

## A. Additional Metrics

### A.1. CRPS

The average marginal CRPS over dimensions  $D$  and over the predicted time steps compared to the test interval of each dataset is reported in Table 3.

Table 3. Test set CRPS comparison (lower is better) of models on six real world data sets. Mean and standard error metrics for TimeGrad obtained by re-training and evaluating 10 times.

Method	Exchange	Solar	Electricity	Traffic	Taxi	Wikipedia
VAR	0.015±0.000	0.595±0.000	0.060±0.000	0.222±0.000	0.410±0.000	4.101±0.002
GARCH	0.024±0.000	0.928±0.000	0.291±0.000	0.426±0.000	-	-
Vec-LSTM ind-scaling	0.013±0.000	0.434±0.012	0.059±0.001	0.168±0.037	0.586±0.004.	0.379±0.004
Vec-LSTM lowrank-Copula	0.009±0.000	0.384±0.010	0.084±0.006	0.165±0.004	0.416±0.004	0.247±0.001
GP scaling	0.017±0.000	0.415±0.009	0.053±0.000	0.140±0.002	0.346±0.348	1.549±1.017
GP Copula	<b>0.008±0.000</b>	0.371±0.022	0.056±0.002	0.133±0.001	0.360±0.201	<b>0.236±0.000</b>
Transformer MAF	0.012±0.003	0.368±0.001	0.052±0.000	0.134±0.001	0.377±0.002	0.274±0.007
<b>TimeGrad</b>	0.009±0.001	<b>0.367±0.001</b>	<b>0.049±0.002</b>	<b>0.110±0.002</b>	<b>0.311±0.03</b>	0.261±0.02

### A.2. MSE

The MSE is defined as the mean squared error over all the time series dimensions  $D$  and over the whole prediction range with respect to the test data. Results comparing the different methods for the datasets considered are given in Table 4.

Table 4. Test set MSE comparison (lower is better) of models on six real world data sets.

Method	Exchange	Solar	Electricity	Traffic	Taxi	Wikipedia
VAR	$4.4 \times 10^{-2}$	$7.0 \times 10^3$	$1.2 \times 10^7$	$5.1 \times 10^{-3}$	-	-
GARCH	$4.0 \times 10^{-2}$	$3.5 \times 10^3$	$1.2 \times 10^6$	$3.3 \times 10^{-3}$	-	-
Vec-LSTM ind-scaling	<b><math>1.6 \times 10^{-4}</math></b>	$9.3 \times 10^2$	$2.1 \times 10^5$	$6.3 \times 10^{-4}$	$7.3 \times 10^1$	$7.2 \times 10^7$
Vec-LSTM lowrank-Copula	$1.9 \times 10^{-4}$	$2.9 \times 10^3$	$5.5 \times 10^6$	$1.5 \times 10^{-3}$	$5.1 \times 10^1$	$3.8 \times 10^7$
GP scaling	$2.9 \times 10^{-4}$	$1.1 \times 10^3$	<b><math>1.8 \times 10^5</math></b>	$5.2 \times 10^{-4}$	$2.7 \times 10^1$	$5.5 \times 10^7$
GP Copula	$1.7 \times 10^{-4}$	$9.8 \times 10^2$	$2.4 \times 10^5$	$6.9 \times 10^{-4}$	$3.1 \times 10^1$	$4.0 \times 10^7$
Transformer MAF	$3.4 \times 10^{-4}$	$9.3 \times 10^2$	$2.0 \times 10^5$	$5.0 \times 10^{-4}$	$4.5 \times 10^1$	<b><math>3.1 \times 10^7</math></b>
<b>TimeGrad</b>	$2.5 \times 10^{-4}$	<b><math>8.8 \times 10^2</math></b>	$1.97 \times 10^5$	<b><math>4.2 \times 10^{-4}</math></b>	<b><math>2.6 \times 10^1</math></b>	$3.8 \times 10^7$