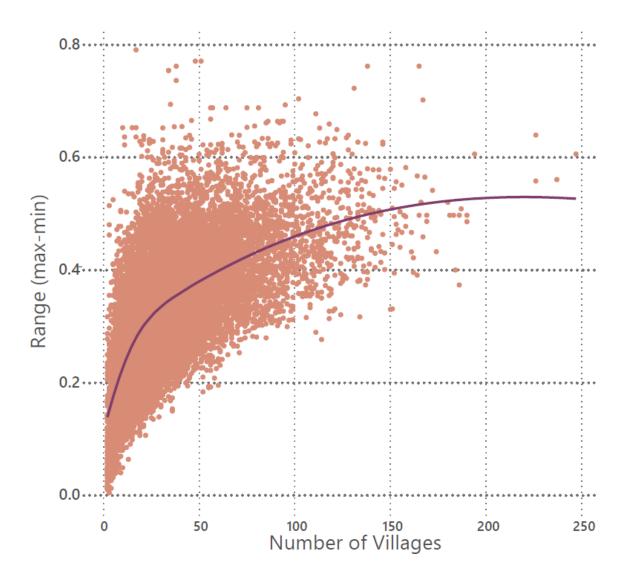
Village Count (per Cluster) / Stunting



(b) Underweight

Scatterplot

Village Count (per Cluster) / Underweight



(c) Wasting

Scatterplot

Village Count (per Cluster) / Wasting

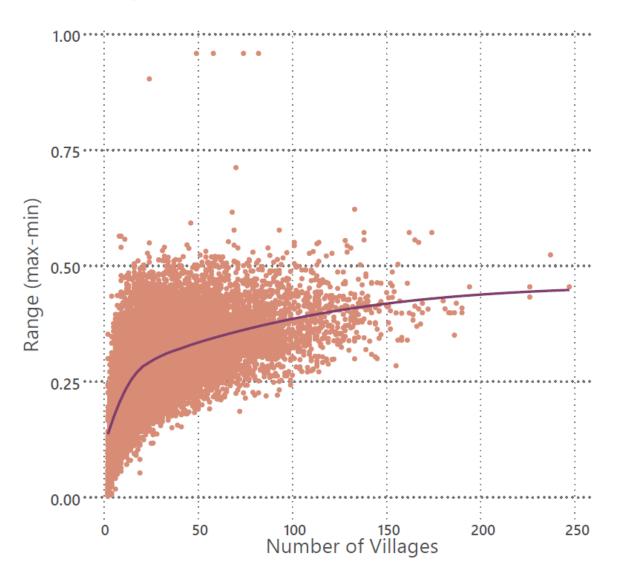


Table S1. List of 2011 Census demographic and amenities features used for prediction models

Variable name	Definition	Note
tot_hh	Total number of household	Raw variable from ML Infomap demographic file
m_mnw	Male Main Workers	Raw variable from ML Infomap demographic file
f_mnw	Female Main Workers	Raw variable from ML Infomap demographic file
m_cult	Male Cultivators	Raw variable from ML Infomap demographic file
f_cult	Female Cultivators	Raw variable from ML Infomap demographic file
m_aglb	Male Agricultural Labourers	Raw variable from ML Infomap demographic file
f_aglb	Female Agricultural Labourers	Raw variable from ML Infomap demographic file
m_mfhh	Male Household industry workers	Raw variable from ML Infomap demographic file
f_mfhh	Female Household industry workers	Raw variable from ML Infomap demographic file
m_oth_w	Male other workers	Raw variable from ML Infomap demographic file
f_oth_w	Female other workers	Raw variable from ML Infomap demographic file
m_mrg_3_6	Male Marginal Workers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
f_mrg_3_6	Female Marginal Workers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
m_m_cl_3_6	Male Marginal Cultivators-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
f_m_cl_3_6	Female Marginal Cultivators- worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
m_mag_3_6	Male Marginal Agricultural Labourers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
f_mag_3_6	Female Marginal Agricultural Labourers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
m_mhh_3_6	Male Marginal Workers in Household Industries-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
f_mhh_3_6	Female Marginal Workers in Household Industries-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
m_mot_3_6	Male Marginal Other Workers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file

f_mot_3_6	Female Marginal Other Workers-worked for 3 months or more but less than 6 months	Raw variable from ML Infomap demographic file
m_mrg_0_3	Male Marginal Workers-worked for less than 3 months	Raw variable from ML Infomap demographic file
f_mrg_0_3	Female Marginal Workers- worked for less than 3 months	Raw variable from ML Infomap demographic file
m_m_cl_0_3	Male Marginal Cultivators-worked for less than 3 months	Raw variable from ML Infomap demographic file
f_m_cl_0_3	Female Marginal Cultivators- worked for less than 3 months	Raw variable from ML Infomap demographic file
m_mag_0_3	Male Marginal Agricultural Labourers- worked for less than 3 months	Raw variable from ML Infomap demographic file
f_mag_0_3	Female Marginal Agricultural Labourers-worked for less than 3 months	Raw variable from ML Infomap demographic file
m_mhh_0_3	Male Marginal Workers in Household Industries-worked for less than 3 months	Raw variable from ML Infomap demographic file
f_mhh_0_3	Female Marginal Workers in Household Industries-worked for less than 3 months	Raw variable from ML Infomap demographic file
m_mot_0_3	Male Marginal Other Workers	Raw variable from ML Infomap demographic file
f_mot_0_3	Female Marginal Other Workers	Raw variable from ML Infomap demographic file
m_nnw	Male Non-workers	Raw variable from ML Infomap demographic file
f_nnw	Female Non-workers	Raw variable from ML Infomap demographic file
population_per_hectare	tot_pop/Total_Geographical_Areain_Hect	Derived variable from ML Infomap demographic file
sex_ratio	f_pop/m_pop	Derived variable from ML Infomap demographic file
child_population	tot_l6/tot_pop	Derived variable from ML Infomap demographic file
child_sex_ratio	f_l6 / m_l6	Derived variable from ML Infomap demographic file
proportion_SC	tot_sc/tot_pop	Derived variable from ML Infomap demographic file
proportion_ST	tot_st/tot_pop	Derived variable from ML Infomap demographic file
proportion_total_literate	tot_lit/(tot_pop-tot_l6)	Derived variable from ML Infomap demographic file
proportion_male_literate	m_lit / (m_pop - m_l6)	Derived variable from ML Infomap demographic file
proportion_female_literate	f_lit / (f_pop - f_l6)	Derived variable from ML Infomap demographic file
proportion_workers	tot_w/(tot_pop - tot_l6)	Derived variable from ML Infomap demographic file
proportion_male_workers	m_w / (m_pop - m_l6)	Derived variable from ML Infomap demographic file
proportion_female_workers	f_w / (f_pop - f_l6)	Derived variable from ML Infomap demographic file
proportion_AGLB	tot_aglb / tot_w	Derived variable from ML Infomap demographic file

proportion_MRW proportion_MRG_3_6 proportion MRG 0 3 District Head Quarter Distance Nearest_Statutory_Town__Distance Total_Geographical_Area_in_Hect Govt_Pre_Primary_School__Nursery Private Pre Primary School Nu Govt Primary School Status Private Primary School Status A Govt Middle School Status Private Middle School Status A Govt_Secondary_School_Status_A_ Private Secondary School Status Govt Senior Secondary School St Private_Senior_Secondary_School Govt_Arts_and_Science_Degree_Co Private Arts and Science Degree Govt Engineering College Status Private Engineering College Sta Govt_Medicine_College_Status_A_ Private_Medicine_College_Statu Govt_Management_Institute_Statu Private Management Institute St Govt Polytechnic Status Private_Polytechnic_Status Govt_Vocational_Training_School_

tot_mrw/tot_w $t_mrg_3_6/tot_w$ t mrg 0.3/tot w District Head Quarter (Distance in km) Nearest Statutory Town (Distance in km) Total Geographical Area (in Hectares) Govt Pre-Primary School (Nursery/LKG/UKG) (Status A(1)/NA(2))Private Pre - Primary School (Nursery/LKG/UKG) (Status A(1)/NA(2))Govt Primary School (Status A(1)/NA(2)) Private Primary School (Status A(1)/NA(2)) Govt Middle School (Status A(1)/NA(2)) Private Middle School (Status A(1)/NA(2)) Govt Secondary School (Status A(1)/NA(2)) Private Secondary School (Status A(1)/NA(2)) Govt Senior Secondary School (Status A(1)/NA(2)) Private Senior Secondary School (Status A(1)/NA(2)) Govt Arts and Science Degree College (Status A(1)/NA(2)) Private Arts and Science Degree College (Status A(1)/NA(2))Govt Engineering College (Status A(1)/NA(2)) Private Engineering College (Status A(1)/NA(2)) Govt Medicine College (Status A(1)/NA(2)) Private Medicine College (Status A(1)/NA(2)) Govt Management Institute (Status A(1)/NA(2)) Private Management Institute (Status A(1)/NA(2)) Govt Polytechnic (Status A(1)/NA(2)) Private Polytechnic (Status A(1)/NA(2)) Govt Vocational Training School/ITI (Status A(1)/NA(2))

Derived variable fromML Infomap demographic file Derived variable fromML Infomap demographic file Derived variable fromML Infomap demographic file Raw variable from village amenities file Raw variable from village amenities file Raw variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Private Vocational Training Scho Government Non Formal Training C Private_Non_Formal_Training_Cent Government_School_For_Disabled___ Private School For Disabled St Government Others Status Private_Others_Status Community_Health_Centre_Status Primary Health Centre Status Primary Health Sub Centre Status Maternity Child Center Status **TB** Clinic Status Hospital_Allopathic_Status Hospital Alternative Status Dispensary Status Veterinary_Hospital_Status Mobile_Health_Clinic_Status Family Welfare Center Status Non_Gov_Medical_Out_Status Non_Gov_Medical_InOut_Status Non Gov Medical Char Status Non Gov Medical MBBS Status Non Gov Medical OtherDeg Status Non Gov Medical NoDeg Status

Private Vocational Training School/ITI (Status A(1)/NA(2))Government Non Formal Training Centre (Status A(1)/NA(2)) Private Non Formal Training Centre (Status A(1)/NA(2)) Government School For Disabled (Status A(1)/NA(2)) Private School For Disabled (Status A(1)/NA(2)) Government Others (Status A(1)/NA(2)) Private Others (Status A(1)/NA(2)) Community_Health_Centre (Status A(1)/NA(2)) Primary Health Centre (Status A(1)/NA(2)) PrimaryHealth Sub Centre (Status A(1)/NA(2)) Maternity And Child Welfare Centre(Status A(1)/NA(2)) TB Clinic(Status A(1)/NA(2)) Hospital Allopathic (Status A(1)/NA(2)) Hospital Alternative Medicine(Status A(1)/NA(2)) Dispensary(Status A(1)/NA(2)) Veterinary Hospital(Status A(1)/NA(2)) Mobile Health Clinic(Status A(1)/NA(2)) Family Welfare Centre(Status A(1)/NA(2)) Non Government Medical facilities Out Patient(Status A(1)/NA(2))Non Government Medical facilities In and Out Patient(Status A(1)/NA(2)) Non Government Medical facilities Charitable(Status A(1)/NA(2))Non Government Medical facilities Medical Practitioner with MBBS Degree(Status A(1)/NA(2)) Non Government Medical facilities Medical Practitioner with Other Degree (Status A(1)/NA(2)) Non Government Medical facilities Medical Practitioner with No Degree(Status A(1)/NA(2))

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file

Non Gov Medical Faith Status Non Gov Medical Shop Status Non Gov Medical Others Status Tap Water Treated Status Tap_Water_Untreated_Status Covered Well Status Uncovered Well Status Hand_Pump_Functioning_Status Tube_Wells_Borehole_Status Spring Functioning Status River Canal Functioning Status Tank Pond Lake Status Other_water_Status Closed_Drainage_Status Open Drainage Status No Drainage Status Open_Pucca_Drainage_Covered_with Open Pucca Drainage Uncovered S Open_Kuccha_Drainage_Status_A_1 Whether Drain water is discharge TSC coverage Community Toilet Complex includ Community Toilet Complex exclud Rural Production Centres or Sani

Non Government Medical facilities Traditional Pracitioner and Faith Healer(Status A(1)/NA(2)) Non Government Medical facilities Medicine Shop(Status A(1)/NA(2))Non Government Medical facilities Other(Status A(1)/NA(2))Tap Water-Treated (Status A(1)/NA(2)) Tap Water Untreated (Status A(1)/NA(2)) Covered Well (Status A(1)/NA(2)) Uncovered Well (Status A(1)/NA(2)) Hand Pump (Status A(1)/NA(2)) Tube Wells/Borehole (Status A(1)/NA(2)) Spring (Status A(1)/NA(2)) River/Canal (Status A(1)/NA(2)) Tank/Pond/Lake (Status A(1)/NA(2)) Others (Status A(1)/NA(2)) Closed Drainage (Status A(1)/NA(2)) Open Drainage (Status A(1)/NA(2)) No Drainage (Status A(1)/NA(2)) Open Pucca Drainage Covered with Tiles Slabs (Status A(1)/NA(2))Open Pucca Drainage Uncovered (Status A(1)/NA(2)) Open KucchaDrainage(Status A(1)/NA(2)) Whether Drain water is discharged directly into water bodies or to sewar plant (Is the Area Covered under Total Sanitation Campaign (TSC)? (Status A(1)/NA(2)) Community Toilet Complex (including Bath) for General Public (Status A(1)/NA(2)) Community Toilet Complex (excluding Bath) for General Public (Status A(1)/NA(2)) Rural Production Centres or Sanitary hardware outlet availability near the village (Status A(1)/NA(2))

Derived variable from village amenities file Derived variable from village amenities file

Rural Production mart or Sanitar Community waste disposal system Community Bio gas or recycle of No_System_Garbage_on_road_stree Post Office Status Sub Post Office Status Post And Telegraph Office Statu Village Pin Code Status Telephone_Landlines_Status PCO Mobile Status Mobile Phone Coverage Status A Internet Cafes Common Service Private Courier Facility Status Public_Bus_Service_Status Private_Bus_Service_Status Railway Station Status Auto Status Taxi Status Vans_Status Tractors_Status Cycle_pulled_Rickshaws____manual Cycle pulled Rickshaws machine Carts Drivens by Animals Status Sea River Ferry Service Status National_Highway_Status State_Highway_Status

Rural Production mart or Sanitary hardware outlet availability near the village (Status A(1)/NA(2)) Community waste disposal system after house to house collection (Status A(1)/NA(2)) Community Bio-gas or recycle of waste for production use (Status A(1)/NA(2))No System(Garbage on road/street) (Status A(1)/NA(2)) Post Office (Status A(1)/NA(2)) Sub Post Office (Status A(1)/NA(2)) Post And Telegraph Office (Status A(1)/NA(2)) Village Pin Code (Status A(1)/NA(2)) Telephone (landlines) (Status A(1)/NA(2)) Public Call Office /Mobile (PCO) (Status A(1)/NA(2)) Mobile Phone Coverage (Status A(1)/NA(2)) Internet Cafes / Common Service Centre (CSC) (Status A(1)/NA(2))Private Courier Facility (Status A(1)/NA(2)) Public Bus Service (Status A(1)/NA(2)) Private Bus Service (Status A(1)/NA(2)) Railway Station (Status A(1)/NA(2)) Auto/Modified Autos (Status A(1)/NA(2)) Taxi (Status A(1)/NA(2)) Vans (Status A(1)/NA(2)) Tractors (Status A(1)/NA(2)) Cycle-pulled Rickshaws (manual driven) (Status A(1)/NA(2))Cycle-pulled Rickshaws (machine driven) (Status A(1)/NA(2)Carts Drivens by Animals (Status A(1)/NA(2)) Sea/River/Ferry Service (Status A(1)/NA(2)) National Highway (Status A(1)/NA(2)) State Highway (Status A(1)/NA(2))

Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file MajorDistrict_Highway_Status OtherDistrict_Highway_Status Pucca Road Status Gravel Road Status Water_Bounded_Macadam__Status_A AllWeather Road Status Navigable_Waterways__River_Canal Footpath Status ATM Status Commercial Bank Status Cooperative_Bank_Status Agricultural_Credit_Societies_S Self Help Group Status Public Distribution System PDS Regular Market Status Weekly_Haat_Status Agricultural_Marketing_Society___ Nutritional Centres ICDS Status Nutritional Centres Anganwadi Ce Nutritional_Centres_Others_Stat ASHA_Status Community_Centre_with_without_TV Sports Field Status Sports_Club_Recreation_Centre__S Cinema_Video_Hall Public_library_status Public reading room Daily Newspaper Supply Status A

Major District Road (Status A(1)/NA(2)) Other District Road (Status A(1)/NA(2)) Black Topped (pucca) Road (Status A(1)/NA(2)) Gravel (kuchha) Roads (Status A(1)/NA(2)) Water Bounded Macadam (Status A(1)/NA(2)) All Weather Road (Status A(1)/NA(2)) Navigable Waterways (River/Canal) (Status A(1)/NA(2)) Foothpath (Status A(1)/NA(2)) ATM (Status A(1)/NA(2)) Commercial Bank (Status A(1)/NA(2)) Cooperative Bank (Status A(1)/NA(2)) Agricultural Credit Societies (Status A(1)/NA(2)) Self - Help Group (SHG) (Status A(1)/NA(2)) Public Distribution System(PDS) Shop (Status A(1)/NA(2))Mandis/Regular Market (Status A(1)/NA(2)) Weekly Haat (Status A(1)/NA(2)) Agricultural Marketing Society (Status A(1)/NA(2)) Nutritional Centres-ICDS (Status A(1)/NA(2)) Nutritional Centres-Anganwadi Centre (Status A(1)/NA(2)) Nutritional Centres-Others (Status A(1)/NA(2)) ASHA (Status A(1)/NA(2)) Community Centre with/without TV(Status A(1)/NA(2)) Sports Field (Status A(1)/NA(2)) Sports Club/Recreation Centre (Status A(1)/NA(2)) Cinema/Video Hall (Status A(1)/NA(2)) Public Library (Status A(1)/NA(2)) Public Reading Room(Status A(1)/NA(2)) Daily Newspaper Supply (Status A(1)/NA(2))

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file Derived variable from village amenities file

Assembly_Polling_StationStatus	As sembly Polling Station (Status A(1)/NA(2))	Derived variable from village amenities file
Birth_and_Death_Registration_Off	Birth and Death Registration Office (Status A(1)/NA(2))	Derived variable from village amenities file
Power_Supply_For_Domestic_Use	Power Supply For Domestic Use (Status A(1)/NA(2))	Derived variable from village amenities file
Power_Supply_For_Agriculture_Use	Power Supply For Agriculture Use(Status A(1)/NA(2))	Derived variable from village amenities file
Power_Supply_For_Commercial_Use	Power Supply For Commercial Use (Status A(1)/NA(2))	Derived variable from village amenities file
Power_Supply_For_All_Users	Power Supply For All Users (Status A(1)/NA(2))	Derived variable from village amenities file
Forest_Area_in_Hectares_	Forest Area (in Hectares)	Raw variable from village amenities file
Area_under_Non_Agricultural_Uses	Area under Non-Agricultural Uses (in Hectares)	Raw variable from village amenities file
BarrenUn_cultivable_Land_Area	Barren & Un-cultivable Land Area (in Hectares)	Raw variable from village amenities file
Permanent_Pastures_and_Other_Gra	Permanent Pastures and Other Grazing Land Area (in Hectares)	Raw variable from village amenities file
Land_Under_Miscellaneous_Tree_Cr	Land Under Miscellaneous Tree Crops etc. Area (in Hectares)	Raw variable from village amenities file
Culturable_Waste_Land_Areain_H	Culturable Waste Land Area (in Hectares)	Raw variable from village amenities file
Fallows_Land_other_than_Current	Fallows Land other than Current Fallows Area (in Hectares)	Raw variable from village amenities file
Current_Fallows_Areain_Hectare	Current Fallows Area (in Hectares)	Raw variable from village amenities file
Net_Area_Sownin_Hectares_	Net Area Sown (in Hectares)	Raw variable from village amenities file
Total_Unirrigated_Land_Areain	Total Unirrigated Land Area (in Hectares)	Raw variable from village amenities file
Area_Irrigated_by_Sourcein_Hec	Area Irrigated by Source (in Hectares)	Raw variable from village amenities file
Canals_Areain_Hectares_	Canals Area (in Hectares)	Raw variable from village amenities file
Wells_Tube_Wells_Areain_Hectar	Wells/TubeWells Area (in Hectares)	Raw variable from village amenities file
Tanks_Lakes_Area_Hectares	Tanks/Lakes Area (in Hectares)	Raw variable from village amenities file
Waterfall_Areain_Hectares_	Waterfall Area (in Hectares)	Raw variable from village amenities file
Other_Area	Other Source (specify) Area (in Hectares)	Raw variable from village amenities file
Nearest_Town_Name	Nearest Town Name	Raw variable from village amenities file
Nearest_Town_Distance_from_Villa	Nearest Town Distance from Village (in Km.)	Raw variable from village amenities file

Dataset S1. (separate file)

Predicted estimates of child anthropometric failures for 597,121 villages

Dataset S2. (separate file)

Analytic codes used for prediction modeling