

L Further Related Work

In this section, we delve deeper into the related work on (medical) time series foundation models. Current research efforts in universal forecasting with time series foundation models can be broadly classified into four categories, as summarized in Table 13:

(i) **Encoder-only models.** Moirai is trained on the LOTSA dataset comprising 27B time points, with model sizes up to 311M parameters [23]. MOMENT, based on the T5 architecture, is pretrained on the Time-series Pile dataset containing approximately 1B time points, reaching up to 385M parameters [78].

(ii) **Encoder-decoder models,** exemplified by Chronos [28], which tokenizes time series data via scaling and quantization, training on both public and synthetic datasets, offers pre-trained models at multiple scales, with the largest containing up to 710M parameters.

(iii) **Decoder-only models.** TimesFM is trained on a corpus of 100B time points, with model sizes up to 500M parameters [25]. Lag-Llama focuses on univariate probabilistic forecasting, utilizing a decoder-only Transformer architecture with up to 200M parameters [81]. Timer is a generative pre-trained Transformer model designed for large-scale time series modeling, with a base version containing 84M parameters and pre-trained on 260B time points [80].

(iv) **Mixture-of-Experts architectures.** Recent models adopt sparse Mixture-of-Experts (MoE) architectures to enhance scalability and efficiency. Time-MoE [24] scales to 2.4B parameters with only a few experts activated per input, while Moirai-MoE [26] achieves token-level specialization without frequency heuristics, improving adaptability and inference cost.

Table 13: Comparison between time series models.

Method	MIRA (Ours)	Sundial (2025)	Time-MoE (2024)	Moirai-MoE (2024)	Moirai (2024)	TimesFM (2024)	Moment (2024)	Chronos (2024)	Timer (2024)	Lag-Llama (2023)	TimeGPT (2023)
Architecture	Decoder	Decoder	Decoder	Decoder	Encoder	Decoder	Encoder	EncDec	Decoder	Decoder	EncDec
(Max) Model Size	2.4B	444M	2.4B	1.1B	311M	500M	385M	710M	67M	200M	Unknown
Input Token	Point	Patch	Point	Patch	Patch	Patch	Patch	Point	Patch	Point	Patch
Dataset Scale	454B	1TB	309B	27B	27B/231B*	100B	1.13B	84B	29B	0.36B	100B
Max Length	512	2880	4096	5000	5000	512	512	512	1440	1024	Unknown
FFN	Sparse	Dense	Sparse	Sparse	Dense	Dense	Dense	Dense	Dense	Dense	Dense

* Depend on the way of calculation according to the original paper.

M Downstream Datasets Statistics

Table 14: Characteristics of the WHO FluNet Dataset.

Dataset	Time Points	Missing Rate	Frequency	Irregular Sampling	Signal Type
WHO FluNet	5.27M	61.52%	Weekly		Case Counts
CINC 2012	5.28M	0%	Variable-specific	✓	ECG, Blood Pressure (ABP/NIBP), Respiration Rate, SpO ₂ , Laboratory Measurements, etc.
MIT-BIH	31.1M	0%	360 Hz		ECG
CinC 2019	39.5M	68.28%	Hourly	✓	Heart Rate, Respiration Rate, Temperature, Blood Pressure
Johns Hopkins COVID-19 Dataset	3.81M	0%	Daily		Case Counts, Death Counts
CDC-Influenza Hospitalizations Admissions	3.07M	27.1%	Weekly		Hospitalized Case Counts

N Further Baseline Result

Table 15: Zero-shot forecasting performance on out-of-distribution datasets. Reported values are predicted by zero-shot forecasting model pre-trained on general corpora.

Models		Lag-llama		TimeGPT		Timer		Moment _{small}		Moment _{base}		Moment _{large}		Moirai-MoE _{small}		Moirai-MoE _{base}	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Heart Rate (10 ⁻¹)	24	1.643	1.385	1.668	1.463	1.899	1.769	3.231	2.035	3.218	2.025	3.204	2.019	1.443	1.137	1.636	1.321
	32	1.702	1.433	2.090	1.755	1.898	1.721	2.985	1.849	2.978	1.857	2.949	1.479	1.777	1.432	1.942	1.595
	48	1.813	1.528	2.592	2.198	1.898	1.667	2.787	1.738	2.764	1.726	2.734	1.711	2.374	1.927	2.523	2.095
	64	1.895	1.605	2.683	2.245	1.907	1.657	2.861	1.784	2.796	1.732	2.781	1.731	2.347	1.906	2.476	1.595
	Avg.	1.764	1.488	2.258	1.915	1.901	1.704	2.966	1.852	2.939	1.835	2.917	1.735	1.833	1.844	2.013	1.651
MIT-BIH	24	0.162	0.132	0.138	0.118	0.179	0.161	0.453	0.258	0.451	0.257	0.450	0.256	0.165	0.135	0.165	0.135
	32	0.185	0.148	0.175	0.141	0.219	0.184	0.407	0.216	0.405	0.214	0.402	0.211	0.185	0.152	0.185	0.151
	48	0.253	0.192	0.291	0.227	0.330	0.278	0.407	0.224	0.406	0.223	0.402	0.218	0.231	0.183	0.231	0.183
	64	0.268	0.203	0.321	0.256	0.292	0.230	0.401	0.216	0.402	0.217	0.396	0.210	0.252	0.198	0.253	0.198
	Avg.	0.217	0.169	0.231	0.185	0.255	0.213	0.417	0.229	0.416	0.228	0.413	0.224	0.197	0.159	0.197	0.159
CDC-IHA (10 ¹)	12	6.354	4.725	6.314	4.671	6.235	4.996	15.864	5.280	15.925	5.289	15.765	5.227	6.943	5.591	8.613	6.640
	18	5.193	3.968	5.208	3.989	4.908	3.727	12.884	4.435	12.941	4.404	12.894	4.416	5.751	4.600	5.868	4.731
	24	7.643	5.528	7.886	5.535	7.670	5.601	22.026	5.921	21.651	5.934	21.735	5.887	7.852	5.715	8.167	6.077
	30	6.935	5.162	7.206	5.245	6.881	5.104	20.436	5.593	20.008	5.524	20.363	5.511	7.852	5.715	7.906	6.001
	Avg.	6.531	4.846	6.641	4.853	6.411	4.849	17.803	5.307	17.631	5.288	17.689	5.260	6.792	5.324	7.666	5.862
JH COVID-19 (10 ²)	7	2.115	0.825	0.160	0.156	3.141	3.141	2.667	0.404	2.662	0.403	2.672	0.404	0.467	0.356	0.487	0.357
	14	3.052	1.214	0.990	0.817	2.268	2.135	3.795	0.642	3.826	0.644	3.811	0.641	0.547	0.396	0.522	0.389
	28	3.874	1.588	1.279	0.892	2.415	2.118	4.096	0.786	4.143	0.795	4.048	0.779	0.924	0.661	0.916	0.674
	90	5.342	2.101	5.087	4.459	2.765	1.917	1.749	0.378	1.756	0.377	1.724	0.373	3.871	1.486	75.64	11.34
	Avg.	3.596	1.432	1.879	1.580	2.647	2.328	3.077	0.553	3.097	0.554	3.064	0.549	1.452	0.725	19.391	3.190
ILI	24	1.385	0.989	0.904	0.507	1.246	1.174	1.548	0.989	1.545	0.992	1.547	0.989	1.634	1.385	1.583	1.349
	36	1.702	1.321	2.729	1.582	2.316	1.751	1.467	0.988	1.466	0.986	1.464	0.983	1.995	1.667	2.033	1.701
	48	1.914	1.493	2.337	1.227	2.051	1.549	1.669	1.134	1.663	1.128	1.662	1.128	2.158	1.801	2.119	1.765
	60	2.118	1.659	2.075	0.991	1.914	1.493	1.594	1.111	1.590	1.110	1.589	1.106	2.218	1.859	2.195	1.842
	Avg.	1.780	1.366	2.011	1.077	1.881	1.492	1.570	1.056	1.566	1.054	1.566	1.052	1.998	1.675	1.978	1.661