A PROOF OF EQUATION (8)

We assume that each dimension of Z is independent and identically distributed (i.i.d.), then The posterior distribution of the MLP can be further elaborated as:

$$
\begin{aligned}
q_\phi(Z|X) &= \prod_{i=1}^{N\times H_{\text{in}}\times d} q_\phi(z_i|x_i) \\
&= \prod_{i=1}^{N\times H_{\text{in}}\times d} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(z_i-\mu_i)^2}{2\sigma_i^2}\right)
\end{aligned} \tag{15}
$$

Here, z_i represents the i-th independent dimension of the latent variable Z, and x_i represents the corresponding input dimension. The independence assumption implies that $q_\phi(z_i|x_i)$ are independent for all i, where $i \in \{1, 2, \ldots, N \times H_{\text{in}} \times d\}$.

The prior distribution is defined as:

$$
\begin{aligned}
p(Z) &= \prod_{i=1}^{N\times H_{\text{in}}\times d} p(z_i) \\
&= \prod_{i=1}^{N\times H_{\text{in}}\times d} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_i^2}{2}\right)
\end{aligned} \tag{16}
$$

Thus, the KL divergence between $q_\phi(Z|X)$ and $p(Z)$ is given by:

$$D_{\text{KL}}(q_\phi(Z|X)||p(Z)) \tag{17}$$

$$= \int q_\phi(Z|X) \log \frac{q_\phi(Z|X)}{p(Z)}\, dZ \tag{18}$$

$$= \int q_\phi(Z|X) \log \frac{\prod_{i=1}^{N\times H_{\text{in}}\times d} q_\phi(z_i|x_i)}{\prod_{i=1}^{N\times H_{\text{in}}\times d} p(z_i)}\, dZ \tag{19}$$

$$= \sum_{i=1}^{N\times H_{\text{in}}\times d} \int \left(\prod_{i=1}^{d} q_\phi(z_i|x_i)\right) \log \frac{q_\phi(z_i|x_i)}{p(z_i)} dZ \tag{20}$$

By the independence assumption, each z_i is independent, and the integral over Z can be factorized into a sum over integrals of z_i:

$$D_{\text{KL}}(q_\phi(Z|X)||p(Z)) \tag{21}$$

$$= \sum_{i=1}^{N\times H_{\text{in}}\times d} \int q_\phi(z_i|x_i)\left(\log\frac{1}{\sigma_i} - \frac{(z_i-\mu_i)^2}{2\sigma_i^2} + \frac{z_i^2}{2}\right) dz_i$$

Since $\int q_\phi(z_i|x_i)dz_i = 1$, the KL divergence becomes:

$$D_{\text{KL}}(q_\phi(Z|X)||p(Z)) = \sum_{i=1}^{N\times H_{\text{in}}\times d} \frac{1}{2}\left(\log\frac{1}{\sigma_i^2} - 1 + \sigma_i^2 + \mu_i^2\right) \tag{22}$$

B INSTRUCTION TUNING FORMAT

The instruction format of our proposed spatio-temporal LLM follows the format used in UrbanGPT [19], as shown in Table 5.

Table 5: Instruction format for inflow and outflow prediction

Instructions	Given the historical data for taxi flow over 12 time steps in a specific region of Chicago, the recorded taxi inflows are [0 2 1 1 1 0 1 1 0 1 0 2], and the recorded taxi outflows are [0 1 0 2 1 2 0 1 0 1 2 0]. The recording time of the historical data is 'January 1, 2021, 00:00, Friday to January 1, 2021, 05:30, Friday, with data points recorded at 30-minute intervals'. Here is the region information: This region is located within the city of Chicago and encompasses various POIs within a four-kilometer radius, covering cafe, secondary_school, hardware_store, supermarket, pharmacy, restaurant, clothing_store, department_store, lodging, doctor categories. Now we want to predict the taxi inflow and outflow for the next 12 time steps during the time period of 'January 1, 2021, 06:00, Friday to January 1, 2021, 11:30, Friday, with data points recorded at 30-minute intervals'.
Additional Information	To improve prediction accuracy, a spatio-temporal model is utilized to encode the historical taxi data as tokens <ST_HIS>, where the first and the second tokens correspond to the representations of taxi inflow and outflow. Please conduct an analysis of the traffic patterns in this region, taking into account the provided time and regional information, and then generate the predictive tokens for regression, in the form <ST_PRE>