

Supplementary Materials for

AGPC: An Annual 500 m Grided Population (1990–2020) for China

Incorporating 3D Building Volume Dynamics

Xiaocong Xu^{1,2}, Shiyu He¹, Jinpei Ou¹, Yan Zhou¹, Xiaoping Liu^{1, 2}

¹Guangdong Key Laboratory for Urbanization and Geo-simulation, School of Geography and Planning, Sun Yat-sen University, Guangzhou, 510275, China

²Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519082, China

Correspondence to: Jinpei Ou (oujinpei3@mail.sysu.edu.cn) and Yan Zhou (zhouy825@mail.sysu.edu.cn)

Contents in this file:

Table S1–S2

Fig S1–S3

Table S1 Comparison among three machine learning approaches

Year	Model	RMSE (person/km ²)		R		rRMSE (%)	
		Training	Validation	Training	Validation	Training	Validation
2010	XGBoost	1086.46	1610.11	0.9590	0.8915	0.12	0.21
	Random Forest	1419.70	1976.61	0.9496	0.8631	0.15	0.23
	SVR	1368.30	1669.778	0.9210	0.8769	0.15	0.19
2020	XGBoost	812.86	1305.03	0.9656	0.9022	0.10	0.16
	Random Forest	961.54	1327.51	0.9580	0.8892	0.12	0.16
	SVR	847.91	1691.96	0.9572	0.8793	0.11	0.21

Table S2 Summary of multi-temporal evaluations of grid population estimates at city-level

Year	R	RMSE ($\times 10^5$ persons)	rRMSE (%)
1990	0.7844	13.67990	21.6
1995	0.8077	13.77777	21.3
2000	0.9102	11.75177	14.6
2005	0.9143	11.64968	13.9
2010	0.9967	2.07928	2.4
2015	0.9676	7.54615	8.0
2020	0.9992	1.40618	1.5

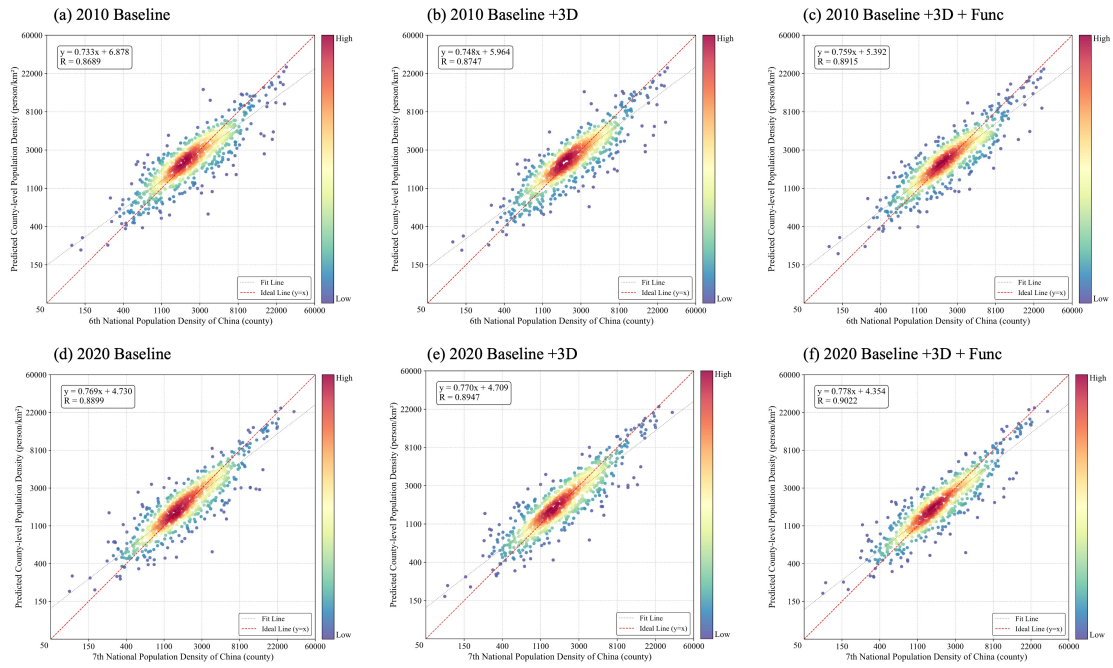


Fig. S1 Model comparisons when incorporating 3D urban environment and building function variables

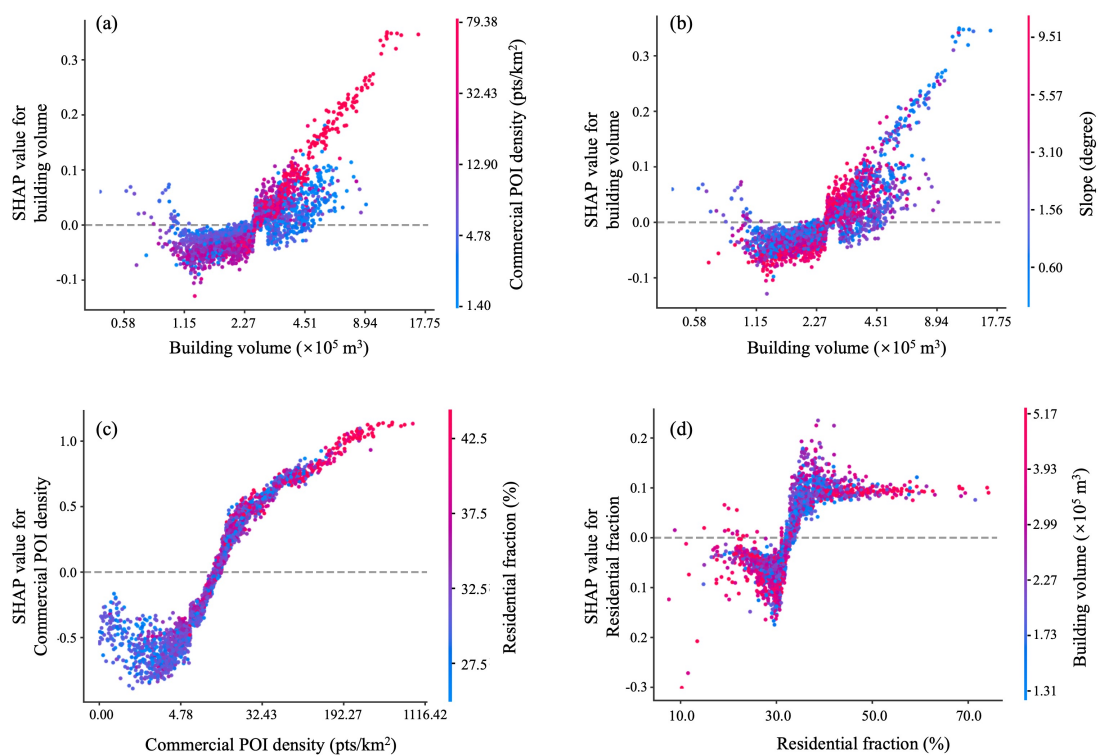
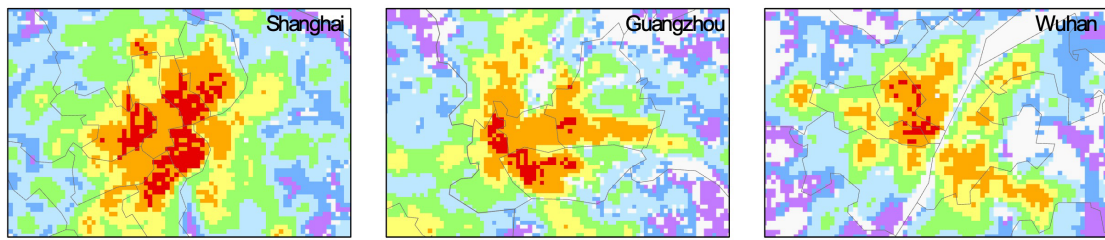
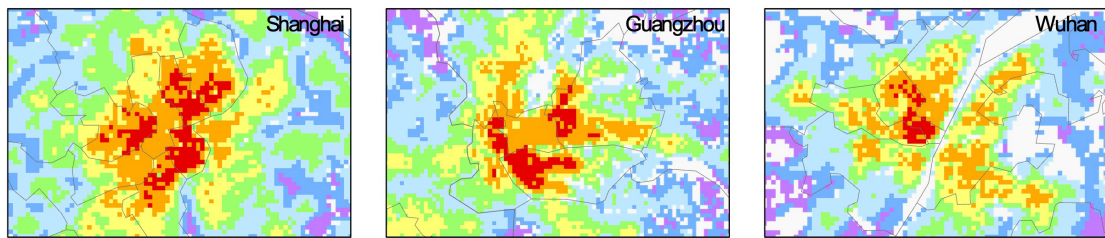


Fig. S2 Dependency plots of covariates identified by the SHAP summary analysis for 2020

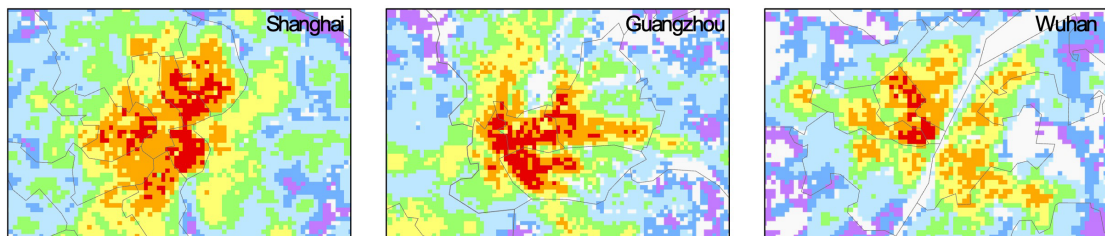
(a) Baseline



(b) Baseline + 3D



(c) Baseline + 3D + Func



0 15 30 km

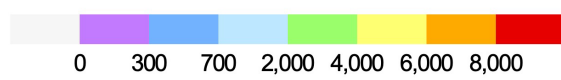


Fig. S3 Comparison of when incorporating 3D urban environment and building function variables