Supplementary Equations

May 2025

S1 Traffic Volume

The average hourly traffic T_x per buffer is calculated as:

$$T_x = \frac{\sum(l_{ix} \cdot T_{ix})}{\sum l_{ix}} \tag{S1}$$

where l corresponds to the length of road segment i in buffer x.

S2 Ensemble Trees

S2.1 Random Forest

The model:

$$\hat{y} = \frac{1}{K}\sum_{k=1}^{K} T_k(x) \tag{S2}$$

Node importance:

$$n_{ij} = w_j C_j - w_{\text{left}(j)} C_{\text{left}(j)} - w_{\text{right}(j)} C_{\text{right}(j)} \tag{S3}$$

Feature importance:

$$fi_i = \frac{\sum_{j:\text{node j splits on feature i}} n_{ij}}{\sum_{k \in \text{all nodes}} n_{ik}} \tag{S4}$$

Normalized importance:

$$\text{norm} fi_i = \frac{fi_i}{\sum_{j \in \text{all features}} fi_j} \tag{S5}$$

S2.2 Gradient Boosting

S2.2.1 LightGBM

$$H_T(x) = \sum_{t=1}^{T} H_t(x), \quad H_t \in I \tag{S6}$$

Learner optimization:

$$h_t(x) = \arg\min_{h \in H} L(y, H_{t-1}(x) + h(x)) \tag{S7}$$

Negative gradient:

$$r_t = \frac{\partial L(y, H_{t-1}(x))}{\partial H_{t-1}(x)} \tag{S8}$$

Fit approximation:

$$h_t(x) = \arg\min \sum (r_t - h_t(x))^2 \tag{S9}$$

Strong learner:

$$H_t(x) = h_t(x) + H_{t-1}(x) \tag{S10}$$

S2.2.2 XGBoost

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathcal{F} \tag{S11}$$

$$\mathcal{F} = \{f(x) = w_{q(x)}\} \tag{S12}$$

Objective function:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{S13}$$

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \|\omega\|^2 \tag{S14}$$

After iteration:

$$L_t = \sum_i l(y, \hat{y}^i_{t-1}) + f_t(x_i) + \Omega(f_t) \tag{S15}$$

Approximation:

$$L_t \approx \sum_{i=1}^{n} \left[l(y, \hat{y}^i_{t-1}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{S16}$$

Optimal weight:

$$\omega_j^* = -\frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{S17}$$

Minimum loss:

$$\tilde{L}_t = -\frac{1}{2} \sum_{j=1}^{T} \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} \tag{S18}$$

S3 Multiple Linear Regression

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_j x_{ij} + \epsilon_i \tag{S19}$$

$$\epsilon_i = y_i - \hat{y}_i \tag{S20}$$

S3.1 Ridge Regression

$$\min_\beta L_2 = (y - X\beta)^2 + \lambda \sum_{i=1}^{p} \beta_i^2 \tag{S21}$$

S3.2 LASSO Regression

$$\min_{\beta} L_1 = (y - X\beta)^2 + \lambda \sum_{i=1}^{p} |\beta_i| \tag{S22}$$

S4 Kriging

$$\gamma(h) = \frac{1}{2N(h)} \sum_{j=1}^{N(h)} [z(s_j + h) - z(s_j)]^2 \tag{S23}$$

S4.1 Ordinary Kriging

$$Z \sim \mathcal{N}(\mu, \Sigma_Z) \tag{S24}$$

$$\Sigma_Z = \delta^2 R(\alpha) + \tau^2 I \tag{S25}$$

S4.2 Universal Kriging

$$\hat{z}(u) = \beta_0 + \beta_1 x_1(u) + \beta_2 x_2(u) + \cdots + \sum_{i=1}^{n} \lambda_i z(u_i) \tag{S26}$$

S5 Mixed Effects Model

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_i + e_{ij} \tag{S27}$$

S6 Feature Selection with Shapley Values

$$\phi_j(val) = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [val(S \cup \{j\}) - val(S)] \tag{S28}$$

References

[1] Shapley, L. S. (1953). *Stochastic games*. Proceedings of the National Academy of Sciences, 39(10), 1095–1100.

[2] Patelli, L., Cameletti, M., Golini, N., & Ignaccolo, R. (2023). A path in regression random forest looking for spatial dependence.

[3] Yao, X., Fu, X., & Zong, C. (2022). Short-term load forecasting method based on feature preference strategy and LightGBM-XGBoost.

[4] Khan, M. et al. (2023). Spatial interpolation of water quality index based on kriging.

[5] Algaba, E., Fragnelli, V., & Sánchez-Soriano, J. (2019). *Handbook of the Shapley Value*. CRC Press.