Traffic dataset, LP_kernel, 24-steps non-overlapping, 29-th time series, total epochs: 300 --- Source: 0 : t₀ --- Target: $t_0 + 1: t_0 + 24$ Prediction: $t_0 + 1 : t_0 + 24$ Attention point: $t_0 + 1$?1 Attention layer 1 for $t_0 + 1$ 0.25 0.20 0.15 0.10 0.05 Attention layer 2 for $t_0 + 1$ 0.08 0.06 0.04 Attention layer 3 for $t_0 + 1$ 0.06 0.05 0.04

48000

50000

52000

44000

46000

42000