Supporting information for "Improved global daily nitrogen dioxide concentrations from 2005 to 2023 derived using a deep learning approach"

Jiangshan Mu[1,2#], Chenliang Tao[3,1#], Yuqiang Zhang[1*], Zhou Liu[1], Yingnan Zhang[4], Na Zhao[1], Bin Luo[1], Qionghui Zhou[1], Qingzhu Zhang[1], Hongliang Zhang[3], Likun Xue[1*]

[1]Environment Research Institute, Shandong University, 266237, Qingdao, China

[2]Nicholas School of the Environment, Duke University, 27708, Durham, USA

[3]Department of Environmental Science and Engineering, Fudan University, 200233, Shanghai, China

[4]Department of Earth System Science, University of California, Irvine, 92697, California, USA

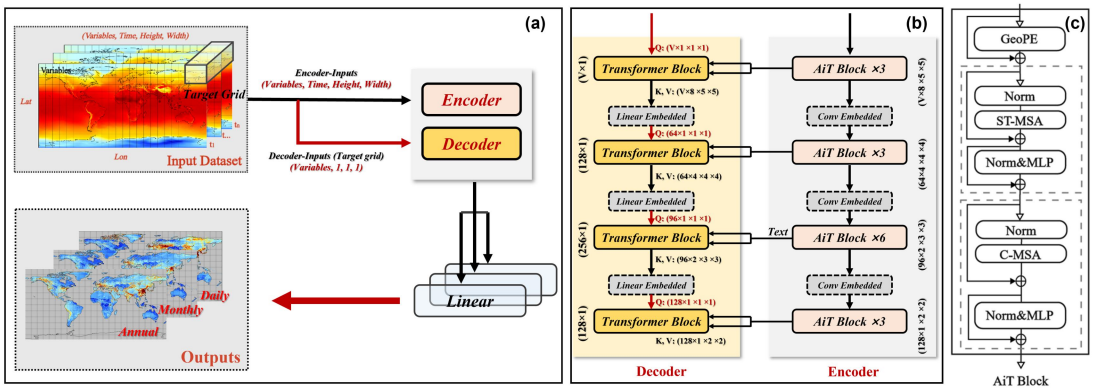

Figure S1: Sequential overview of Full Model Workflow for global surface NO_2 estimations: (a), schematic diagram of Air Transformer (AiT) model; (b), the structure of the Decoder and Encoder in the model; (c), the structure of the AiT Block. GeoPE, Norm, MLP, ST-MSA, and C-MSA respectively indicate positional embedding, layer normalization, multi-layer perceptron, spatial–temporal multi-head self-attention, and multi-channel (multi-variables) multi-head self-attention.

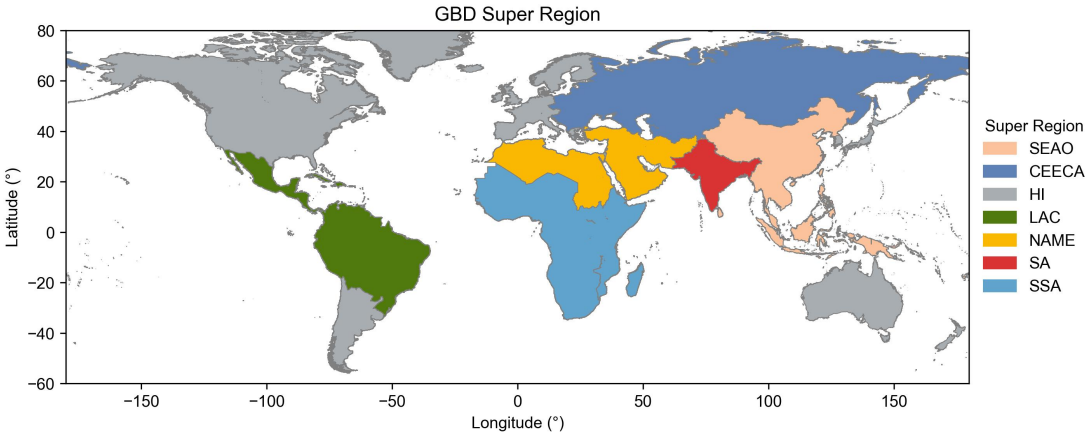

Figure S2. GBD Super-regions based on epidemiological similarity and geographic closeness. The abbreviations are as follows: South-East Asia, East Asia & Oceania (SEAO), Central Europe, Eastern Europe & Central Asia (CEECA), High-income (HI), Latin America & Caribbean (LAC), North Africa & Middle East (NAME), South Asia (SA), and Sub-Saharan Africa (SSA).

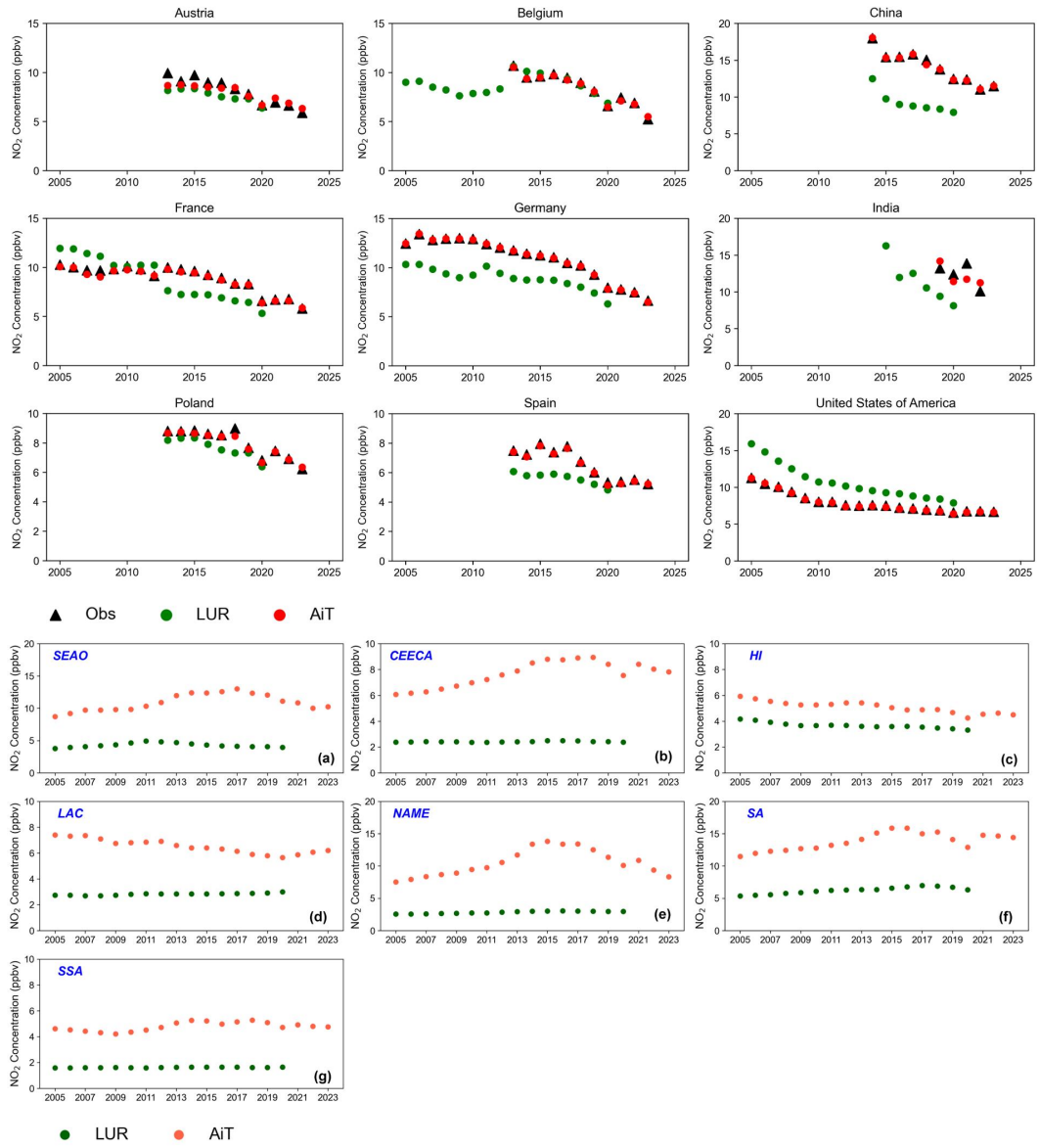

Figure S3. Comparison of NO_2 concentrations between LUR and AiT. top panel for comparisons from 1990 to 2023 across nine countries: Austria, Belgium, China, France, Germany, India, Poland, Spain, and the United States of America, including surface observation (black triangles), LUR (green dots), and AiT (red dots); bottom panel shows the comparisons between LUR and AiT for the seven super regions.

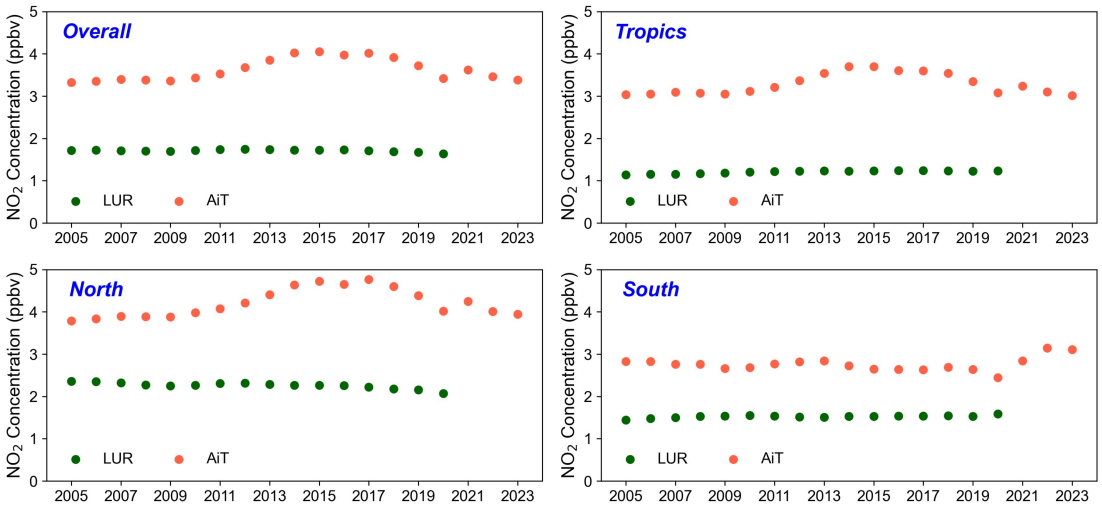

Figure S4: Comparison of area-weighted NO_2 concentrations estimated by LUR and AiT methods across different

global regions (a, global; b, the Tropics; c, Northern Hemisphere; d, Southern Hemisphere) during 2005-2023.

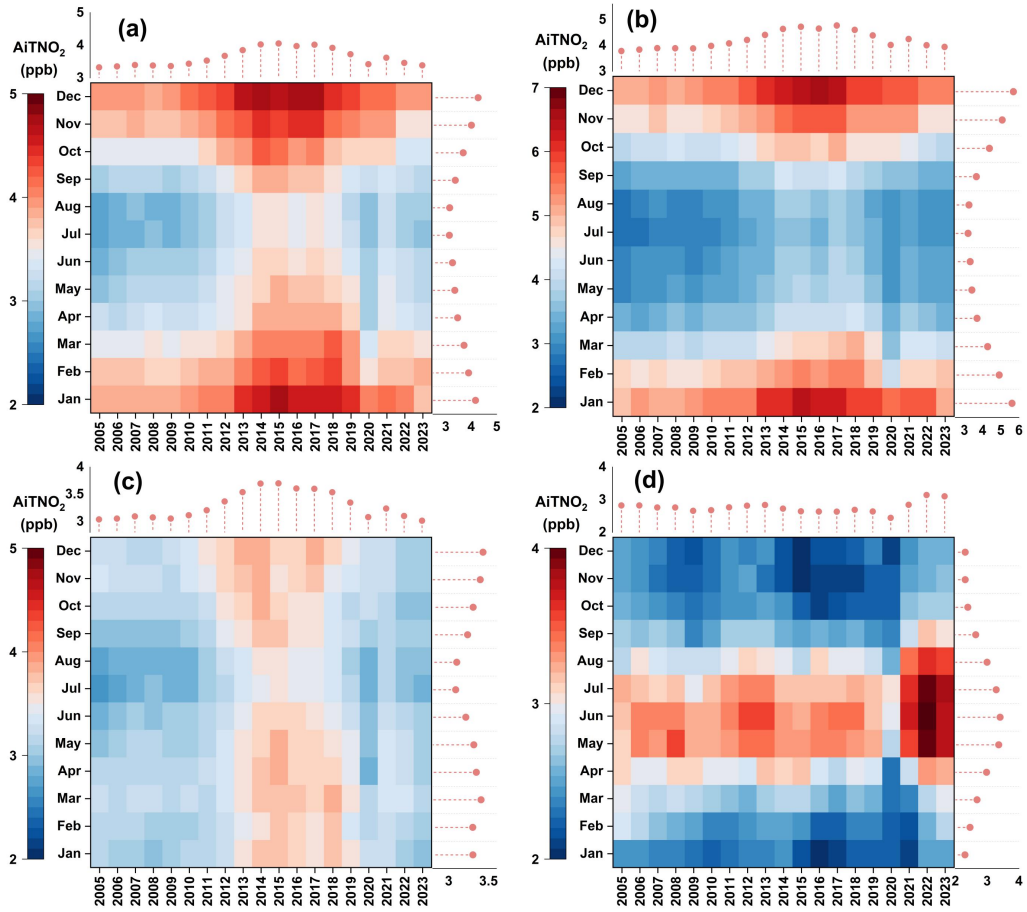

Figure S5: Temporal variations in area-weighted average $AiTNO_2$ concentrations (2005-2023) with heat maps representing monthly data (x-axis: year, y-axis: month), annual averages displayed as dot plots above, and monthly averages displayed as dot plots on the right for global (a), the Northern Hemisphere (30°N to 60°N) (b), the tropics (30°N to 30°S) (c), and the Southern Hemisphere (30°S to 60°S) (d).

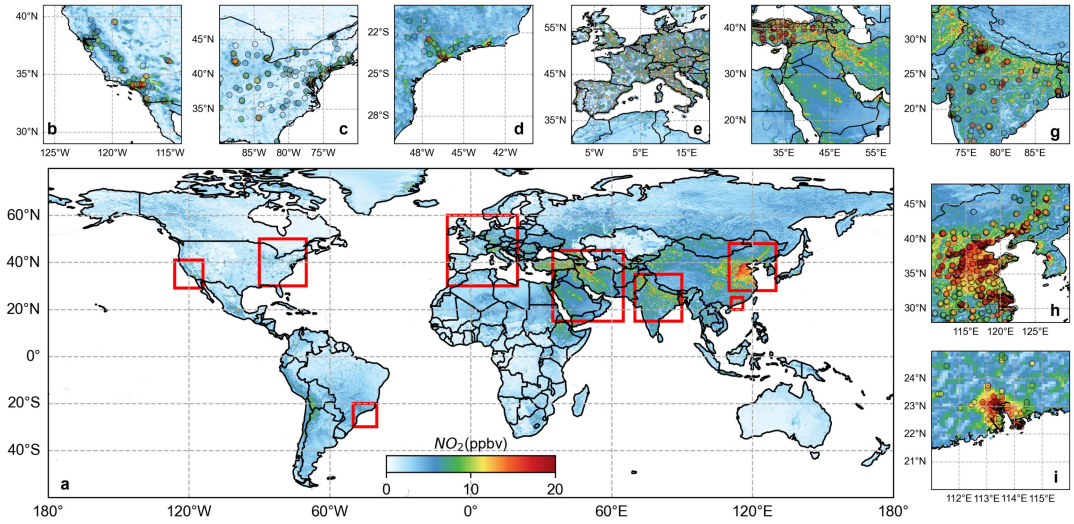

Figure S6: a: 0.1 degree gapless NO_2 global map based on annual average concentration in 2020 (unit: ppbv). Zoomed-in maps (outlined by red rectangles) showing AiT-NO_2 and ground-based NO_2 measurements (colored dots) over the b: western United States, c: eastern United States, d: central South America, e: Europe, f: Arabian Peninsula, g: India, h: eastern China, and i: southern China.

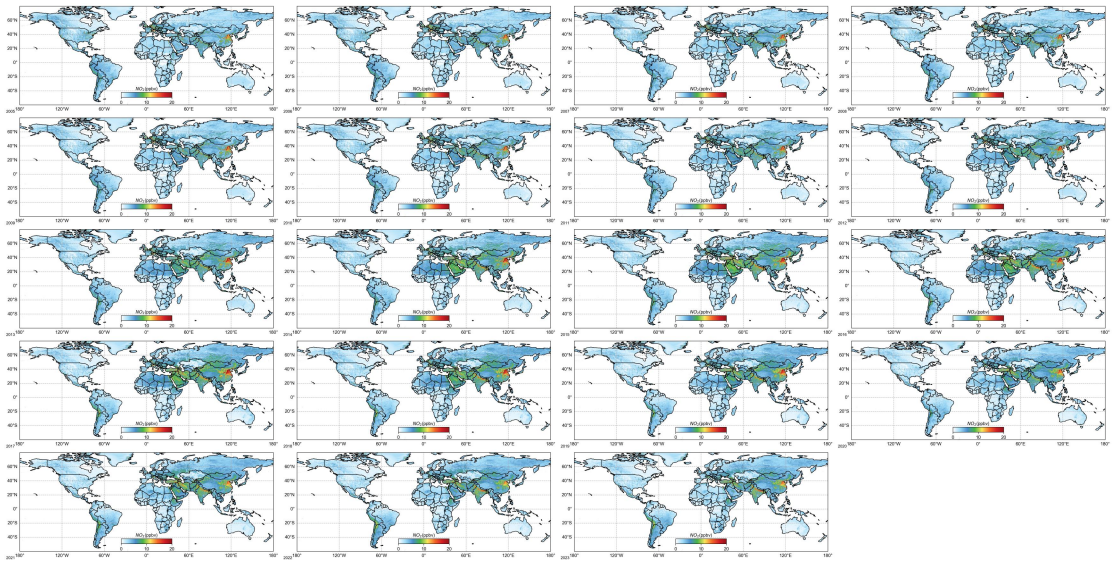

Figure S7: Annual global NO_2 concentration (AiTNO2) at 0.1° × 0.1° from 2005 to 2023. Each subfigure represents the data for a specific year, arranged from left to right and top to bottom, starting with 2005 (top left) and ending with 2023 (bottom right).

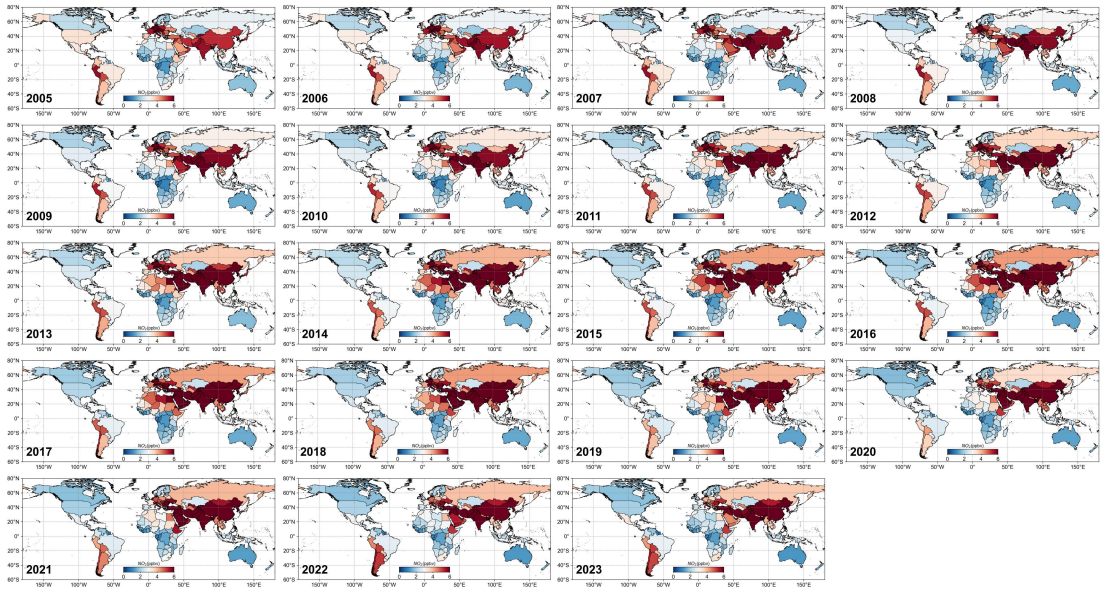

Figure S8: Annual area-weighted average NO_2 concentration data ($AiTNO_2$) by country from 2005 to 2023. Each subfigure represents the data for a specific year, arranged from left to right and top to bottom, starting with 2005 (top left) and ending with 2023 (bottom right).

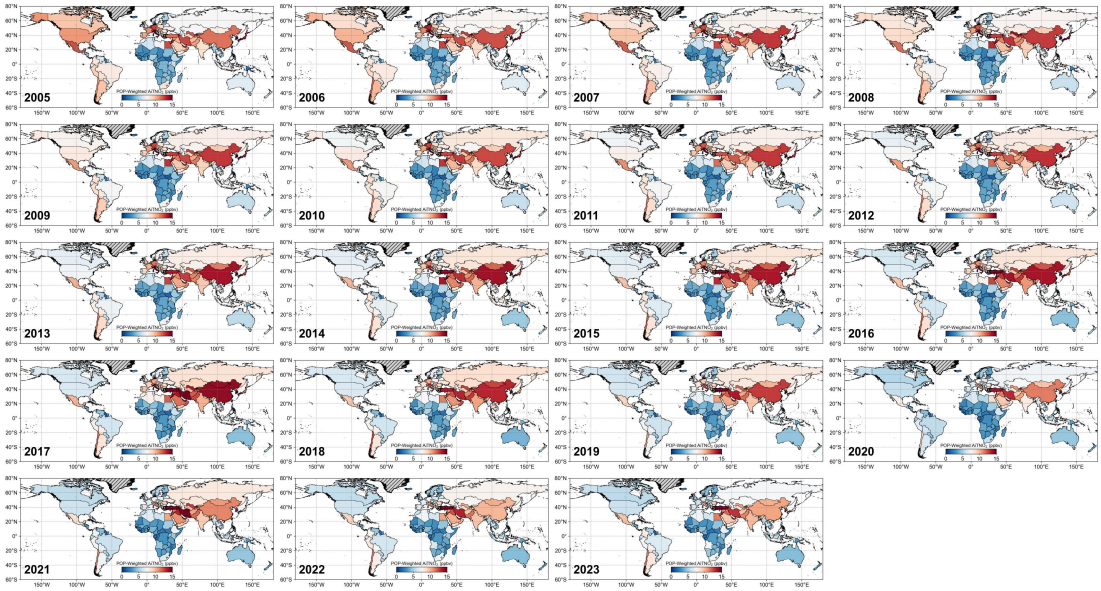

Figure S9: Annual population-weighted average NO_2 concentration by country from 2005 to 2023. Each subfigure represents the data for a specific year, arranged from left to right and top to bottom, starting with 2005 (top left) and ending with 2023 (bottom right).

Table S1. The regional NO_2 concentration underestimation through LUR method compared with the AiT method from 2005 to 2020.

Year	SEAO	CEECA	HI	LAC	NAME	SA	SSA
2005	-56.7%	-60.8%	-29.5%	-63.0%	-66.1%	-53.4%	-65.5%
2006	-57.0%	-61.2%	-28.9%	-62.7%	-67.5%	-54.3%	-64.8%
2007	-58.3%	-61.3%	-29.2%	-63.4%	-68.9%	-54.7%	-63.9%
2008	-56.8%	-62.9%	-29.9%	-62.0%	-69.5%	-53.6%	-63.1%
2009	-55.8%	-64.1%	-30.4%	-59.5%	-69.8%	-53.6%	-61.9%
2010	-52.9%	-66.0%	-30.6%	-58.7%	-71.2%	-52.5%	-63.2%
2011	-52.3%	-67.4%	-30.5%	-58.3%	-71.9%	-52.8%	-64.7%
2012	-55.7%	-68.5%	-32.3%	-58.9%	-72.9%	-53.5%	-65.7%
2013	-60.7%	-69.5%	-33.5%	-56.9%	-75.0%	-55.2%	-67.9%
2014	-63.8%	-71.5%	-31.9%	-55.6%	-77.7%	-58.0%	-68.9%
2015	-65.2%	-71.7%	-29.0%	-55.7%	-78.0%	-58.5%	-68.5%
2016	-66.7%	-71.4%	-26.1%	-54.7%	-77.1%	-57.3%	-66.9%
2017	-68.3%	-72.2%	-27.5%	-53.4%	-77.5%	-53.4%	-68.2%
2018	-67.0%	-72.9%	-29.2%	-51.2%	-76.2%	-54.8%	-69.3%
2019	-66.4%	-71.2%	-26.9%	-49.6%	-73.9%	-52.4%	-68.4%
2020	-64.5%	-68.6%	-21.9%	-46.9%	-70.7%	-51.0%	-65.2%
Average	-60.5%	-67.6%	-29.2%	-56.9%	-72.7%	-54.3%	-66.0%