

A Model Details & Hyperparameters

A.1 Basic Geographic Features

We extract urban geographic features from OpenStreetMap (OSM) ⁵ and WorldPop ⁶ data. The original geographic features include the following three aspects.

- POI features (denoted as $f^{(poi)}$) reflect the function of a region. We attempt to construct the POI semantic features of a region by counting the quantities of various POI categories. However, the quantities of different POI categories exhibit significant imbalance, for instance, the total number of commercial POIs is far larger than that of residential POIs. This imbalance can easily cause the model to overlook the influence of POI categories with smaller quantities. Therefore, we employ the TF-IDF algorithm [39] to extract POI features. Specifically, we treat each POI category as a "word", consider all POIs within a region as a "document", and define the entire city as the "corpus". When calculating the importance of a word for a document, the algorithm automatically incorporates the word's frequency in the corpus for weighted analysis.
- Road features (denoted as $f^{(road)}$) reflect the transportation attribute of a region. We calculate the total length of all categories of road segments within a region. Regions with dense road networks and a large number of trunk roads usually have convenient transportation and tend to generate a higher volume of traffic flow.
- Population features (denoted as $f^{(pop)}$) reflect the traffic potential of a region. Regions with higher population density are more likely to generate a higher volume of traffic flow. We obtained the United Nations (UN)-adjusted 100m resolution national population data from *WorldPop* and counted the population number in each manually partitioned rectangular regions.

A.2 Optimal Transport Problem

Optimal Transport (OT) is a mathematical problem aiming to find the most efficient way to move mass from source distribution to target distribution. It was introduced by *Gaspard Monge* in 1781. When both the source and target distributions are represented by enumerable samples (N_s samples for source and N_t samples for target), the OT problem can be formally defined as

$$\begin{aligned} \mathbf{T}^* &= \arg \min_{\mathbf{T} \in \mathbb{R}_+^{N_s \times N_t}} \sum_{i,j} \mathbf{T}_{i,j} \cdot \mathbf{D}_{i,j}, \\ s.t. \quad \mathbf{T}\mathbf{1} &= \mathbf{w}_s \text{ and } \mathbf{T}^\top \mathbf{1} = \mathbf{w}_t, \end{aligned} \tag{15}$$

where $\mathbf{D} \in \mathbb{R}_+^{N_s \times N_t}$ is the cost matrix (distance matrix) defining the cost to move mass from source distribution to target distribution, $\mathbf{w}_s \in \mathbb{R}^{N_s}$ and $\mathbf{w}_t \in \mathbb{R}^{N_t}$ are the weights of each samples in the source and target distribution. The total weights of both \mathbf{w}_s and \mathbf{w}_t are equal to 1. The objective of the OT problem is to find a transportation plan \mathbf{T}^* that minimizes the total transportation cost under the weights-equal constraint.

OT problem has two main functions: (1) Measuring the distance between two distributions; (2) Finding the correspondences between two distributions. We employed both. Specifically, we treated the geographic representations of regions in the source and target cities as two mass distributions, with each region assigned the same weight. we used the solution of the OT problem (also known as the Wasserstein distance) to measure the geographic representation distance of the source and target cities, and treated it as a loss to optimize the spatial encoder, thereby pulling the correspondent regions in representation space.

As illustrated in Equation 15, the OT problem is a linear programming problem, and we use an OT solver based on the network simplex algorithm [2] to address it. Thanks to the Python Optimal Transport (POT) ⁷ tool, we can conveniently calculate the solution of OT problem.

⁵<https://www.openstreetmap.org>

⁶<https://www.worldpop.org/>

⁷<https://github.com/PythonOT/POT>

Table 3: Hyperparameters setting for CRAFT

Hyperparameter	Setting value
Diffusion steps (n)	500
$\beta_1 \sim \beta_n$	0.0002 \sim 0.04 (linear)
GraphTransformer (GT) layers	3
Temporal encoder layers	2
GT attention heads	4
Retrieval top-K value	5
Batch size	256
Learning rate	5×10^{-6}
Training epochs	300
Regional geographic representation (h_i) dimensions	128
Temporal encoder hidden dimensions (dimensions of $x_{i,t}$)	256
Hour embedding (t_{hour}) dimensions	64
Weekday embedding (t_{week}) dimensions	64
Month embedding (t_{month}) dimensions	64
Condition ($c_{i,t}$) dimensions	256

A.3 Implementation Details

For the proposed CRAFT method, we provide the hyperparameter settings in Table 3 to facilitate the reproducibility by researchers. All these parameters are recommended values, not fixed, and can be adjusted according to the dataset and experimental environment. During training, the AdamW optimizer was used. To enhance stability, the EMA (Exponential Moving Average) mechanism was adopted to train the diffusion model.

B Details of Experimental Settings

B.1 Experimental Environment

All neural network models (including CRAFT and other baselines) are implemented in PyTorch and trained on a single NVIDIA RTX 3090 GPU. The experimental machine ran on Ubuntu 20.04.6 LTS, was equipped 24-core Intel(R) Xeon(R) Silver CPU, and had 503 GB of RAM. The training time for all models on a single dataset did not exceed 16 hours.

B.2 Datasets and Pre-processing

Table 4: Data description

Datasets	Chicago	Washington D.C.	Toronto	New York City
# Trips ($\times 10^3$)	5136	4011	2395	35080
Time range	2023.01-2023.12	2023.01-2023.12	2020.01-2020.12	2023.01-2023.12
# Regions	73	82	61	96
# POIs	17205	14070	20621	50776

We conducted experiments using the traffic flow datasets of four cities, namely Chicago, Washington D.C., Toronto, and New York City. The original data are all trip records of shared bicycles, which include the latitude and longitude of the starting and ending points of users’ trips as well as timestamps. We associated the trips with the manually partitioned urban regions, and counted the number of bicycles entering and leaving each region within each hour, which served as the traffic flow values. Details of the datasets are presented in Table 4.

In fact, the user trip data is sparse, which leads to the instability of the traffic flow trend in original data. This is also a common problem in researches about traffic flow data. In response to this, we have adopted two processing methods: (1) We have filled in the missing values in the traffic flow sequence through linear interpolation. For the situation where there are values at the previous and subsequent time steps but missing values in the middle, we have filled them with the average of the

previous and subsequent values. (2) We have used the moving average method to smooth the values within a window with a time length of 3. For the processed data, we have set a sliding window with a length of T to extract the training and test samples. If there are still missing values exceeding 5% in a certain sample, we will discard that sample.

B.3 Baseline Methods

The baseline models used in the experiment include two types of deep learning traffic flow generation models that have emerged in the literature in recent years: *Static Flow Generation Models* and *Dynamic Flow Generation Models*. *Static Flow Generation Models* directly estimate traffic flow using the geographic features and temporal information of urban regions, and they include the following two models.

- GMEL [27]: It uses two graph neural networks to extract features and can simultaneously predict the inflow and outflow as well as the OD (origin-destination) flow between regions.
- DFG [40]: It uses a deep cross network to extract the POI (point of interest) features inside and outside the region, conducts supervised training using the intention-aware pedestrian flow, and predicts the inflow and outflow of the region.

Dynamic Flow Generation Models use deep generative models to learn the complex distribution of data and map random noise into data samples. The selected baseline models are as follows.

- KSTDiff [58]: It uses the urban knowledge graph to extract the representations of urban geographic entities, and then constructs a knowledge-enhanced spatiotemporal diffusion model to generate the inflow and outflow of regions.
- CGAN [29]: This is a conditional generative adversarial network. In the experiment, we use the original static geographic features and time embeddings as conditions to guide the generation of the GAN.
- Diffwave [23]: This is a diffusion model suitable for generating time series data and is often used in speech synthesis. In this experiment, we use it to learn the distribution of traffic flow data.
- DiT [31]: This is a diffusion probabilistic model with a Transformer as the noise estimator.
- DDPM [12]: This is a diffusion probabilistic model U-Net as noise estimator for image synthesizing. To be applied in traffic flow data generation, we replaced 2D convolutions with 1D convolutions.
- CVAE [10]: This is a conditional variational autoencoder. Condition information is added to both the Encoder and the Decoder to guide it to learn the conditional probability distribution of the data. We use the original static geographic features and time embeddings as conditions.

B.4 Evaluation metrics

We used three commonly used metrics in traffic flow generation research to evaluate the quality of generated data, including Common Part of Commuters (CPC), Normalized Mean Absolute Error (NMAE), Normalized Root Mean Square Error (NRMSE). Since the actual traffic flow values in different cities vary greatly, the traffic flow data generated by our model are all normalized, which can reflect the relative magnitudes of traffic flows in different regions and at different time in the same city. All the metrics are calculated based on the normalized data, and the calculation formulas are as follows

$$\begin{aligned} \text{CPC} &= \frac{2 \sum_{i=1}^M \min(\hat{y}_i, y_i)}{\sum_{i=1}^M \hat{y}_i + \sum_{i=1}^M y_i}, \\ \text{NMAE} &= \frac{\frac{1}{M} \sum_{i=1}^M |\hat{y}_i - y_i|}{\max(y_i) - \min(y_i)}, \\ \text{NRMSE} &= \frac{\sqrt{\frac{1}{M} \sum_{i=1}^M (\hat{y}_i - y_i)^2}}{\max(y_i) - \min(y_i)}, \end{aligned} \quad (16)$$

where \hat{y}_i is the generated value, y_i is the real value, and M is the number of values of all samples in the test dataset. To enhance the stability of evaluation results, we grouped the data by the Region ID, month, weekday and hour of the target city's test samples and calculated the differences between the group-averaged values.

C Additional Experiments

C.1 Sensitivity Analysis Results

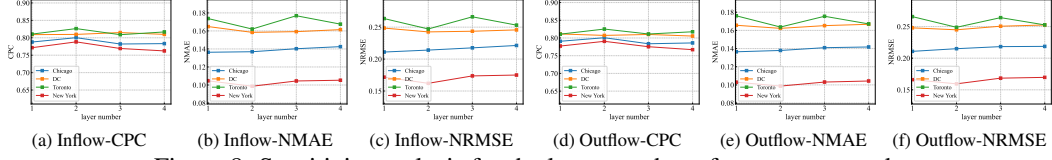


Figure 8: Sensitivity analysis for the layer number of sequence encoder

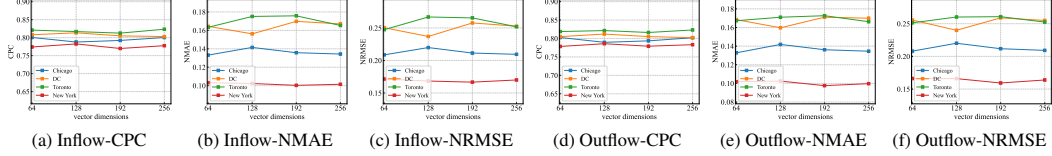


Figure 9: Sensitivity analysis for the dimension of temporal embedding

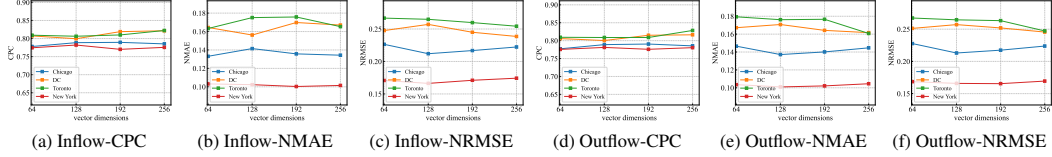


Figure 10: Sensitivity analysis for the dimension of geographic representation

To validate how different hyperparameter configurations affect model performance, we plotted the variation curves of evaluation metrics against three key hyperparameters.: (1) the layer number of the sequence encoder; (2) the dimension of temporal embedding; (3) the dimension of the geographic representation. The results of sensitivity analyses across all datasets and evaluation metrics are shown in the Fig. 8, Fig. 9 and Fig. 10. We can observe that CRAFT demonstrates excellent robustness across all datasets and evaluation metrics, eliminating the need for meticulous hyperparameter tuning to ensure superior model performance.

C.2 Temporal Length Extension Results

Overall Zero-shot Generation Performance in Target City. In the main experiments of the paper, we set the generation time window length to $T = 24$. We then progressively extended the time length T to the following values: $\{48, 72, 96, 120, 144, 168\}$. The experimental setup remained consistent: training the model using data from three cities and performing zero-shot generation on the fourth city. Detailed evaluation results are presented in Table 6. From these tables, we can conclude that: (1) CRAFT achieved best performance in over 98.6% of cases across all datasets and evaluation metrics, demonstrating that CRAFT exhibits state-of-the-art (SOTA) zero-shot generalization capabilities for sequence data generation of varying temporal lengths. (2) CRAFT demonstrates stable performance, while methods such as Diffwave and DDPM exhibit significant performance fluctuations under different temporal length settings.

Data Utility Comparison. We compared the utility of generated data across different temporal lengths for traffic flow prediction tasks. Regardless of the value of T , we used the first $T/2$ historical sequence as input to predict the $T/2$ future flow sequence. The downstream models are trained on generated data from various methods and tested on real data. Detailed results are presented in Table 7 and 8. We found that in 86.5% of cases, the downstream model trained with CRAFT’s generated data achieved best performance, indicating that CRAFT’s generated data has the better utility than other baselines.

Computational Cost. As the length of temporal length increases, the model size and training time will also grow correspondingly. To observe this phenomenon in detail, we conducted experiments on a single NVIDIA RTX 3090 GPU, collected relevant statistical data during the training phase, and the results are shown in Table 5. We observe that our model’s computational cost grows linearly with temporal length increases, but this effect is weak. Extending the time length is acceptable in terms of computational cost.

Table 5: The relationship between computational cost and temporal length

T	Model size (Byte)	Train time (s/epoch)	Valid time (s/epoch)	Avg memory (GB)
24	51781096	40.999	5.685	0.343
48	51830248	41.156	5.531	0.344
72	51879400	43.729	6.053	0.344
96	51928552	44.555	5.631	0.345
120	51977704	46.712	6.429	0.345
144	52026856	48.883	6.666	0.345
168	52076008	51.896	6.351	0.346

C.3 Visualization of the Traffic Flow Spatial Heatmap

We display the average traffic flow of real data and generated data on maps. The visualizations for all datasets and baselines are shown in Fig. 11. This comparison intuitively demonstrates that the traffic data generated by CRAFT exhibits the highest similarity to real data in terms of spatial distribution, indicating that CRAFT can effectively capture the universal mapping relationship between geographic representations and traffic flow across different cities.

C.4 Visualization of Geographic Feature Alignment (GFA)

We employed t-SNE analysis to visualize the impact of the Traffic Flow Alignment (TFA) and Cross City Alignment (CCA) modules in Geographic Feature Alignment. The results of the four experiments are presented in Fig. 12. Regardless of which city was chosen as the target, the alignment results exhibited similar conclusions: (1) Under the combined action of TFA and CCA, regions with high and low traffic volumes were well-separated in the representation space. (2) Without TFA, representations from different cities tended to cluster into multiple groups, with high- and low-traffic regions intermingled within the same cluster, making it difficult to distinguish and reducing the quality of conditions. (3) Without CCA, significant domain shift occurred between the source and target cities. Specifically, a portion of the target city’s region representations deviated from the concentrated representation area of the source city, potentially leading to poorer transferability.

Table 6: Cross-city traffic flow generation results with extended temporal length

Method	City	$T = 48$			$T = 72$			$T = 96$			$T = 120$			$T = 144$			$T = 168$		
		CPC	NMAE	NRMSE	CPC	NMAE	NRMSE	CPC	NMAE	NRMSE	CPC	NMAE	NRMSE	CPC	NMAE	NRMSE	CPC	NMAE	NRMSE
GMEL	Chicago(Inflow)	0.590	0.215	0.310	0.645	0.201	0.288	0.741	0.212	0.268	0.490	0.246	0.351	0.700	0.183	0.257	0.512	0.246	0.358
DFG		0.155	0.328	0.457	0.153	0.334	0.462	0.152	0.340	0.467	0.156	0.326	0.454	0.156	0.328	0.455	0.155	0.330	0.457
KSTDiff		0.000	0.358	0.488	0.460	0.288	0.399	0.003	0.370	0.499	0.540	0.444	0.533	0.552	0.539	0.616	0.032	0.360	0.489
CGAN		0.584	0.379	0.488	0.539	0.371	0.487	0.521	0.386	0.500	0.461	0.359	0.476	0.546	0.492	0.582	0.509	0.437	0.535
Diffwave		0.490	0.504	0.613	0.384	0.456	0.570	0.464	0.430	0.533	0.492	0.500	0.609	0.360	0.437	0.555	0.479	0.437	0.556
DT		0.572	0.325	0.407	0.564	0.362	0.448	0.593	0.333	0.405	0.580	0.356	0.445	0.561	0.341	0.425	0.510	0.340	0.438
DDPM		0.398	0.300	0.419	0.417	0.306	0.427	0.455	0.313	0.428	0.490	0.291	0.397	0.481	0.300	0.411	0.509	0.313	0.417
CVAE		0.476	0.269	0.388	0.555	0.286	0.404	0.616	0.282	0.392	0.509	0.280	0.395	0.525	0.319	0.427	0.527	0.297	0.402
CRAFT		0.789	0.145	0.223	0.798	0.145	0.224	0.791	0.150	0.229	0.785	0.149	0.228	0.777	0.154	0.233	0.645	0.251	0.356
GMEL	Chicago(Outflow)	0.667	0.197	0.278	0.455	0.266	0.373	0.720	0.199	0.259	0.606	0.213	0.310	0.429	0.266	0.371	0.502	0.248	0.352
DFG		0.152	0.333	0.460	0.150	0.339	0.465	0.149	0.346	0.470	0.153	0.332	0.457	0.152	0.333	0.458	0.152	0.335	0.460
KSTDiff		0.000	0.363	0.491	0.214	0.335	0.462	0.006	0.375	0.501	0.533	0.628	0.708	0.569	0.478	0.554	0.151	0.356	0.479
CGAN		0.587	0.367	0.476	0.552	0.370	0.483	0.529	0.379	0.493	0.438	0.368	0.486	0.540	0.526	0.612	0.505	0.433	0.534
Diffwave		0.459	0.445	0.561	0.350	0.469	0.584	0.421	0.482	0.591	0.298	0.428	0.544	0.353	0.460	0.575	0.500	0.490	0.601
DT		0.546	0.346	0.434	0.531	0.346	0.437	0.579	0.346	0.423	0.550	0.349	0.438	0.550	0.364	0.449	0.536	0.342	0.434
DDPM		0.405	0.301	0.418	0.423	0.307	0.427	0.446	0.319	0.434	0.491	0.292	0.398	0.479	0.302	0.413	0.510	0.315	0.420
CVAE		0.472	0.274	0.392	0.562	0.287	0.404	0.620	0.283	0.392	0.507	0.281	0.395	0.534	0.316	0.422	0.532	0.294	0.398
CRAFT		0.791	0.146	0.223	0.797	0.147	0.225	0.797	0.148	0.225	0.792	0.146	0.223	0.789	0.148	0.224	0.656	0.246	0.349
GMEL	Washington, D.C.(Inflow)	0.728	0.234	0.284	0.609	0.289	0.383	0.649	0.272	0.360	0.764	0.223	0.261	0.726	0.227	0.292	0.711	0.234	0.301
DFG		0.697	0.246	0.348	0.696	0.252	0.354	0.698	0.254	0.355	0.692	0.252	0.354	0.692	0.254	0.356	0.691	0.256	0.358
KSTDiff		0.000	0.465	0.595	0.494	0.356	0.458	0.289	0.421	0.542	0.039	0.464	0.591	0.000	0.477	0.603	0.641	0.477	0.584
CGAN		0.596	0.342	0.465	0.510	0.420	0.538	0.575	0.440	0.550	0.583	0.391	0.507	0.653	0.480	0.593	0.442	0.391	0.514
Diffwave		0.578	0.505	0.624	0.274	0.466	0.584	0.407	0.557	0.663	0.450	0.577	0.677	0.641	0.506	0.623	0.457	0.580	0.673
DT		0.621	0.351	0.440	0.627	0.354	0.432	0.617	0.369	0.446	0.608	0.359	0.451	0.598	0.386	0.479	0.630	0.357	0.437
DDPM		0.476	0.355	0.477	0.417	0.418	0.534	0.488	0.403	0.510	0.541	0.349	0.458	0.609	0.340	0.443	0.463	0.370	0.475
CVAE		0.296	0.402	0.534	0.480	0.368	0.486	0.423	0.388	0.511	0.566	0.347	0.455	0.433	0.399	0.518	0.464	0.374	0.491
CRAFT		0.791	0.185	0.275	0.786	0.191	0.282	0.777	0.202	0.301	0.781	0.191	0.281	0.781	0.194	0.285	0.778	0.197	0.288
GMEL	Washington, D.C.(Outflow)	0.683	0.239	0.321	0.557	0.303	0.401	0.616	0.297	0.390	0.534	0.310	0.408	0.633	0.271	0.363	0.688	0.245	0.317
DFG		0.696	0.246	0.347	0.694	0.252	0.353	0.695	0.254	0.355	0.689	0.252	0.353	0.689	0.254	0.355	0.688	0.256	0.357
KSTDiff		0.000	0.470	0.598	0.581	0.338	0.412	0.004	0.488	0.612	0.000	0.476	0.601	0.536	0.503	0.622	0.653	0.513	0.629
CGAN		0.583	0.354	0.476	0.543	0.427	0.540	0.596	0.441	0.551	0.554	0.408	0.519	0.653	0.497	0.610	0.423	0.401	0.519
Diffwave		0.502	0.481	0.605	0.321	0.434	0.562	0.254	0.546	0.650	0.289	0.476	0.596	0.202	0.546	0.655	0.450	0.421	0.549
DT		0.580	0.371	0.461	0.579	0.361	0.450	0.623	0.377	0.464	0.583	0.357	0.452	0.589	0.387	0.478	0.621	0.372	0.463
DDPM		0.474	0.359	0.479	0.422	0.416	0.531	0.496	0.395	0.503	0.543	0.353	0.460	0.607	0.345	0.448	0.462	0.372	0.475
CVAE		0.301	0.404	0.533	0.484	0.368	0.483	0.429	0.386	0.507	0.551	0.355	0.461	0.436	0.400	0.516	0.464	0.376	0.489
CRAFT		0.790	0.188	0.277	0.783	0.195	0.285	0.773	0.207	0.304	0.781	0.193	0.280	0.781	0.196	0.284	0.778	0.199	0.288
GMEL	Toronto(Inflow)	0.748	0.230	0.281	0.721	0.240	0.309	0.639	0.283	0.366	0.558	0.313	0.407	0.672	0.260	0.344	0.703	0.255	0.326
DFG		0.275	0.421	0.535	0.271	0.432	0.544	0.269	0.441	0.551	0.275	0.423	0.536	0.274	0.425	0.537	0.272	0.430	0.542
KSTDiff		0.002	0.500	0.622	0.672	0.398	0.472	0.685	0.479	0.606	0.670	0.393	0.468	0.378	0.427	0.536	0.268	0.485	0.601
CGAN		0.631	0.358	0.474	0.394	0.436	0.546	0.528	0.399	0.513	0.546	0.378	0.478	0.556	0.387	0.495	0.421	0.404	0.512
Diffwave		0.550	0.424	0.553	0.634	0.462	0.588	0.498	0.523	0.633	0.567	0.475	0.594	0.506	0.488	0.605	0.583	0.471	0.587
DT		0.638	0.381	0.469	0.608	0.378	0.458	0.595	0.388	0.477	0.600	0.377	0.460	0.597	0.378	0.467	0.611	0.377	0.460
DDPM		0.625	0.360	0.471	0.649	0.367	0.479	0.600	0.388	0.498	0.628	0.375	0.484	0.608	0.373	0.483	0.602	0.381	0.490
CVAE		0.714	0.302	0.402	0.712	0.305	0.414	0.699	0.305	0.412	0.679	0.302	0.416	0.583	0.368	0.469	0.592	0.368	0.461
CRAFT		0.817	0.178	0.265	0.824	0.176	0.264	0.826	0.174	0.260	0.811	0.181	0.269	0.816	0.182	0.273	0.800	0.194	0.289
GMEL	Toronto(Outflow)	0.736	0.235	0.292	0.632	0.286	0.370	0.716	0.248	0.317	0.734	0.236	0.297	0.688	0.254	0.334	0.722	0.240	0.308
DFG		0.274	0.420	0.533	0.270	0.432	0.543	0.268	0.441	0.550	0.274	0.422	0.534	0.274	0.425	0.536	0.271	0.430	0.540
KSTDiff		0.002	0.498	0.619	0.682	0.411	0.492	0.685	0.479	0.604	0.673	0.402	0.482	0.646	0.419	0.513	0.310	0.492	0.609
CGAN		0.637	0.359	0.472	0.523	0.420	0.526	0.526	0.405	0.515	0.544	0.379	0.478	0.547	0.393	0.499	0.559	0.368	0.457
Diffwave		0.469	0.432	0.548	0.617	0.446	0.564	0.558	0.438	0.561	0.420	0.471	0.589	0.574	0.432	0.547	0.555	0.467	0.581
DT		0.619	0.381	0.469	0.588	0.394	0.482	0.619	0.381	0.465	0.611	0.365	0.444	0.595	0.397	0.489	0.604	0.380	0.467
DDPM		0.628	0.360	0.470	0.650	0.373	0.485	0.603	0.390	0.498	0.633	0.375	0.483	0.609	0.376	0.486	0.607	0.382	0.490
CVAE		0.709	0.307	0.408	0.705	0.313	0.420	0.695	0.310	0.415	0.680	0.305	0.417	0.591	0.365	0.463	0.606	0.358	0.449
CRAFT		0.814	0.182	0.270	0.821	0.180	0.270	0.822	0.181	0.269	0.810	0.182	0.271	0.814	0.188	0.280	0.803	0.195	0.289
GMEL	New York City(Inflow)	0.676	0.178	0.238	0.621	0.256	0.336	0.598	0.297	0.362	0.611	0.263	0.331	0.602	0.282	0.347	0.511	0.429	0.500
DFG		0.583	0.182	0.279	0.585	0.182	0.278	0.584	0.182	0.279	0.583	0.181	0.278	0.583	0.181	0.278	0.582	0.181	0.277
KSTDiff		0.355	0.442	0.562	0.035	0.458	0.568	0.000	0.470	0.594	0.372	0.471	0.586	0.169	0.486	0.596	0.184	0.286	0.396
CGAN		0.425	0.561	0.667	0.419	0.585	0.682	0.396	0.610	0.703	0.454	0.377	0.559	0.411	0.549	0.644	0.440	0.380	0.496
Diffwave		0.313	0.500	0.621	0.213	0.388	0.531	0.376	0.468	0.735	0.413	0.405	0.548	0.236	0.435	0.576	0.219	0.540	0.661
DT		0.517	0.384	0.499	0.481	0.324	0.422	0.470	0.335	0.420	0.440	0.398	0.480	0.471	0.351	0.435	0.447	0.375	0.464
DDPM		0.457	0.375	0.499	0.481	0.381	0.389	0.519	0.247	0.362	0.485	0.335	0.454	0.484	0.330	0.445	0.456	0.306	0.421
CVAE		0.442	0.406	0.535	0.483	0.381	0.454	0.472	0.343	0.471	0.473	0.373	0.406	0.573	0.406	0.470	0.329	0.477	0.587
CRAFT		0.76																	

Table 7: Data utility comparison on traffic flow prediction (LSTM)

Gen	Pred	Chicago				Washington, D.C.				Toronto				New York City			
		Inflow		Outflow		Inflow		Outflow		Inflow		Outflow		Inflow		Outflow	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Real	LSTM ($T = 48$)	0.097	0.151	0.100	0.154	0.094	0.144	0.097	0.150	0.109	0.163	0.111	0.164	0.056	0.101	0.056	0.101
GMEL		0.167	0.230	0.166	0.228	0.202	0.254	0.210	0.268	0.241	0.303	0.251	0.320	0.092	0.144	0.091	0.142
DFG		0.283	0.418	0.289	0.423	0.134	0.202	0.136	0.206	0.319	0.426	0.321	0.427	0.070	0.126	0.069	0.125
KSTDiff		0.362	0.493	0.368	0.497	0.469	0.599	0.475	0.603	0.507	0.629	0.505	0.627	0.348	0.389	0.762	0.814
CGAN		0.199	0.292	0.212	0.309	0.283	0.394	0.286	0.399	0.287	0.388	0.286	0.382	0.151	0.248	0.163	0.257
Diffwave		0.221	0.299	0.232	0.297	0.276	0.354	0.331	0.441	0.327	0.434	0.286	0.362	0.141	0.211	0.138	0.211
DiT		0.232	0.285	0.246	0.292	0.266	0.311	0.288	0.330	0.299	0.337	0.307	0.343	0.252	0.280	0.209	0.245
DDPM		0.113	0.174	0.115	0.171	0.124	0.183	0.122	0.183	0.117	0.171	0.122	0.175	0.065	0.113	0.066	0.110
CVAE		0.175	0.255	0.172	0.249	0.308	0.429	0.304	0.421	0.158	0.230	0.162	0.233	0.138	0.219	0.135	0.213
CRAFT		0.099	0.153	0.103	0.158	0.103	0.156	0.111	0.167	0.120	0.173	0.127	0.180	0.063	0.107	0.064	0.105
Real	LSTM ($T = 72$)	0.104	0.156	0.107	0.159	0.095	0.145	0.099	0.151	0.114	0.169	0.115	0.170	0.059	0.106	0.059	0.105
GMEL		0.169	0.234	0.175	0.246	0.227	0.285	0.240	0.302	0.258	0.329	0.270	0.349	0.093	0.140	0.097	0.153
DFG		0.303	0.435	0.306	0.438	0.142	0.213	0.151	0.226	0.368	0.481	0.370	0.482	0.073	0.129	0.073	0.129
KSTDiff		0.256	0.340	0.282	0.381	0.320	0.400	0.314	0.372	0.336	0.380	0.340	0.403	0.224	0.361	0.205	0.343
CGAN		0.208	0.300	0.206	0.294	0.294	0.401	0.294	0.396	0.407	0.510	0.409	0.519	0.262	0.327	0.267	0.335
Diffwave		0.244	0.310	0.239	0.342	0.312	0.352	0.296	0.363	0.360	0.462	0.405	0.492	0.175	0.236	0.142	0.214
DiT		0.259	0.300	0.228	0.283	0.292	0.332	0.306	0.354	0.321	0.366	0.312	0.355	0.169	0.231	0.204	0.244
DDPM		0.116	0.171	0.117	0.176	0.133	0.186	0.134	0.190	0.121	0.177	0.127	0.183	0.072	0.122	0.074	0.119
CVAE		0.175	0.256	0.177	0.258	0.229	0.319	0.233	0.321	0.174	0.248	0.177	0.250	0.108	0.178	0.109	0.180
CRAFT		0.109	0.165	0.111	0.167	0.111	0.166	0.116	0.171	0.124	0.179	0.127	0.180	0.066	0.111	0.067	0.108
Real	LSTM ($T = 96$)	0.108	0.161	0.110	0.163	0.094	0.144	0.098	0.149	0.117	0.175	0.119	0.175	0.062	0.111	0.063	0.110
GMEL		0.174	0.245	0.182	0.248	0.231	0.288	0.248	0.312	0.261	0.319	0.262	0.313	0.102	0.152	0.139	0.198
DFG		0.322	0.460	0.326	0.462	0.152	0.224	0.158	0.230	0.351	0.455	0.351	0.454	0.081	0.140	0.081	0.142
KSTDiff		0.371	0.502	0.376	0.505	0.416	0.540	0.494	0.617	0.480	0.608	0.480	0.606	0.234	0.372	0.235	0.371
CGAN		0.261	0.372	0.264	0.376	0.397	0.515	0.409	0.519	0.335	0.453	0.336	0.454	0.170	0.240	0.164	0.232
Diffwave		0.252	0.341	0.236	0.321	0.319	0.378	0.297	0.350	0.311	0.384	0.327	0.412	0.149	0.229	0.182	0.271
DiT		0.234	0.287	0.242	0.287	0.297	0.335	0.296	0.333	0.316	0.357	0.302	0.341	0.230	0.260	0.230	0.260
DDPM		0.116	0.172	0.119	0.177	0.122	0.179	0.122	0.184	0.128	0.183	0.135	0.190	0.074	0.125	0.074	0.118
CVAE		0.152	0.225	0.159	0.233	0.301	0.405	0.300	0.402	0.170	0.240	0.174	0.241	0.129	0.202	0.131	0.203
CRAFT		0.112	0.170	0.115	0.171	0.111	0.162	0.115	0.171	0.129	0.188	0.134	0.191	0.069	0.118	0.068	0.112
Real	LSTM ($T = 120$)	0.106	0.160	0.108	0.161	0.095	0.145	0.099	0.149	0.121	0.183	0.122	0.183	0.062	0.111	0.063	0.110
GMEL		0.184	0.243	0.187	0.256	0.213	0.260	0.218	0.271	0.277	0.333	0.262	0.312	0.099	0.152	0.105	0.155
DFG		0.306	0.445	0.311	0.448	0.168	0.249	0.172	0.252	0.352	0.467	0.356	0.472	0.089	0.154	0.089	0.154
KSTDiff		0.390	0.434	0.624	0.702	0.458	0.588	0.473	0.599	0.334	0.377	0.337	0.388	0.767	0.819	0.390	0.444
CGAN		0.275	0.374	0.283	0.387	0.293	0.404	0.286	0.395	0.314	0.409	0.313	0.410	0.136	0.228	0.137	0.227
Diffwave		0.236	0.294	0.233	0.299	0.387	0.495	0.324	0.413	0.307	0.383	0.294	0.358	0.151	0.245	0.143	0.235
DiT		0.231	0.284	0.242	0.286	0.276	0.319	0.277	0.321	0.302	0.344	0.299	0.344	0.270	0.297	0.175	0.219
DDPM		0.116	0.172	0.118	0.177	0.123	0.181	0.122	0.181	0.132	0.188	0.141	0.196	0.070	0.118	0.069	0.112
CVAE		0.151	0.217	0.155	0.222	0.214	0.304	0.220	0.310	0.166	0.239	0.164	0.234	0.133	0.215	0.130	0.208
CRAFT		0.110	0.168	0.112	0.170	0.110	0.166	0.116	0.170	0.129	0.187	0.131	0.190	0.073	0.115	0.074	0.114
Real	LSTM ($T = 144$)	0.107	0.160	0.110	0.162	0.095	0.144	0.098	0.148	0.125	0.186	0.127	0.186	0.062	0.109	0.061	0.108
GMEL		0.184	0.242	0.221	0.327	0.209	0.262	0.238	0.303	0.261	0.317	0.263	0.317	0.086	0.137	0.091	0.138
DFG		0.315	0.454	0.319	0.456	0.155	0.227	0.157	0.229	0.368	0.487	0.372	0.490	0.086	0.145	0.086	0.145
KSTDiff		0.524	0.585	0.456	0.505	0.481	0.607	0.350	0.386	0.341	0.419	0.377	0.423	0.200	0.334	0.188	0.294
CGAN		0.394	0.495	0.420	0.515	0.459	0.579	0.470	0.590	0.362	0.460	0.355	0.452	0.172	0.255	0.176	0.260
Diffwave		0.244	0.335	0.312	0.425	0.290	0.358	0.295	0.346	0.333	0.440	0.366	0.459	0.167	0.251	0.142	0.210
DiT		0.226	0.280	0.224	0.276	0.306	0.345	0.305	0.342	0.304	0.344	0.311	0.352	0.195	0.234	0.206	0.240
DDPM		0.117	0.176	0.121	0.183	0.123	0.182	0.125	0.185	0.132	0.188	0.137	0.194	0.076	0.121	0.078	0.117
CVAE		0.209	0.293	0.210	0.293	0.278	0.370	0.279	0.370	0.226	0.314	0.228	0.316	0.154	0.255	0.153	0.252
CRAFT		0.113	0.170	0.115	0.170	0.114	0.169	0.118	0.171	0.130	0.189	0.137	0.196	0.075	0.122	0.077	0.118
Real	LSTM ($T = 168$)	0.112	0.169	0.115	0.172	0.098	0.146	0.101	0.152	0.121	0.179	0.122	0.179	0.061	0.107	0.061	0.107
GMEL		0.204	0.280	0.205	0.283	0.198	0.256	0.211	0.258	0.225	0.268	0.226	0.279	0.098	0.148	0.097	0.148
DFG		0.299	0.431	0.304	0.433	0.153	0.223	0.156	0.228	0.312	0.414	0.313	0.414	0.089	0.154	0.089	0.153
KSTDiff		0.348	0.480	0.335	0.462	0.443	0.524	0.505	0.623	0.401	0.516	0.379	0.481	0.199	0.339	0.195	0.305
CGAN		0.412	0.504	0.426	0.525	0.395	0.496	0.390	0.484	0.335	0.412	0.340	0.416	0.181	0.269	0.173	0.261
Diffwave		0.238	0.306	0.227	0.295	0.334	0.388	0.444	0.573	0.307	0.354	0.331	0.407	0.166	0.260	0.147	0.220
DiT		0.235	0.296	0.236	0.295	0.289	0.326	0.285	0.324	0.315	0.351	0.317	0.357	0.218	0.250	0.168	0.214
DDPM		0.118	0.175	0.121	0.182	0.131	0.181	0.134	0.188	0.125	0.182	0.137	0.195	0.073	0.119	0.075	0.119
CVAE		0.228	0.312	0.228	0.311	0.316	0.416	0.313	0.411	0.187	0.251	0.190	0.251	0.180	0.260	0.180	0.256
CRAFT		0.114	0.170	0.117	0.173	0.111	0.165	0.115	0.171	0.124	0.180	0.133	0.188	0.078	0.128	0.078	0.122

Table 8: Data utility comparison on traffic flow prediction (Transformer)

Gen	Pred	Chicago				Washington, D.C.				Toronto				New York City			
		Inflow		Outflow		Inflow		Outflow		Inflow		Outflow		Inflow		Outflow	
		MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
Real	Transformer ($T = 48$)	0.099	0.158	0.103	0.162	0.094	0.149	0.098	0.155	0.111	0.170	0.114	0.172	0.053	0.102	0.054	0.103
GMEL		0.165	0.228	0.165	0.227	0.201	0.260	0.199	0.253	0.220	0.274	0.223	0.284	0.080	0.133	0.081	0.131
DFG		0.300	0.424	0.300	0.418	0.184	0.250	0.176	0.240	0.341	0.453	0.328	0.435	0.085	0.134	0.081	0.130
KSTDiff		0.367	0.500	0.377	0.510	0.469	0.604	0.476	0.609	0.571	0.723	0.530	0.670	0.337	0.378	0.756	0.807
CGAN		0.159	0.221	0.171	0.237	0.282	0.379	0.284	0.382	0.239	0.310	0.235	0.304	0.128	0.193	0.135	0.201
Diffwave		0.220	0.292	0.212	0.279	0.278	0.345	0.284	0.354	0.291	0.356	0.285	0.347	0.156	0.228	0.140	0.211
DiT		0.235	0.289	0.236	0.287	0.269	0.312	0.282	0.320	0.296	0.339	0.298	0.337	0.181	0.225	0.158	0.212
DDPM		0.111	0.169	0.115	0.173	0.113	0.166	0.115	0.172	0.119	0.174	0.123	0.178	0.063	0.111	0.065	0.110
CVAE	Transformer ($T = 72$)	0.162	0.228	0.164	0.229	0.230	0.311	0.237	0.318	0.172	0.222	0.176	0.226	0.114	0.181	0.116	0.183
CRAFT		0.097	0.148	0.100	0.152	0.097	0.149	0.105	0.161	0.124	0.174	0.124	0.173	0.060	0.103	0.060	0.102
Real	Transformer ($T = 96$)	0.102	0.159	0.105	0.162	0.094	0.147	0.098	0.153	0.116	0.176	0.118	0.177	0.055	0.100	0.056	0.100
GMEL		0.186	0.251	0.186	0.258	0.249	0.311	0.240	0.298	0.252	0.303	0.248	0.298	0.083	0.138	0.083	0.132
DFG		0.362	0.491	0.350	0.486	0.166	0.222	0.166	0.222	0.391	0.511	0.390	0.509	0.069	0.116	0.070	0.118
KSTDiff		0.349	0.514	0.364	0.519	0.383	0.511	0.376	0.488	0.356	0.401	0.352	0.413	0.216	0.355	0.207	0.359
CGAN		0.179	0.243	0.187	0.249	0.254	0.345	0.252	0.333	0.391	0.493	0.359	0.458	0.146	0.190	0.149	0.194
Diffwave		0.239	0.297	0.228	0.294	0.299	0.352	0.279	0.328	0.324	0.369	0.339	0.382	0.143	0.211	0.143	0.209
DiT		0.243	0.291	0.225	0.281	0.285	0.328	0.287	0.331	0.321	0.364	0.311	0.354	0.177	0.234	0.188	0.234
DDPM		0.108	0.161	0.113	0.167	0.132	0.187	0.131	0.186	0.128	0.184	0.130	0.185	0.067	0.117	0.070	0.118
CVAE	Transformer ($T = 120$)	0.175	0.243	0.176	0.244	0.224	0.305	0.227	0.306	0.192	0.255	0.198	0.259	0.106	0.165	0.109	0.170
CRAFT		0.103	0.155	0.106	0.158	0.106	0.158	0.109	0.161	0.121	0.174	0.124	0.176	0.060	0.104	0.060	0.104
Real	Transformer ($T = 144$)	0.105	0.162	0.108	0.165	0.093	0.143	0.096	0.147	0.115	0.175	0.117	0.176	0.059	0.105	0.059	0.104
GMEL		0.186	0.256	0.180	0.240	0.218	0.276	0.232	0.294	0.231	0.281	0.236	0.286	0.091	0.141	0.145	0.184
DFG		0.341	0.467	0.323	0.442	0.175	0.237	0.168	0.228	0.354	0.467	0.360	0.472	0.068	0.116	0.069	0.117
KSTDiff		0.387	0.526	0.397	0.538	0.420	0.546	0.492	0.616	0.448	0.564	0.484	0.604	0.243	0.386	0.236	0.375
CGAN		0.195	0.268	0.197	0.267	0.362	0.447	0.385	0.476	0.300	0.388	0.296	0.380	0.124	0.175	0.133	0.187
Diffwave		0.240	0.297	0.235	0.304	0.272	0.320	0.274	0.319	0.294	0.342	0.297	0.348	0.144	0.209	0.139	0.205
DiT		0.231	0.287	0.227	0.280	0.285	0.327	0.278	0.321	0.308	0.346	0.297	0.339	0.146	0.208	0.151	0.208
DDPM		0.114	0.169	0.118	0.173	0.122	0.175	0.121	0.176	0.130	0.183	0.138	0.191	0.070	0.116	0.071	0.115
CVAE	Transformer ($T = 168$)	0.153	0.214	0.158	0.219	0.255	0.332	0.261	0.336	0.170	0.230	0.174	0.231	0.113	0.169	0.117	0.169
CRAFT		0.106	0.160	0.108	0.162	0.104	0.150	0.108	0.158	0.122	0.178	0.128	0.183	0.063	0.109	0.064	0.108
Real	Transformer ($T = 192$)	0.100	0.155	0.103	0.159	0.090	0.140	0.093	0.144	0.116	0.175	0.118	0.177	0.057	0.105	0.057	0.105
GMEL		0.192	0.256	0.193	0.261	0.222	0.290	0.232	0.304	0.249	0.305	0.250	0.306	0.094	0.150	0.098	0.156
DFG		0.310	0.425	0.303	0.413	0.150	0.203	0.151	0.208	0.349	0.460	0.354	0.465	0.071	0.122	0.072	0.123
KSTDiff		0.390	0.436	0.621	0.699	0.509	0.654	0.507	0.640	0.328	0.380	0.329	0.386	0.721	0.771	0.341	0.420
CGAN		0.270	0.369	0.282	0.390	0.283	0.375	0.288	0.381	0.273	0.349	0.261	0.338	0.108	0.175	0.109	0.173
Diffwave		0.221	0.281	0.222	0.284	0.270	0.342	0.262	0.321	0.309	0.350	0.294	0.343	0.140	0.207	0.137	0.202
DiT		0.222	0.280	0.222	0.276	0.265	0.318	0.268	0.315	0.306	0.349	0.303	0.347	0.173	0.220	0.149	0.208
DDPM		0.109	0.161	0.113	0.166	0.112	0.163	0.113	0.166	0.129	0.184	0.135	0.190	0.067	0.111	0.066	0.109
CVAE	Transformer ($T = 216$)	0.153	0.214	0.158	0.220	0.202	0.272	0.212	0.283	0.170	0.231	0.173	0.231	0.106	0.171	0.105	0.168
CRAFT		0.107	0.162	0.109	0.163	0.103	0.152	0.107	0.156	0.120	0.173	0.123	0.177	0.063	0.104	0.063	0.104
Real	Transformer ($T = 252$)	0.101	0.155	0.104	0.159	0.091	0.140	0.093	0.144	0.118	0.177	0.120	0.180	0.055	0.104	0.056	0.104
GMEL		0.211	0.281	0.209	0.285	0.242	0.301	0.234	0.284	0.264	0.317	0.263	0.312	0.088	0.140	0.090	0.142
DFG		0.323	0.450	0.307	0.427	0.188	0.248	0.185	0.247	0.399	0.521	0.391	0.510	0.069	0.120	0.069	0.119
KSTDiff		0.524	0.586	0.459	0.510	0.488	0.618	0.369	0.423	0.367	0.462	0.385	0.446	0.190	0.315	0.166	0.239
CGAN		0.342	0.431	0.392	0.480	0.371	0.463	0.352	0.449	0.284	0.360	0.288	0.368	0.133	0.186	0.135	0.184
Diffwave		0.221	0.280	0.215	0.273	0.291	0.334	0.296	0.353	0.290	0.357	0.288	0.359	0.139	0.205	0.136	0.202
DiT		0.223	0.279	0.221	0.276	0.290	0.331	0.283	0.322	0.297	0.338	0.301	0.342	0.147	0.208	0.143	0.203
DDPM		0.113	0.167	0.117	0.173	0.122	0.173	0.124	0.176	0.127	0.181	0.131	0.185	0.071	0.113	0.069	0.110
CVAE	Transformer ($T = 288$)	0.192	0.260	0.193	0.261	0.277	0.359	0.283	0.363	0.234	0.297	0.240	0.305	0.124	0.198	0.124	0.196
CRAFT		0.107	0.160	0.110	0.163	0.106	0.154	0.111	0.159	0.125	0.180	0.129	0.183	0.066	0.107	0.065	0.105
Real	Transformer ($T = 336$)	0.104	0.160	0.108	0.164	0.090	0.140	0.093	0.144	0.115	0.173	0.118	0.176	0.054	0.102	0.055	0.103
GMEL		0.216	0.288	0.217	0.287	0.219	0.281	0.218	0.277	0.238	0.301	0.241	0.297	0.093	0.148	0.095	0.150
DFG		0.315	0.427	0.313	0.424	0.162	0.217	0.167	0.226	0.377	0.493	0.374	0.492	0.071	0.122	0.072	0.123
KSTDiff		0.344	0.473	0.340	0.467	0.444	0.526	0.510	0.625	0.395	0.508	0.374	0.473	0.199	0.339	0.196	0.308
CGAN		0.385	0.469	0.385	0.476	0.324	0.418	0.323	0.412	0.319	0.385	0.346	0.413	0.140	0.200	0.145	0.210
Diffwave		0.228	0.288	0.221	0.283	0.281	0.328	0.267	0.321	0.290	0.339	0.289	0.341	0.142	0.206	0.136	0.202
DiT		0.231	0.288	0.227	0.282	0.283	0.325	0.267	0.318	0.305	0.342	0.308	0.347	0.192	0.231	0.154	0.206
DDPM		0.115	0.169	0.119	0.174	0.129	0.178	0.132	0.182	0.127	0.178	0.136	0.187	0.071	0.112	0.069	0.110
CVAE	Transformer ($T = 384$)	0.193	0.257	0.195	0.259	0.285	0.369	0.287	0.370	0.196	0.247	0.202	0.252	0.138	0.189	0.138	0.188
CRAFT		0.106	0.158	0.108	0.162	0.102	0.149	0.106	0.154	0.122	0.176	0.125	0.179	0.069	0.114	0.067	0.111

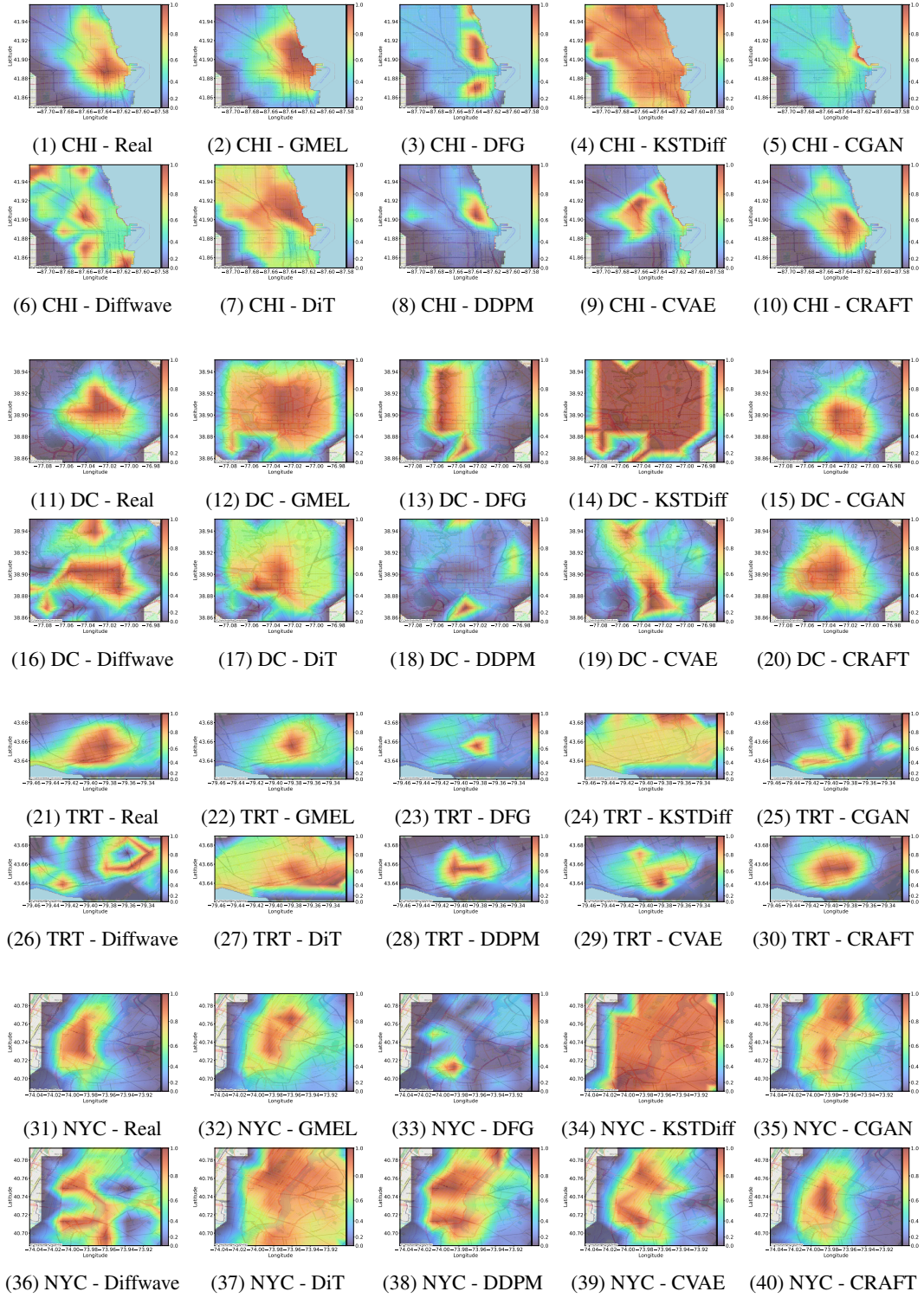


Figure 11: Heatmap of Traffic Flow in Different Cities (CHI stands for Chicago, DC stands for Washington, D.C., TRT stands for Toronto and NYC stands for New York City)

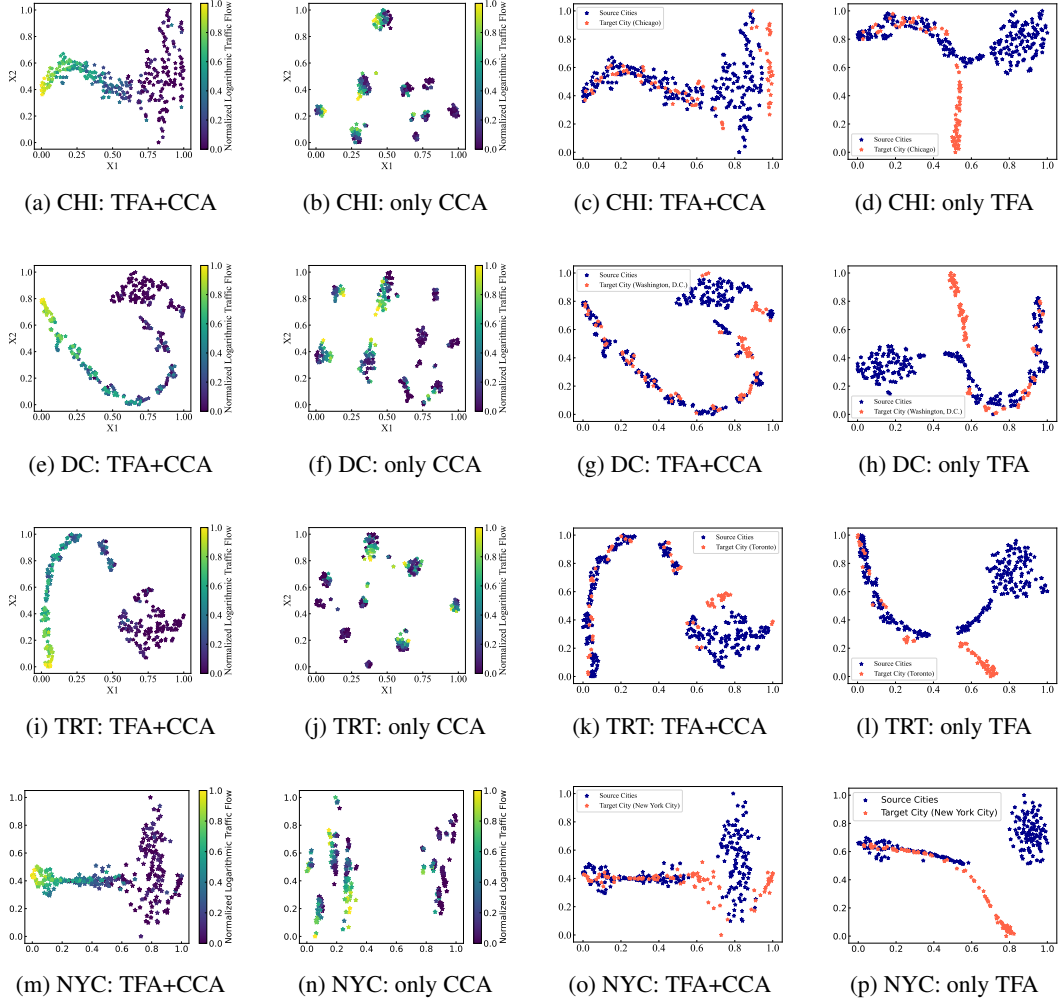


Figure 12: Visualization analysis for TFA and CCA (CHI stands for Chicago, DC stands for Washington, D.C., TRT stands for Toronto and NYC stands for New York City)