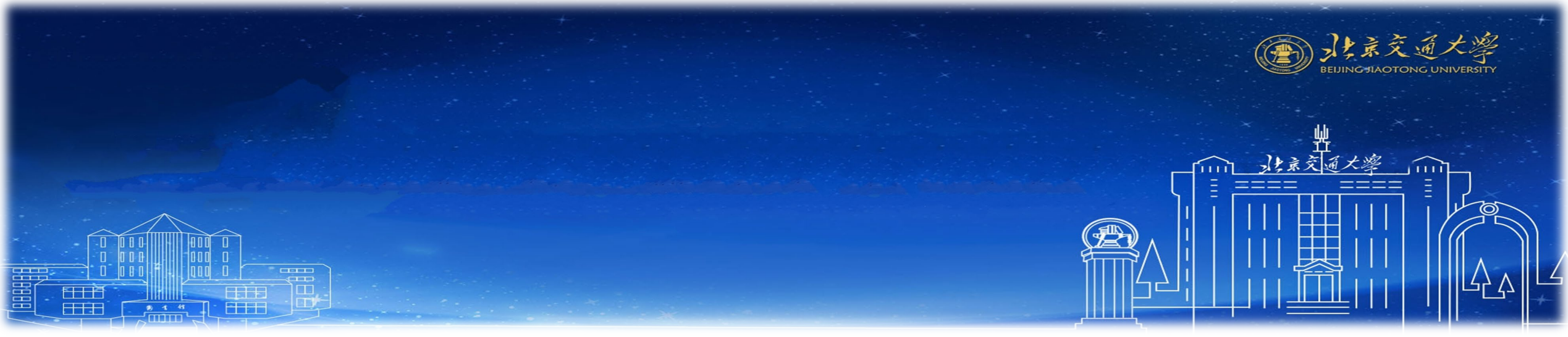
WITRAN——重塑RNN在长程时间序列预测上的应用

北京交通大学　贾宇欣　2023-11-04



一、研究背景与意义

- **时间序列预测的重要性：**
  - 时间序列预测是可用于能源、交通、天气等多个领域的实用工具

- **长程时间序列预测的重要性：**
  - 与时间序列短程预测相比，长程预测可以便于人们有充足的时间进行准备，并进一步做好决策工作。

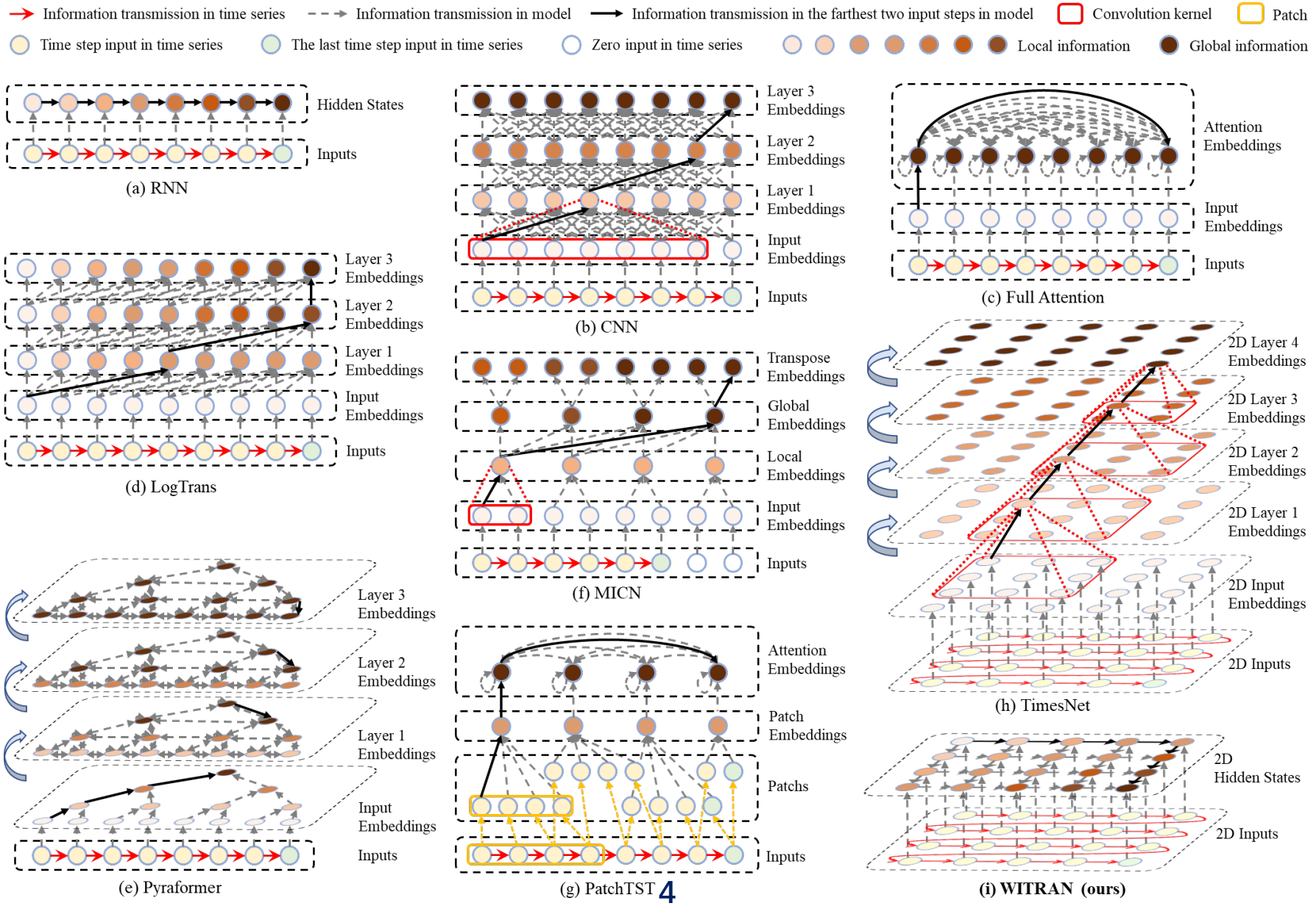
避开
拥堵路段

提前规划
出行以及
穿衣

二、研究动机

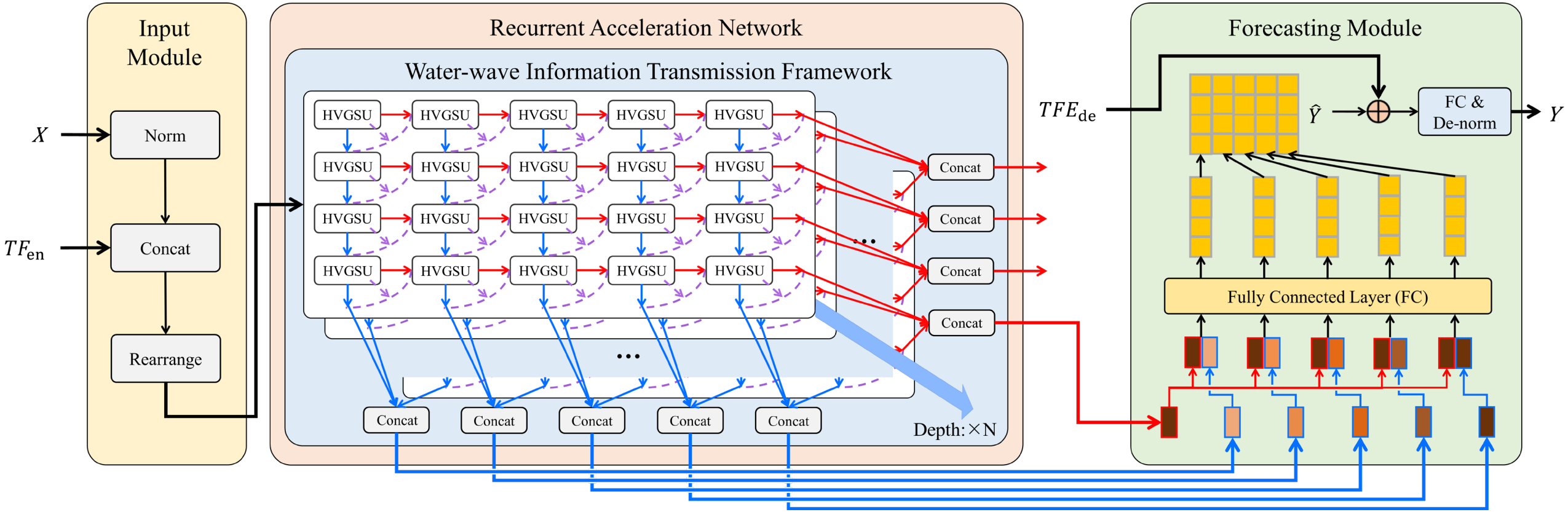
- **数据的充分性：**
  - 应有足够长的历史输入数据，以提升预测效果
- **语义信息的重要性：**
  - （1）全局-局部语义信息：局部语义信息通常包含的是序列的短程变化，全局语义信息通常反映的是序列的长程趋势。
  - （2）长期和短期周期性语义信息：时间序列通常在不同的时间粒度上均表现出重复模式，例如：小时周期性，天周期性等。
- **运算的高效性：**
  - 在保证准确率的同时，降低复杂度。

三、研究现状

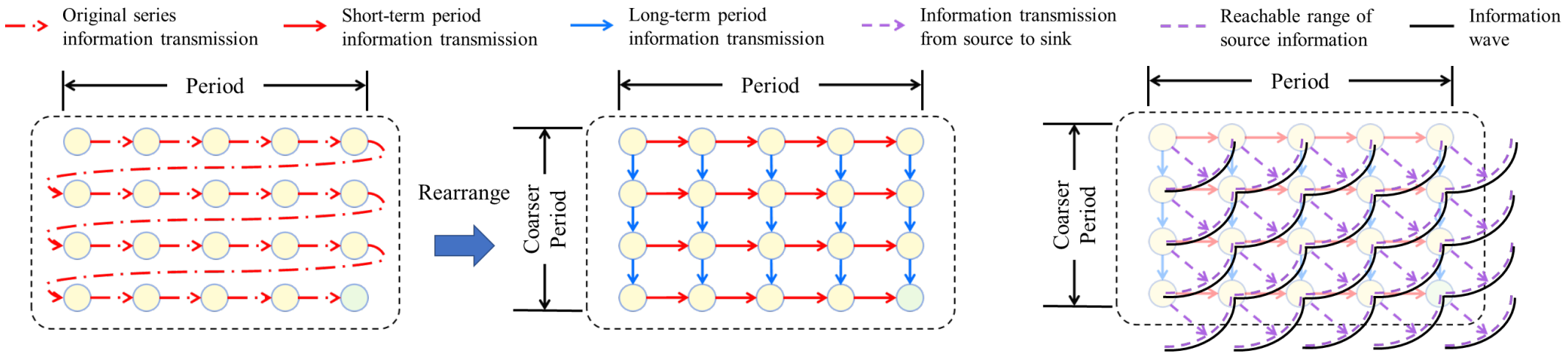

Information transmission in time series
Information transmission in model
Information transmission in the farthest two input steps in model
Convolution kernel
Patch
Time step input in time series
The last time step input in time series
Zero input in time series
Local information
Global information
Hidden States
Inputs
(a) RNN
Layer 3 Embeddings
Layer 2 Embeddings
Layer 1 Embeddings
Input Embeddings
Inputs
(b) CNN
Attention Embeddings
Input Embeddings
Inputs
(c) Full Attention
(d) LogTrans
Transpose Embeddings
Global Embeddings
Local Embeddings
(f) MICN
2D Layer 4 Embeddings
2D Layer 3 Embeddings
2D Layer 2 Embeddings
2D Layer 1 Embeddings
2D Input Embeddings
2D Inputs
(h) TimesNet
(e) Pyraformer
Patch Embeddings
Patchs
(g) PatchTST
2D Hidden States
(i) WITRAN (ours)

四、模型设计

- **输入：** 以历史H个时间步作为模型输入 $X = \{x_1, x_2, \dots, x_H\} \in \mathbb{R}^{H \times c_{in}}$
- **输出：** 对未来P个时间步进行预测 $Y \in \mathbb{R}^{P \times c_{out}}$

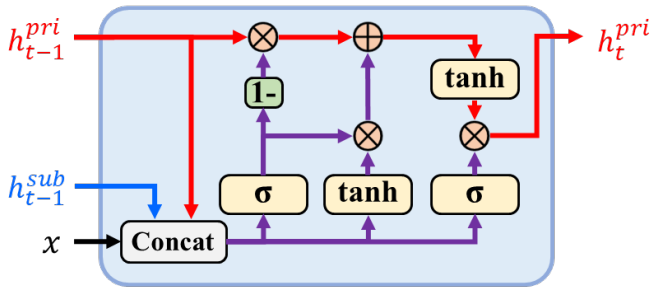


四、模型设计

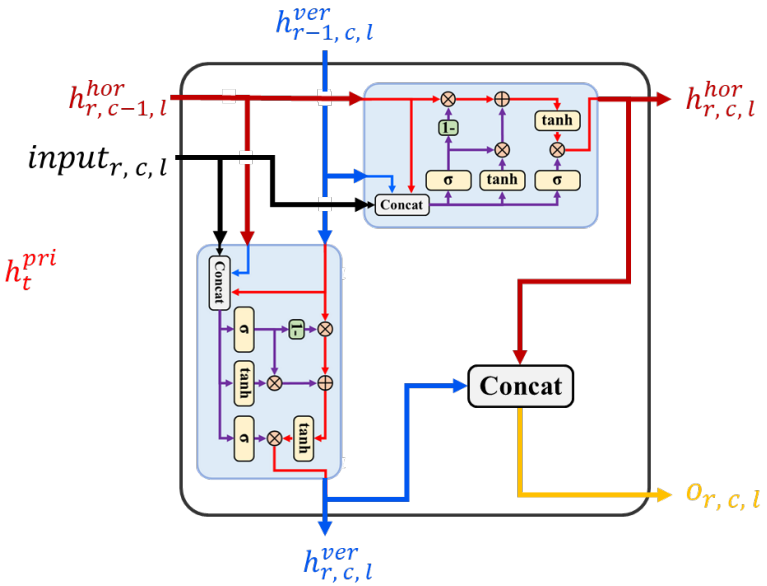


(a) Information transmission before and after the input sequence rearrangement

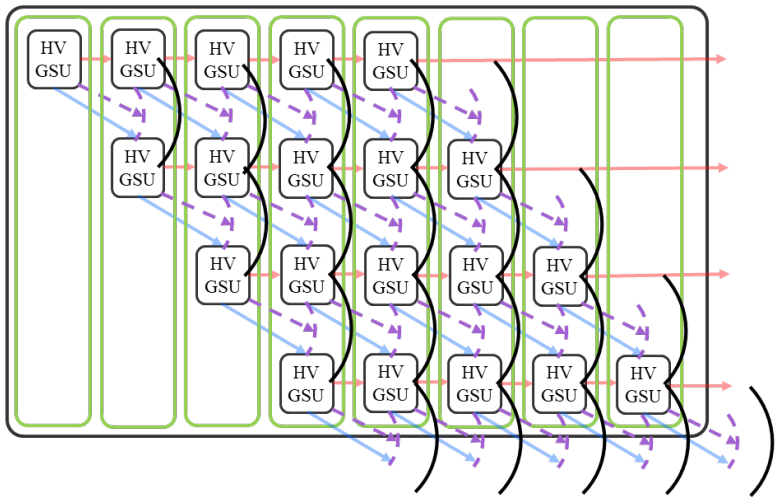
(b) The water-wave information transmission process



(c) GSC structure



(d) HVGSU structure



(e) RAN structure

五、实验设置

- **数据集：** ECL、traffic、ETTh、Weather
- **划分方式：** 按照时间先后按6:2:2划分训练集、验证集、测试集
- **任务设置：** 长程预测任务、超长程预测任务
- **对比方法：**

为了公平比较每个模型的性能，我们设置了相同的搜索空间，以便于每个模型在各个任务上都能达到最佳性能。

Dataset	Sample Length	Dimension	Usage Frequency	Recorded Frequency
ETTh	17420	8	1h	15min
ECL	26304	322	1h	1h
Traffic	17544	863	1h	1h
Weather	35064	22	1h	10min

六、实验设置

NEURAL INFORMATION PROCESSING SYSTEMS

➢长程序列预测任务：WITRAN较次优方法，MSE上降低5.803%

Methods		WITRAN(Ours)		MICN		TimesNet		PatchTST		DLinear		FiLM		FEDformer		Pyraformer		Autoformer		Informer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	168-168	**0.2397**	**0.3519**	0.3168	0.4067	0.2825	0.3797	0.2980	0.3832	0.2605	0.3579	0.2587	0.3557	0.3028	0.4020	0.2651	0.3802	0.3496	0.4337	0.3779	0.4594	0.3036	0.4068
	168-336	**0.2607**	**0.3721**	0.3002	0.4053	0.3505	0.4253	0.3840	0.4393	0.3080	0.3946	0.3062	0.3922	0.3522	0.4394	0.5392	0.5271	0.4733	0.5120	0.5037	0.5301	0.3583	0.4435
	336-336	**0.2517**	**0.3627**	0.3092	0.4132	0.3702	0.4307	0.4377	0.4654	0.2740	0.3720	0.2722	0.3659	0.3378	0.4303	0.2994	0.4030	0.5153	0.5304	0.4591	0.4991	0.5771	0.5643
	336-720	**0.3084**	**0.4055**	0.3820	0.4704	0.3879	0.4531	0.5502	0.5438	0.3208	0.4188	0.3171	0.4152	0.3813	0.4634	0.4856	0.5243	0.5045	0.5393	0.6545	0.5975	0.4368	0.4920
	720-720	**0.2478**	**0.3651**	0.3463	0.4381	0.3537	0.4386	0.5927	0.5742	0.3203	0.4202	0.3158	0.4154	0.4023	0.4769	0.3115	0.4218	0.9639	0.7520	0.4850	0.5238	0.3992	0.4640
Traffic	168-168	**0.1377**	**0.2051**	0.2428	0.3543	0.1490	0.2293	0.1622	0.2320	0.1519	0.2195	0.1501	0.2143	0.2469	0.3479	0.2979	0.3815	0.2378	0.3490	0.3363	0.3994	1.5204	0.9594
	168-336	**0.1321**	**0.2059**	0.2401	0.3514	0.1499	0.2356	0.1641	0.2364	0.1468	0.2210	0.1453	0.2165	0.2426	0.3449	0.5838	0.5652	0.2683	0.3803	0.5891	0.5608	0.6953	0.6015
	336-336	**0.1306**	**0.2054**	0.2413	0.3549	0.1446	0.2300	0.1546	0.2332	0.1325	0.2114	0.1324	0.2104	0.2339	0.3365	0.4703	0.4964	0.2460	0.3567	0.5447	0.5384	0.8482	0.6424
	336-720	**0.1391**	**0.2175**	0.2422	0.3513	0.1584	0.2440	0.1747	0.2536	0.1449	0.2252	0.1438	0.2229	0.2987	0.3976	0.5235	0.5292	0.2849	0.3956	1.2044	0.8254	0.7320	0.6233
	720-720	0.1408	**0.2191**	0.2552	0.3709	0.1546	0.2410	0.1543	0.2441	0.1410	0.2241	**0.1383**	0.2208	0.2667	0.3685	0.4811	0.4962	0.2959	0.4045	1.2954	0.9205	1.1963	0.8271
$ETTh_1$	168-168	0.1105	0.2589	0.1257	0.2803	0.1133	0.2612	0.1212	0.2704	0.1122	0.2605	**0.1091**	**0.2558**	0.1284	0.2826	0.1534	0.3287	0.1318	0.2872	0.1563	0.3299	0.1504	0.3257
	168-336	0.1189	0.2714	0.1422	0.3006	0.1202	0.2732	0.1287	0.2808	0.1251	0.2794	**0.1187**	**0.2708**	0.1271	0.2810	0.1665	0.3419	0.1315	0.2878	0.1663	0.3335	0.1599	0.3324
	336-336	**0.1112**	**0.2638**	0.1576	0.3159	0.1279	0.2846	0.1496	0.3039	0.1261	0.2803	0.1196	0.2738	0.1252	0.2794	0.1408	0.3087	0.1384	0.2959	0.1648	0.3291	0.1438	0.3121
	336-720	**0.1494**	**0.3092**	0.2219	0.3729	0.1501	0.3127	0.2092	0.3659	0.1942	0.3462	0.1793	0.3335	0.1534	0.3178	0.3984	0.5202	0.1928	0.3450	0.1522	0.3203	0.1644	0.3304
	720-720	**0.1296**	**0.2868**	0.2959	0.4402	0.1510	0.3118	0.2178	0.3694	0.1920	0.3435	0.1845	0.3379	0.1386	0.2995	0.1563	0.3253	0.2388	0.3869	0.1595	0.3259	0.1730	0.3414
$ETTh_2$	168-168	**0.2389**	**0.3813**	0.2734	0.4162	0.2655	0.4051	0.2582	0.3983	0.2556	0.3944	0.2546	0.3942	0.2844	0.4285	0.2746	0.4080	0.2903	0.4326	0.3764	0.4863	0.3043	0.4365
	168-336	**0.2277**	**0.3778**	0.3017	0.4429	0.2725	0.4163	0.3206	0.4515	0.2891	0.4256	0.2894	0.4263	0.2961	0.4355	0.2392	0.3834	0.4447	0.4964	0.3364	0.4583	0.3662	0.4671
	336-336	**0.2432**	**0.3922**	0.3472	0.4796	0.3184	0.4431	0.3559	0.4779	0.2950	0.4329	0.2951	0.4347	0.2884	0.4314	0.2610	0.4010	0.2805	0.4255	0.3709	0.4785	0.3218	0.4412
	336-720	0.2373	0.3888	0.4248	0.5268	0.2858	0.4253	0.4936	0.5592	0.4125	0.5136	0.4158	0.5162	0.3425	0.4656	**0.2341**	**0.3818**	0.3372	0.4625	0.3572	0.4675	0.3582	0.4629
	720-720	**0.2635**	**0.4018**	0.3549	0.4805	0.2936	0.4238	0.5243	0.5745	0.3495	0.4749	0.4045	0.5105	0.3275	0.4534	0.2795	0.4151	0.4668	0.5477	0.3585	0.4699	0.3087	0.4320
Weather	168-168	**0.2050**	**0.3338**	0.2231	0.3489	0.2420	0.3608	0.2469	0.3597	0.2421	0.3578	0.2426	0.3544	0.2583	0.3774	0.2144	0.3451	0.2670	0.3813	0.2639	0.3926	0.2200	0.3438
	168-336	**0.2197**	**0.3470**	0.2663	0.3837	0.2821	0.3885	0.3040	0.4049	0.2918	0.3975	0.2981	0.3988	0.2909	0.4030	0.2594	0.3833	0.2990	0.4096	0.2798	0.4061	0.2230	0.3488
	336-336	**0.2163**	**0.3482**	0.2701	0.3804	0.2684	0.3752	0.3149	0.4145	0.2905	0.3969	0.2943	0.3969	0.2791	0.3984	0.2310	0.3591	0.3066	0.4162	0.2898	0.4129	0.2308	0.3556
	336-720	**0.2054**	**0.3424**	0.3086	0.4138	0.2930	0.4045	0.4358	0.4937	0.3897	0.4739	0.4096	0.4767	0.2648	0.3915	0.3241	0.4300	0.3468	0.4592	0.2483	0.3778	0.2334	0.3570
	720-720	**0.2008**	**0.3417**	0.2828	0.3969	0.2967	0.4070	0.5701	0.5491	0.3724	0.4614	0.3999	0.4661	0.2416	0.3728	0.2378	0.3684	0.4309	0.5085	0.3545	0.4569	0.2463	0.3722

六、实验设置

NEURAL INFORMATION PROCESSING SYSTEMS

➢超长程序列预测任务：WITRAN较次优方法，MSE上降低14.275%

Methods		WITRAN(Ours)		MICN		TimesNet		PatchTST		DLinear		FiLM		FEDformer		Pyraformer		Autoformer		Informer		Transformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ECL	720-1440	**0.2499**	**0.3727**	1.0460	0.7765	0.6119	0.5962	0.8243	0.6704	0.4923	0.5473	0.4730	0.5336	0.4833	0.5393	0.3250	0.4332	1.4957	0.9533	0.5064	0.5317	0.4030	0.4797
	1440-1440	**0.2408**	**0.3680**	2.2862	1.2207	0.5720	0.5712	0.9053	0.7328	0.5146	0.5615	0.4849	0.5429	0.5142	0.5571	0.4895	0.5280	1.7873	1.0283	0.7247	0.6292	0.5531	0.5524
	1440-2880	**0.3359**	**0.4383**	2.8936	1.3717	0.7683	0.6846	1.1282	0.8087	0.8355	0.7193	0.6847	0.6493	3.9018	1.5276	0.4320	0.5161	1.2867	0.8878	0.6152	0.5953	0.5243	0.5460
Traffic	720-1440	0.1672	0.2449	0.2876	0.3916	0.1882	0.2656	0.1904	0.2685	0.1639	**0.2412**	**0.1638**	0.2448	0.2753	0.3650	0.4463	0.4609	0.3104	0.4095	0.7614	0.6469	0.9876	0.7445
	1440-1440	**0.1543**	**0.2325**	0.2950	0.3923	0.1598	0.2388	0.1817	0.2764	0.1599	0.2411	0.1602	0.2437	0.2848	0.3681	0.4710	0.4916	0.2970	0.3999	0.7375	0.6414	0.7430	0.6492
	1440-2880	**0.1425**	**0.2333**	0.2823	0.3874	0.1560	0.2409	0.2029	0.3100	0.1550	0.2472	0.1744	0.2693	0.2952	0.3844	0.5165	0.5305	0.3035	0.3982	0.9849	0.7618	0.6000	0.5877
$ETTh_1$	720-1440	**0.1331**	**0.2943**	0.4640	0.5836	0.1391	0.3049	0.3708	0.4906	0.2952	0.4370	0.2949	0.4388	0.1768	0.3409	0.1666	0.3315	0.3298	0.4741	0.1378	0.3051	0.1905	0.3555
	1440-1440	**0.1304**	**0.2902**	0.5650	0.6293	0.1801	0.3372	0.4475	0.5329	0.2200	0.3714	0.2294	0.3759	0.3574	0.4878	0.3487	0.4866	0.4531	0.5507	0.1430	0.3156	0.1972	0.3630
	1440-2880	**0.1850**	**0.3452**	0.7591	0.7215	0.2732	0.4094	0.9617	0.8271	0.3773	0.4794	0.6834	0.7096	0.4269	0.5252	0.5857	0.6760	1.3566	0.9235	0.3177	0.4733	0.3495	0.4911
$ETTh_2$	720-1440	**0.2915**	**0.4289**	0.4922	0.5649	0.4186	0.5092	0.9401	0.7680	0.5037	0.5645	0.7166	0.6628	0.3731	0.4827	0.2952	0.4336	0.5633	0.5996	0.4025	0.4991	0.3712	0.4805
	1440-1440	**0.2815**	**0.4220**	0.5030	0.5644	0.4409	0.5218	0.7860	0.6704	0.5176	0.5734	0.7446	0.6590	0.3906	0.4951	0.2946	0.4316	0.8029	0.7140	0.3484	0.4786	0.3797	0.4818
	1440-2880	**0.3280**	0.4585	0.5549	0.5886	1.5304	0.9026	2.0561	1.1595	0.5053	0.5584	3.2835	1.6030	1.7167	0.9698	0.3345	0.4544	4.1031	1.7198	0.3335	**0.4482**	0.3737	0.4787
Weather	720-1440	**0.1872**	**0.3312**	0.3999	0.4848	0.2407	0.3694	0.5453	0.5631	0.4406	0.5264	0.6360	0.5997	0.2352	0.3733	0.6810	0.6352	0.8599	0.7064	0.2466	0.3849	0.2188	0.3512
	1440-1440	**0.1907**	**0.3366**	0.2873	0.4201	0.2869	0.4033	0.5371	0.5559	0.3147	0.4417	0.6002	0.5880	0.2226	0.3609	0.2401	0.3777	0.9766	0.7739	0.2556	0.3969	0.2610	0.3823
	1440-2880	**0.1769**	**0.3257**	0.3570	0.4810	0.2199	0.3563	0.9061	0.7220	0.3197	0.4533	1.2605	0.8805	0.2138	0.3599	0.1852	0.3332	1.7465	1.0962	0.2126	0.3600	0.1993	0.3436

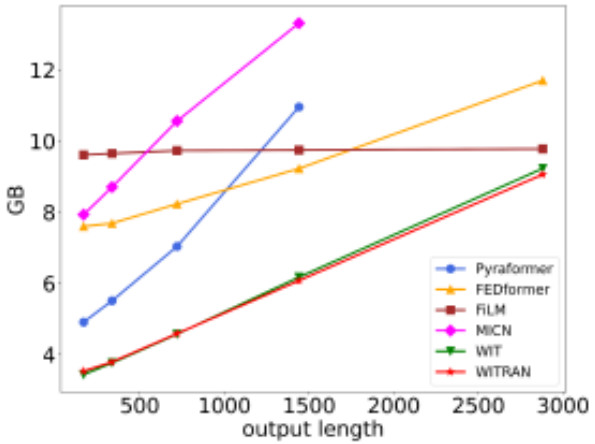
六、实验设置

➢实际运行速度和空间占用

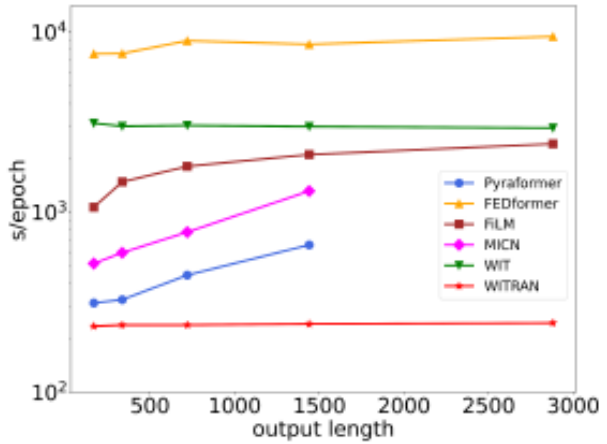
方法	Pyraformer	FEDformer	FiLM	MICN	WITRAN
时间复杂度	$O(L)$	$O(L)$	$O(L)$	$O(L)$	$O(\sqrt{L})$
空间复杂度	$O(L)$	$O(L)$	$O(L)$	$O(L)$	$O(L)$

输入固定为720，改变输出长度

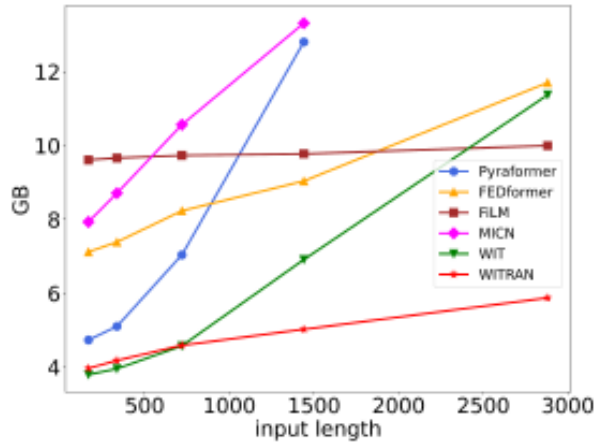
输出固定为720，改变输入长度



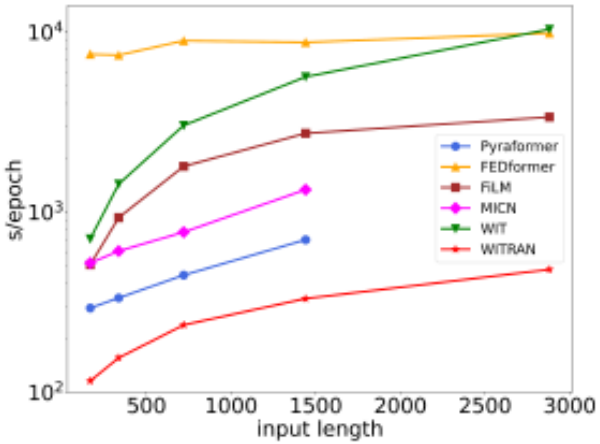
(a) Memory consumption (input length is fixed as 720)



(b) Time consumption (input length is fixed as 720)



(c) Memory consumption (output length is fixed as 720)



(d) Time consumption (output length is fixed as 720)

七、总结与展望

- 数据的充分性：
  - WITRAN证明了，将更长的历史输入数据利用好，是可以提升预测效果的
- 语义信息的重要性：
  - WITRAN作为一种全新的信息传递范式，可以同时捕获全局-局部语义信息和长期-短期周期性语义信息，从而提升预测效果
- 运算的高效性：
  - RAN在保持$O(L)$空间复杂度的同时，将时间复杂度降为$O(\sqrt{L})$。

WITRAN作为一种RNN的全新变体，

重塑了RNN在长程时间序列预测上的应用！

NEURAL INFORMATION PROCESSING SYSTEMS

WITRAN: Water-wave Information Transmission and Recurrent Acceleration Network for Long-range Time Series Forecasting

谢谢大家！

Paper：https://neurips.cc/virtual/2023/poster/69972

Code：https://github.com/Water2sea/WITRAN