Pathways to Sustainability: Carbon-Aware Routing for Global AI Data Transfers

Nikolas EJ Schmitz, Dayana Savostianova, Leon Niggemeier, Martin Strauch, Peter Boor

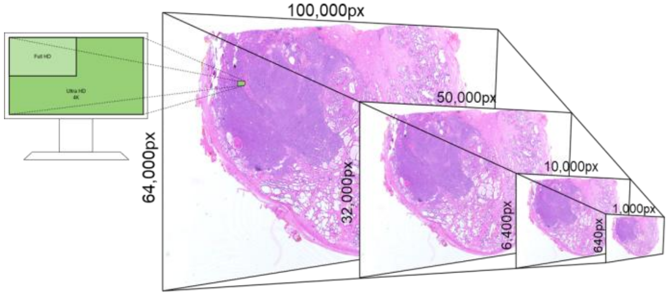
Institute of Pathology
RWTH Aachen University Hospital
Germany

NEURAL INFORMATION PROCESSING SYSTEMS

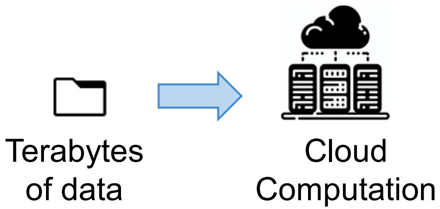
LaBooratory of Nephropathology | UNIKLINIK RWTHAACHEN

Motivation

- **Background**
  - Processing WSIs in digital pathology requires **powerful hardware**
  - Resulting in considerable **carbon dioxide equivalent (CO_2eq) emissions**[1]
  - Transferring datasets to remote computation sites may further increase emissions
- **Research Question**
  - How can we reduce carbon emissions (CE) of data transmissions?
  - What is the optimal time of transmission to reduce CE?



Whole Slide Image (WSI)

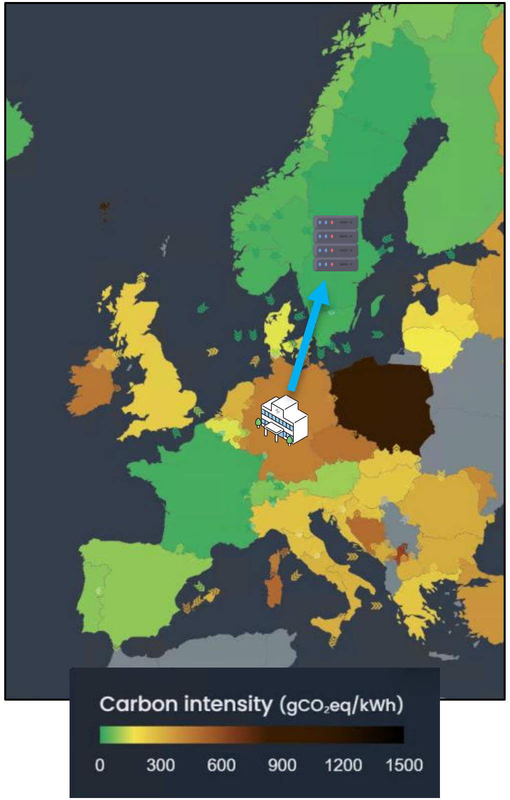


[1]Sadr et al.; Operational greenhouse-gas emissions of deep learning in digital pathology: a modelling study; Lancet DH, 2024

Carbon Intensity

- Carbon emissions associated with energy generation depend on a country's energy mix.
- *Example*: Renewable sources (wind, solar) → **Low Carbon Intensity (CI)**

➔ **Optimizing computation workloads** by allocating them to cloud data centers in **low-CI countries** can significantly reduce CO_2eq emissions.



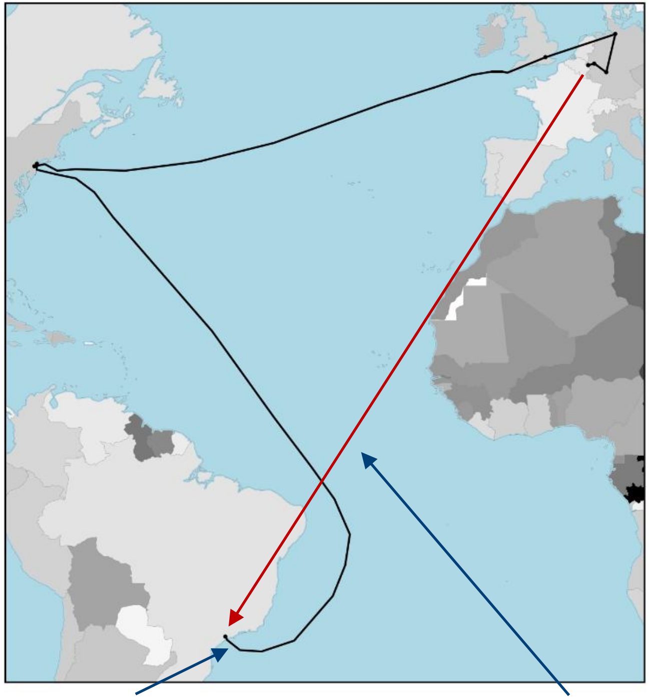
app.electricitymaps.com/map/

Network Routing

- Internet data is sent as IP packets, forwarded router by router toward the destination.
- Packets don't know the full path in advance; each router decides the next hop.

In this work

1. PerfOps'[1] MTR (MyTraceRoute) was used to trace sequence of hops between client and server
   - Hops were geolocated by IP and serve as anchor points
   - Provides closest real-path approximation
2. Network graph built from ITU Broadband Map[2] and Submarine Cable Map[3] data

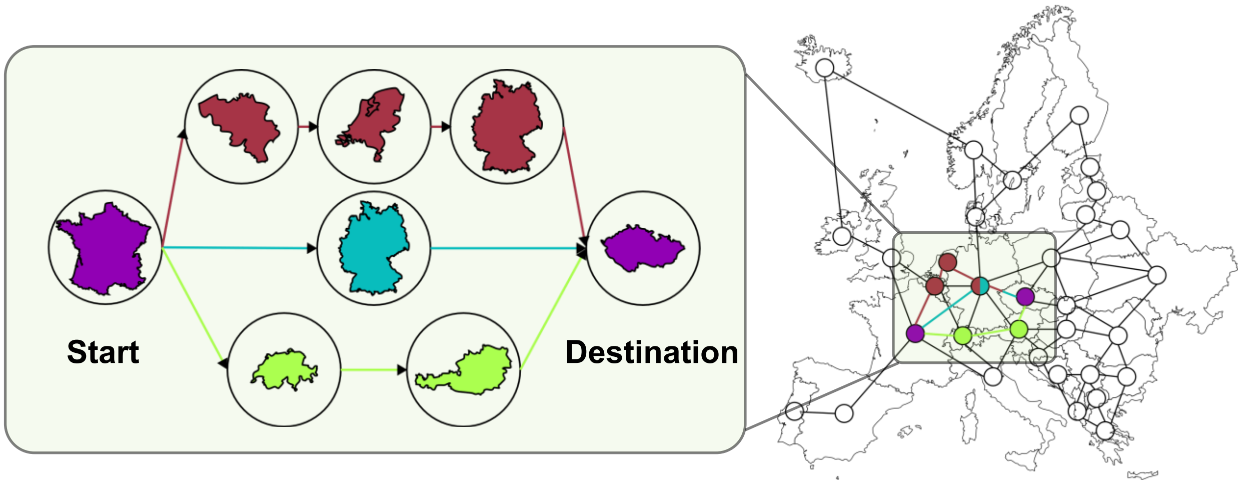


Real path of data transmission

Line of sight

[1]https://perfops.net/ [2]https://bbmaps.itu.int/bbmaps/ [3]https://www.submarinecablemap.com

Routing Methods



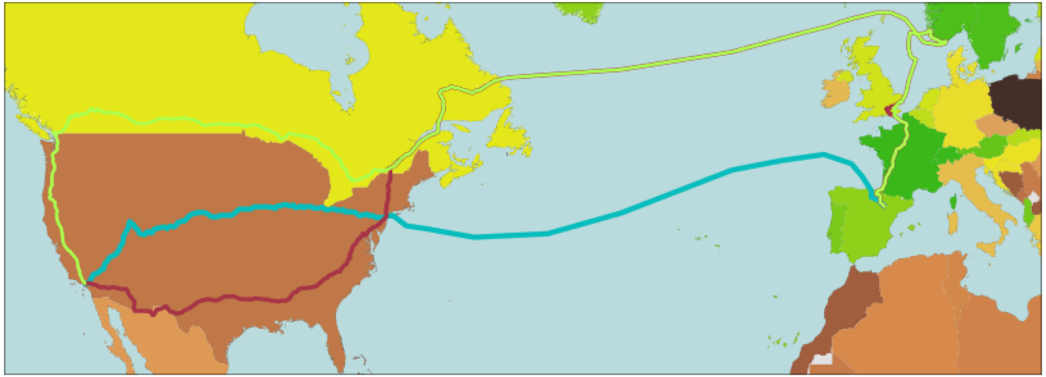
Baseline: MTR hop sequence with minimum CI routing between traced hops

Shortest: Path with minimum total distance

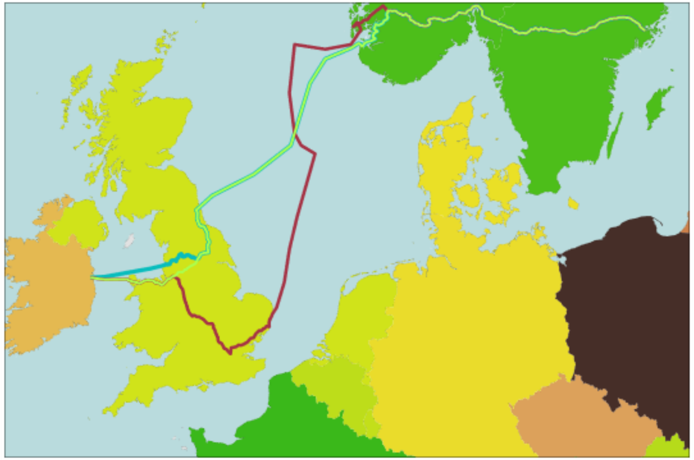
Lowest CI: Path minimizing CI

Dijkstra‘s shortest path algorithm is used to compute paths with the lowest total weight.

Exemplary Routes



Los Angeles, USA to Zaragoza, Spain with carbon costs 38.16, 51.67 (+35.4%), 22.41 (-41.3%)



Dublin, Ireland to Stockholm, Sweden with carbon costs 2.93, 2.55 (-13.1%), 2.07 (-29.5%)

CI in gCO_2eq/kWh

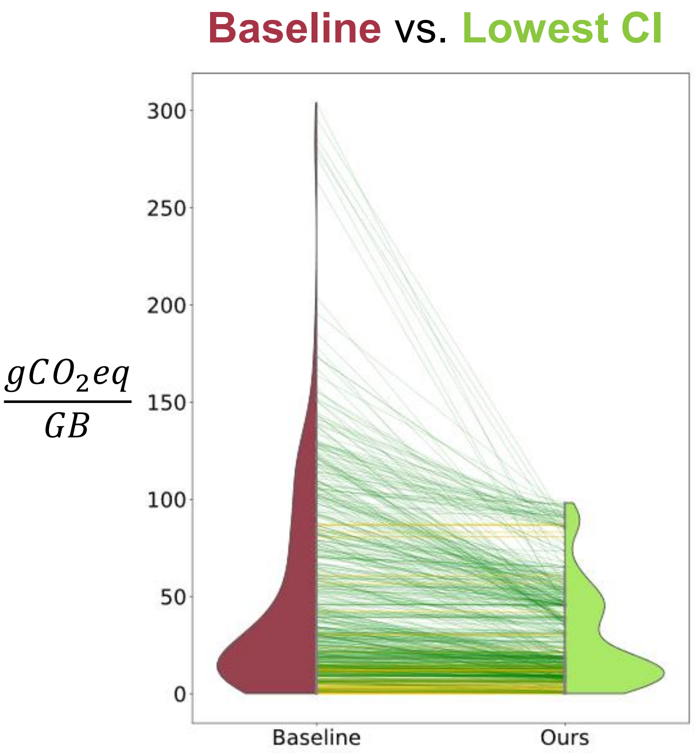
0 700

Baseline
Shortest
Lowest CI

Results for selected routes

From City	To City	Baseline		Shortest			Lowest CI (ours)		
		gCO2/GB	std (±)	gCO2/GB	$\Delta_{Base.}$ (%)	std (±)	gCO2/GB	$\Delta_{Base.}$ (%)	std (±)
Tokyo	São Paulo	76.77	4.08	78.26	1.95	12.72	58.05	**-24.39**	5.57
Singapore	Zürich	32.77	4.19	55.65	69.80	2.51	31.95	**-2.51**	3.31
Amsterdam	San Jose	39.07	3.44	45.69	16.92	4.26	18.06	**-53.79**	1.99
Chicago	Jakarta	124.02	9.18	88.35	-28.76	5.05	46.66	**-62.37**	2.97
Los Angeles	Zaragoza	38.16	1.58	51.67	35.38	2.77	22.41	**-41.29**	1.39
Seattle	Stockholm	15.28	1.81	9.09	-40.51	0.88	9.07	**-40.67**	0.88
Dublin	Stockholm	2.93	2.09	2.55	-13.14	1.61	2.07	**-29.54**	1.32
Frankfurt am Main	Dār Kulayb	13.20	2.66	19.71	49.33	0.91	13.20	**0.00**	2.66
Athens	Milan	8.40	0.97	4.37	-47.95	0.48	3.77	**-55.14**	0.36
Frankfurt am Main	Melbourne	101.49	15.60	94.11	-7.27	6.68	76.99	**-24.14**	8.38

Average Improvements



Pairwise Route Comparison

- Corresponding pairs of routes are connected by lines.
- Green lines: reduction in CE
  Yellow lines: equal CE
- N = 670 route pairs

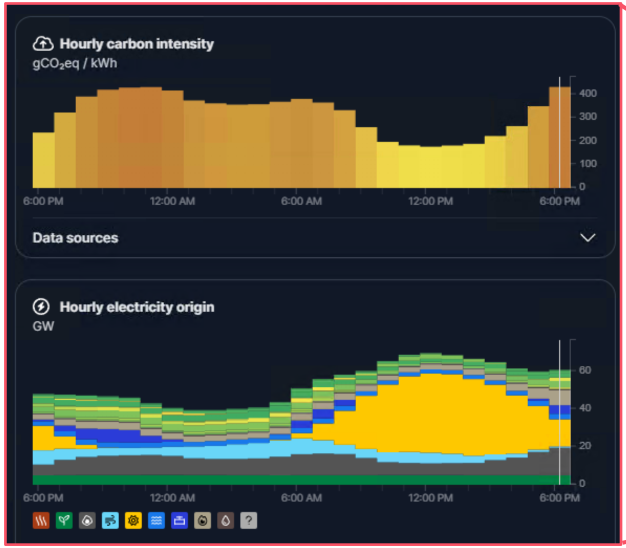
Mean (standard deviation):

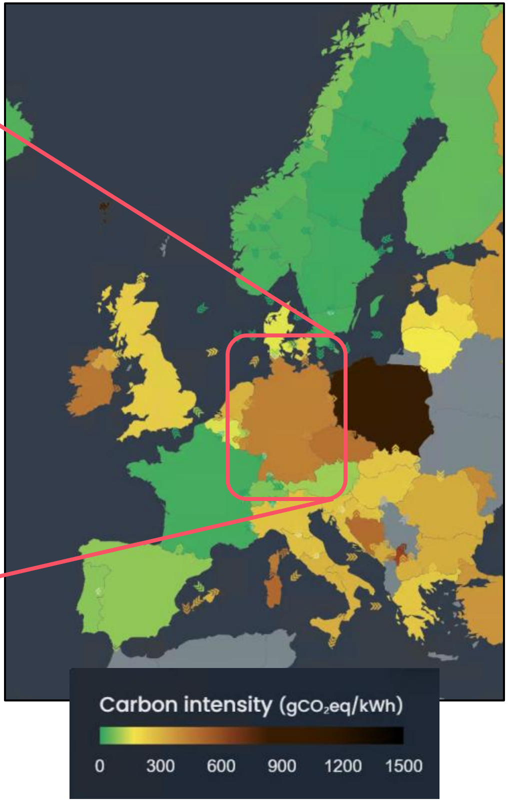
Baseline: 46.89 (49.05)
Lowest CI: **28.1** (25.34)
Shortest: 43.98 (32.68)

Time-of-Day Variation in CI

Varies by time of day ▶

Energy mix ▶

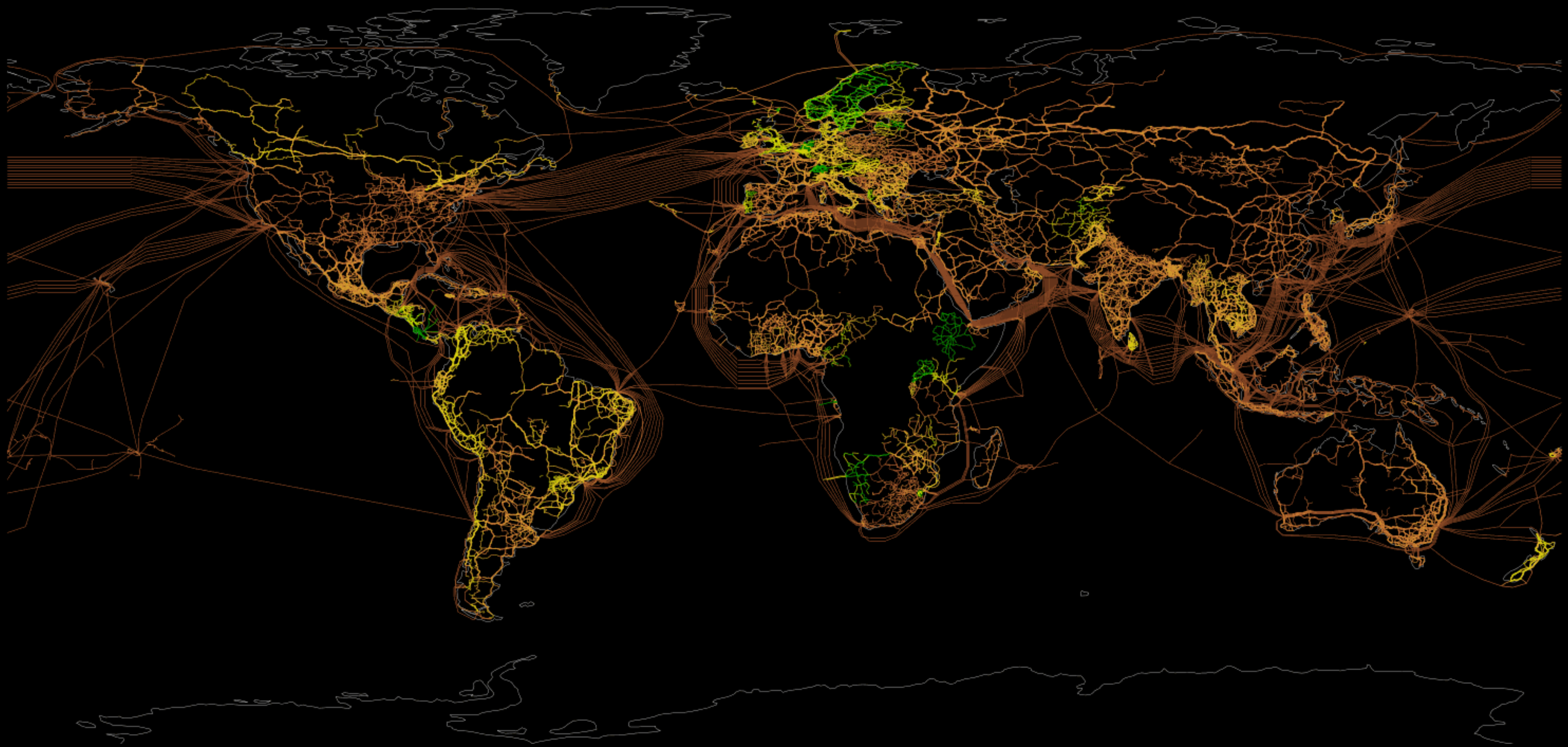




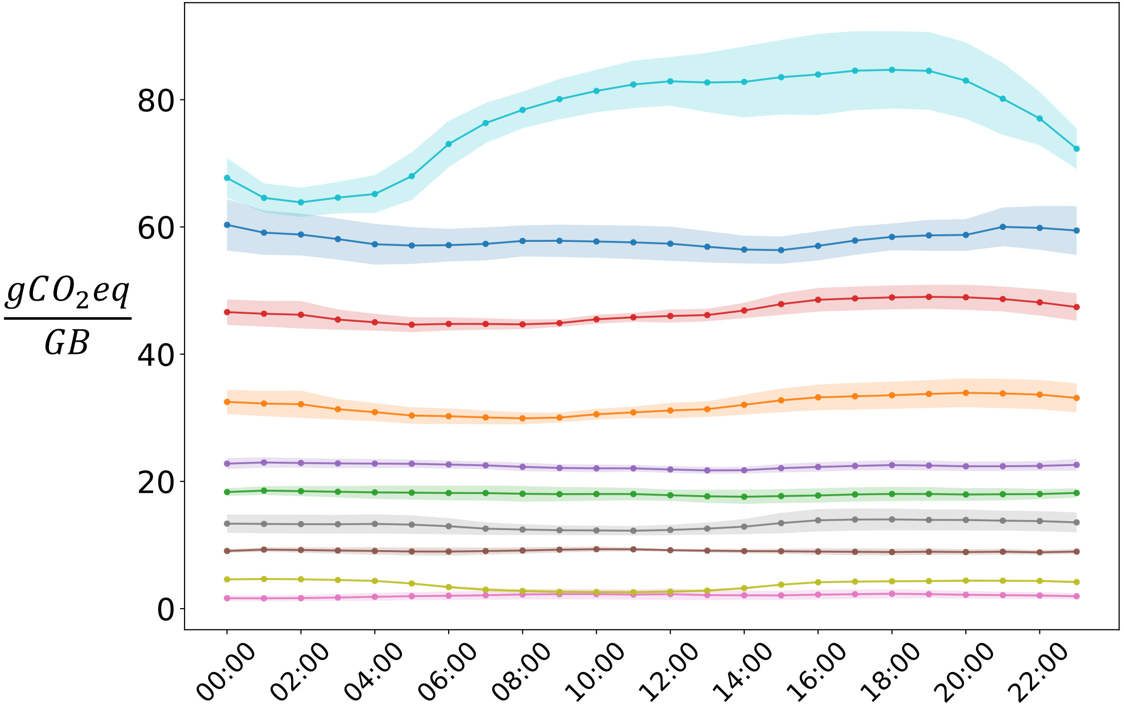
➔ Recompute graph weights for every hour!

app.electricitymaps.com/map/

Carbon Emission Graph



Time-of-Day Variation

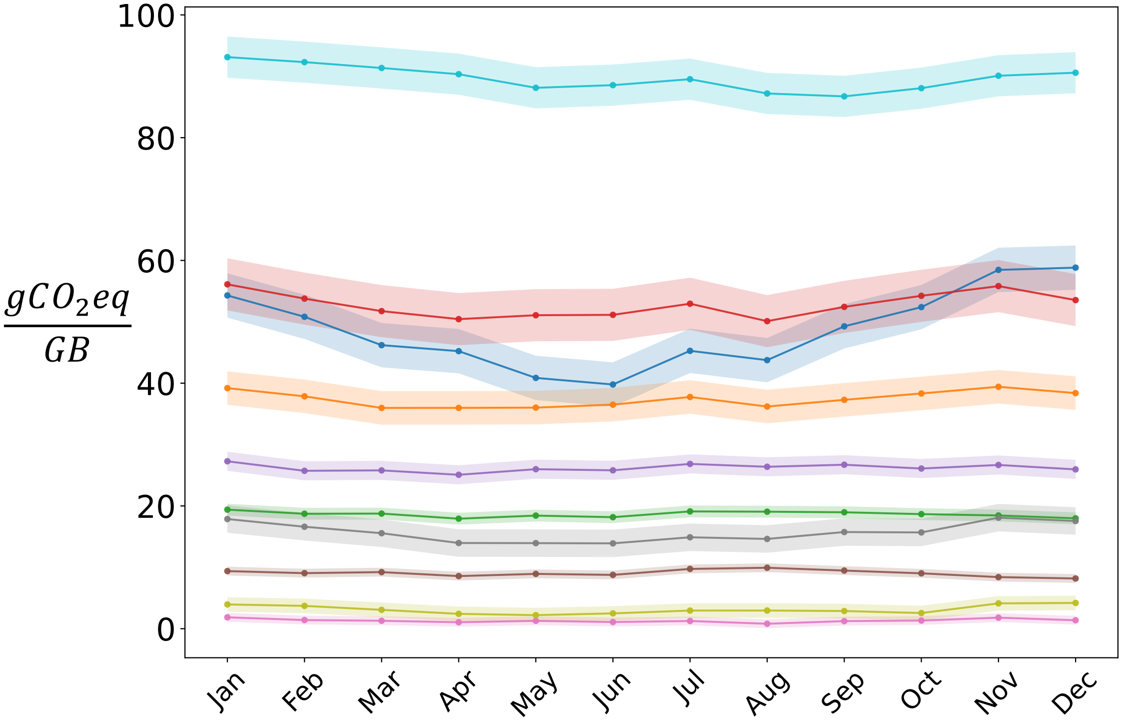


Data

- 4 days of 4 months (January, April, June, October)
- 24h per day
- Mean and SEM of Lowest CI paths

From City	To City	Baseline	Shortest	Lowest CI
Frankfurt am Main	Melbourne	33647	20678	32433
Tokyo	São Paulo	35292	24342	26568
Chicago	Jakarta	35606	20968	31944
Singapore	Zürich	19929	12016	19408
Los Angeles	Zaragoza	14761	11048	15414
Amsterdam	San Jose	24981	11106	13501
Frankfurt am Main	Dār Kulayb	9988	5743	9988
Seattle	Stockholm	15365	11475	11493
Athens	Milan	3105	1824	1905
Dublin	Stockholm	2671	2149	2156

Seasonal Variation



Data

- All days and months of 2024
- 1h per day (11:00)
- Mean and SD of Lowest CI paths

From City	To City	Baseline	Shortest	Lowest CI
Frankfurt am Main	Melbourne	33647	20678	32433
Tokyo	São Paulo	35292	24342	26568
Chicago	Jakarta	35606	20968	31944
Singapore	Zürich	19929	12016	19408
Los Angeles	Zaragoza	14761	11048	15414
Amsterdam	San Jose	24981	11106	13501
Frankfurt am Main	Dār Kulayb	9988	5743	9988
Seattle	Stockholm	15365	11475	11493
Athens	Milan	3105	1824	1905
Dublin	Stockholm	2671	2149	2156

Summary

Conclusion

- We propose a **graph-based**, **carbon-aware** routing approach for internet data transfers.
- Incorporates **temporal and spatial variations** in grid carbon intensity.
- Achieves ~40% emission reduction compared to the baseline (N = 670 route pairs).

Limitations

- Internet routing is not (yet) path-aware; routes cannot currently be precomputed
  - Initiatives like SCION[1] (ETH Zürich spin-off) may enable path-aware routing in the future.
- Data transfer duration is not considered in the current model.

[1] https://www.scion.org

Thanks for your attention