
Assimilation of Sparse Vehicle Trajectories with a Macroscopic Traffic Model

Julien Moreau^{1,2} Marc Lelarge¹

¹ Inria Paris, DI ENS, PSL Research University ² Renault Group
`{julien.moreau, marc.lelarge}@inria.fr`

Abstract

Accurate macroscopic traffic prediction could lead to smarter vehicles and safer roads. It is however dependent on reliable and frequent measurements. Recent research in scientific machine learning has shown progress on reconstruction tasks using probe vehicles, hinting at potential gains for a sequential prediction model. In this work, we adopt the classical data assimilation framework and gather insight on how to design effective predictors. Our results highlight the effectiveness of macroscopic solvers given sparse data while confirming the flexibility of Kalman filtering. However, we report the difficulty of calibrating the solution given the low amount of measures and their high variance.

1 Introduction

The goal of Traffic State Estimation (TSE) is to infer physical traffic quantities (velocity, density) from partial observations. In this paper, we study a fleet of connected vehicles that continuously transmit their GPS coordinates and speed to a remote computing center. We are interested in predicting the state of traffic in real time even when no probe measurement is available. Due to the low penetration rate of probe vehicles (typically 1-5%), direct estimation of macroscopic traffic quantities is challenging. TSE algorithms exploit physical knowledge of the traffic dynamics, they often consider the speed of individual vehicles as a noisy and sparse measure of a macroscopic mean velocity.

Several models have been introduced in the TSE literature to address this particular challenge. In this study, we build upon the most common modeling blocks and evaluate their effectiveness across diminishing probe penetration rates. We investigate what conditions are required to achieve a reliable estimation of congestion patterns and how to find reasonable values for the physical and statistical hyperparameters. Our implementation is available at <https://github.com/J-Moreau/traffic-assim>.

2 Background

Macroscopic Traffic PDEs Car traffic can be represented with tools from compressible fluid dynamics. The Eulerian formulation defines continuous scalar fields for velocity $v(x, t)$ and density $\rho(x, t)$ which depend on space and time coordinates. The evolution of traffic quantities is described with partial differential equations (PDEs), notably the one-dimensional transport equation from Navier-Stokes $\partial_t \rho + \partial_x Q = 0$ where $Q = \rho v$ is the traffic flow (vehicles per time unit).

A simple relation for speed and density The most common formulation is the LWR model [Lighthill and Whitham, 1955, Richards, 1956] which constrains v and Q to be functions of ρ . A popular choice is Greenshields' model where the density-velocity relation is affine $v = V_e(\rho) = V_{\max}(1 - \rho/\rho_{\max})$ and the flow is a quadratic function $Q = Q_e(\rho) = \rho V_{\max}(1 - \rho/\rho_{\max})$ as illustrated

in Figure 1. This is strongly related to the inviscid Burgers equation and exhibits the same shockwave behavior. The parameters (V_{\max} , ρ_{\max}) depend on the road characteristics and are usually estimated from real data or domain knowledge.

The relation between ρ , Q and v is referred to as the Fundamental Diagram. Other formulations are possible for Q_e such as triangular or trapezoidal [Daganzo, 1994, 1995, Newell, 1982]. If the fundamental diagram is invertible, as is the case for Greenshields' model, it is possible to define the inverse function for density $\rho = R_e(v)$. The LWR model can thus be equivalently formulated as a PDE on velocity [Work et al., 2010].

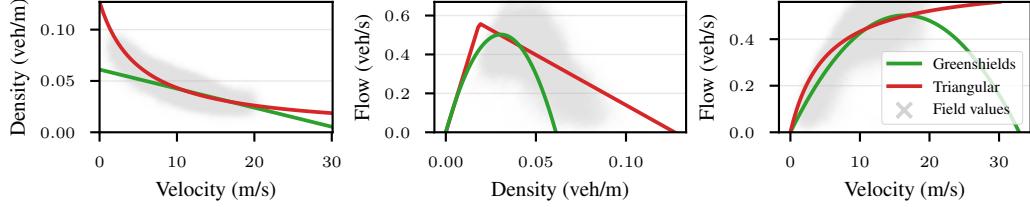


Figure 1: Macroscopic Fundamental Diagram fitted to the NGSIM US-101 data.

Numerical solutions A popular solution for integrating hyperbolic PDEs are finite volume methods such as Godunov's scheme [Godunov and Bohachevsky, 1959] which divides the road into discrete cells. The application of Godunov to LWR with trapezoidal flow is known as the Cell Transmission Model [Daganzo, 1994, Lebacque, 1996]. If conditions at the boundaries are known (e.g. through loop detectors) it is possible to unroll the physical model to obtain an estimate of the state [Fan and Seibold, 2013]. However, the measures are expected to differ from the simulated physics and one has to reconcile the two.

Data Assimilation Incorporation of measurements into a physical model has been studied extensively in the geophysics community and is referred to as Data Assimilation (DA). DA Algorithms minimize the discrepancy between the observed measurements and the inferred state, they can be classified into two main categories: variational (e.g. 3D, 4DVar) and sequential (e.g. Kalman and Particle Filters). In TSE problems, the state dimension is reasonable, and the solver is computationally cheap, making sequential methods the standard. Given a linear state space model with Gaussian noise, the Kalman Filter (KF) iteratively returns the optimal interpolation between physical process and measurements at the present time. The Extended Kalman filter (EKF) linearizes the process at each step to handle non-linear fluid mechanics solvers.

3 Related Work

Sequential Data Assimilation Applying a linearized KF for predicting the traffic state has shown good performance in practice [Herrera and Bayen, 2010]. However, solvers such as Godunov are not differentiable everywhere which can lead to error accumulation [Blandin et al., 2012]. An alternative is to use the Ensemble KF [Work et al., 2010, Seo et al., 2015] or a differentiable solver such as finite differences [Wang and Papageorgiou, 2005]. Later research has shown the benefits of adopting Lagrangian coordinates [Duret and Yuan, 2017, Yang et al., 2019] which require additional processing to get back to an Eulerian prediction.

Deep forward solvers Extensive research was led on training neural substitutes for fluid dynamics solvers [Li et al., 2021, Sun et al., 2023, Brandstetter et al., 2022]. They are usually trained in an autoregressive fashion on clean simulated data and are not meant to work directly on sparse noisy inputs. Another research direction is to train models on noisy or partial simulation measurements which can later be applied to a data assimilation framework [Cheng et al., 2023, Singh et al., 2023, Rozet and Louppe, 2023]. Bridging the gap between simulation and the real setting remains a non-trivial problem.

Inverse solvers for reconstruction Many studies have focused on solving the inverse problem, i.e. inferring missing data or parameters from partial observations. A popular solution is to leverage

Physics informed deep learning [Raissi et al., 2019] to obtain a solution that respects both the measures and the PDE prior. PINNs have been applied for TSE on probes [Shi et al., 2021, 2022] or loop measurements [Goatin and Inzunza, 2023] and are able to learn the physical parameters. However, they can struggle to learn the discontinuous shockwaves of hyperbolic PDEs [De Ryck et al., 2024]. Networks can also be trained to reconstruct traffic on simulations and later applied to real data [Thodi et al., 2022, 2024]. These architectures are designed to infer the gaps in the traffic state history and are not directly suited for sequential prediction.

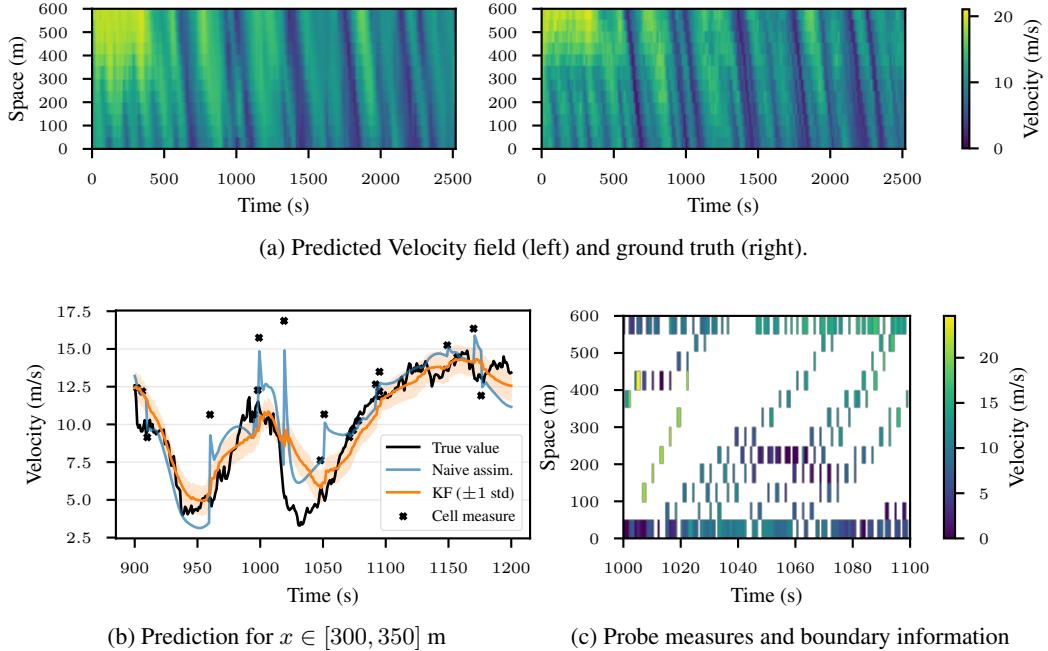


Figure 2: Triangular-EKF prediction results on NGSIM US-101. Fig. 2a and 2b: The EKF is able to reconstruct coherent values even when measurements strongly deviate from the macroscopic mean value. In contrast, the naive assimilation procedure is perturbed by the noisy observations. Fig. 2c shows the input data: probe penetration 2%, boundary information 20%.

4 Data Assimilation Model

We use the EKF framework and model the system with a Linear State Space and Gaussian noise. Time is discretized into steps of length Δt indexed by $t \in [1, N]$, space is divided into cells of length Δx indexed by $k \in [1, C]$. Let $\mathbf{v}_t \in \mathbb{R}^C$ be the mean velocity in cells $[c_k, c_{k+1}]$. Let $\mathbf{x}_t^{\text{meas}}$ and $\dot{\mathbf{x}}_t^{\text{meas}}$ vectors of \mathbb{R}^P be the measured position and speed of probe vehicles indexed by $i \in [1, P]$. We consider the following discretized model of traffic flow:

$$\begin{cases} \mathbf{v}_{t+1} = \mathcal{F}_\theta(\mathbf{v}_t) + \mathbf{w}_t, & \mathbf{w}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}) \\ \dot{\mathbf{x}}_t^{\text{meas}} = \mathcal{H}(\mathbf{x}_t^{\text{meas}})\mathbf{v}_t + \boldsymbol{\epsilon}_t, & \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{R}) \end{cases} \quad \begin{array}{l} (\text{Macroscopic Process}) \\ (\text{Measurement}) \end{array} \quad (1)$$

Where \mathcal{F}_θ is the fluid mechanics solver (see Appendix A for details), with physical parameters θ , and \mathbf{w}_t , $\boldsymbol{\epsilon}_t$ are i.i.d. Gaussian error terms. $\mathcal{H} : \mathbb{R}^P \rightarrow \mathbb{R}^{P \times C}$ is the mapping between cell k and probe vehicle i : $\mathcal{H}(\mathbf{x})_{i,k} = \mathbb{1}(x_i \in [c_k, c_{k+1}])$. The discrete scheme has to respect the CFL condition $\Delta x \geq V_{\max} \Delta t$.

The EKF prediction and correction equations are:

$$\begin{cases} \mathbf{v}_{t|t-1} = \mathcal{F}_\theta(\mathbf{v}_{t-1|t-1}) \\ \mathbf{P}_{t|t-1} = \mathbf{F}_{t-1}\mathbf{P}_{t-1|t-1}\mathbf{F}_{t-1}^T + \mathbf{Q} \end{cases} \quad \begin{cases} \mathbf{K}_t = \mathbf{P}_{t|t-1}\mathbf{H}_t^T(\mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T + \mathbf{R})^{-1} \\ \mathbf{v}_{t|t} = \mathbf{v}_{t|t-1} + \mathbf{K}_t(\dot{\mathbf{x}}_t^{\text{meas}} - \mathbf{H}_t\mathbf{v}_{t|t-1}) \\ \mathbf{P}_{t|t} = (\mathbf{I} - \mathbf{K}_t\mathbf{H}_t)\mathbf{P}_{t|t-1} \end{cases} \quad (2)$$

Where \mathbf{F}_t is the Jacobian of the solver \mathcal{F}_θ (see Appendix C) and $\mathbf{H}_t = \mathcal{H}(\mathbf{x}_t^{\text{meas}})$.

5 Experiments

Dataset Experiments are led on the NGSIM dataset [USDT, 2016] which contains trajectories of all vehicles on two U.S. freeway segments, US-101 and I-80. We acknowledge that incoherences have been detected in the trajectory interpolation [Coifman and Li, 2017], but we hypothesize that macroscopic quantities are correctly represented. We sample the position and speed of vehicles every 3 seconds to simulate periodic fleet measurements.

Boundaries The Godunov solver can handle partial boundary conditions with the use of ghost cells, letting waves exit the domain and assuming constant flow when no wave is entering. To mitigate the lack of information at the boundaries, we add a larger fraction ($p=20\%$) of probe information at the borders. This emulates contributions from fleet vehicles outside the studied domain.

Hyperparameters Fundamental Diagrams are fit to the US-101 data: parameters for the Triangular FD are chosen to be the same as Herrera and Bayen [2010] and the Greenshields FD is fitted with least-squares on the $Q(v)$ relation. We use the discretization $\Delta t = 1s$, $\Delta x = 50m$, measurement variance for ϵ is computed by taking the mean of the empirical variance for each cell and timestep.

We argue that reasonable values for physical parameters and measure variance \mathbf{R} can be inferred from domain knowledge. The Kalman process variance \mathbf{Q} however depends on the correctness of the model approximation and is trickier to determine beforehand. We use an initial value of $\sigma_Q^2 = 0.1m^2s^{-2}$ and use a single step of Desroziers' method [Desroziers et al., 2005] to get a variance estimate $\mathbb{E}[(\mathbf{v}_{t+1|t+1} - \mathcal{F}(\mathbf{v}_{t|t}))^2]$. This has no significant effect on performance.

6 Results

In Table 1, the benchmarked physical models are the Greenshields and Triangular diagrams as well as a physics-agnostic baseline where the solver is the Identity. We also compare the Kalman analysis to a naive assimilation protocol where the state is set to the average between measure and prediction, this is the equivalent of setting a constant gain $K=1/2$. This model obtains surprisingly good results on US-101 which we attribute to the greater trust it gives to measures. Figure 2 highlights a successful reconstruction.

Table 1: EKF MSE (m^2s^{-2}), avg \pm std over 10 probe sampling seeds. The Triangular function consistently outperforms Greenshields. Tri. ($K=1/2$) is the Triangular diagram with naive assimilation.

Model	US-101			I-80		
	p=1%	p=2%	p=5%	p=1%	p=2%	p=5%
Identity	10.4 ± 1.4	$7.37 \pm .60$	$4.41 \pm .18$	8.06 ± 3.2	5.75 ± 1.5	$3.43 \pm .27$
Greenshields	4.57 ± 1.11	$3.74 \pm .24$	$3.27 \pm .34$	$3.23 \pm .42$	$3.01 \pm .30$	$2.83 \pm .22$
Triangular	4.43 ± 1.18	$3.16 \pm .37$	$2.37 \pm .27$	$2.31 \pm .35$	$2.07 \pm .19$	$1.75 \pm .11$
Tri. ($K=1/2$)	5.18 ± 1.51	$3.19 \pm .42$	$2.10 \pm .16$	$3.34 \pm .82$	$3.28 \pm .39$	$3.14 \pm .30$

This benchmark suggests the importance of choosing a valid physical model, the amount of measures is too small to correctly guide a physics-agnostic filter. While it is not straight forward to parametrize the process and measurement variance, the EKF exhibits consistent performance and is a sound choice for the assimilation.

7 Future Perspectives

An interesting future work direction would be to include learning-based components to improve the algorithm. The presented EKF is differentiable and could be optimized with gradient descent, however a direct grid search seems more appropriate given the low number of parameters. Training a more expressive solver on probe data only is challenging due to the low supervision information contained in the sparse and noisy trajectories. Exploring whether these issues can be addressed with a larger, potentially simulated, dataset could be a topic for further investigation.

Acknowledgement

This material is based upon work supported by the ANRT (Association nationale de la recherche et de la technologie) with a CIFRE fellowship granted to Julien Moreau.

References

- S. Blandin, A. Couque, A. Bayen, and D. Work. On sequential data assimilation for scalar macroscopic traffic flow models. *Physica D: Nonlinear Phenomena*, 241(17):1421–1440, Sept. 2012. ISSN 0167-2789. doi: 10.1016/j.physd.2012.05.005.
- J. Brandstetter, D. Worrall, and M. Welling. Message passing neural pde solvers. *ICLR 2022*, 2022.
- S. Cheng, C. Quilodrán-Casas, S. Ouala, A. Farchi, C. Liu, P. Tandeo, R. Fablet, D. Lucor, B. Iooss, J. Brajard, D. Xiao, T. Janjic, W. Ding, Y. Guo, A. Carrassi, M. Bocquet, and R. Arcucci. Machine Learning With Data Assimilation and Uncertainty Quantification for Dynamical Systems: A Review. *IEEE/CAA Journal of Automatica Sinica*, 10(6):1361–1387, June 2023. ISSN 2329-9274. doi: 10.1109/JAS.2023.123537.
- B. Coifman and L. Li. A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset. *Transportation Research Part B: Methodological*, 105:362–377, Nov. 2017. ISSN 0191-2615. doi: 10.1016/j.trb.2017.09.018.
- C. F. Daganzo. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, 28(4):269–287, Aug. 1994. ISSN 0191-2615. doi: 10.1016/0191-2615(94)90002-7.
- C. F. Daganzo. The cell transmission model, part II: Network traffic. *Transportation Research Part B: Methodological*, 29(2):79–93, 1995. ISSN 0191-2615. doi: 10.1016/0191-2615(94)90022-R.
- T. De Ryck, S. Mishra, and R. Molinaro. wpinns: Weak physics informed neural networks for approximating entropy solutions of hyperbolic conservation laws. *SIAM Journal on Numerical Analysis*, 62(2):811–841, 2024. doi: 10.1137/22M1522504. URL <https://doi.org/10.1137/22M1522504>.
- G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis-error statistics in observation space. In *Quarterly Journal of the Royal Meteorological Society*, volume 131, pages 3385–3396, 2005.
- A. Duret and Y. Yuan. Traffic state estimation based on Eulerian and Lagrangian observations in a mesoscopic modeling framework. *Transportation Research Part B: Methodological*, 101:51–71, July 2017. ISSN 0191-2615. doi: 10.1016/j.trb.2017.02.008.
- S. Fan and B. Seibold. Data-fitted first-order traffic models and their second-order generalizations: Comparison by trajectory and sensor data. *Transportation Research Record*, 2391(1):32–43, 2013. doi: 10.3141/2391-04. URL <https://doi.org/10.3141/2391-04>.
- P. Goatin and D. Inzunza. A PINN approach for traffic state estimation and model calibration based on loop detector flow data. In *MT-ITS 2023 - 8th International Conference on Models and Technologies for Intelligent Transportation Systems*, Saint-Laurent-Du-Var, France, June 2023.
- S. K. Godunov and I. Bohachevsky. Finite difference method for numerical computation of discontinuous solutions of the equations of fluid dynamics. *Matematicheskij sbornik*, 47(89)(3):271–306, 1959.
- J. C. Herrera and A. M. Bayen. Incorporation of Lagrangian measurements in freeway traffic state estimation. *Transportation Research Part B: Methodological*, 44(4):460–481, May 2010. ISSN 0191-2615. doi: 10.1016/j.trb.2009.10.005.
- JP. Lebacque. The godunov scheme and what it means for first order traffic flow models. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Lyon, France, July*, volume 2426, 1996.

- R. J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2002. ISBN 978-0-521-00924-9. doi: 10.1017/CBO9780511791253.
- Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. Fourier neural operator for parametric partial differential equations. 2021. URL <https://doi.org/10.48550/arXiv.2010.08895>.
- M. J. Lighthill and G. B. Whitham. On kinematic waves II: A theory of traffic flow on long, crowded roads. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 229(1178):317–345, 1955.
- G. F. Newell. *Applications of Queueing Theory*. Chapman and Hall, London, 2 edition, 1982.
- M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, Feb. 2019. ISSN 0021-9991. doi: 10.1016/j.jcp.2018.10.045.
- P. I. Richards. Shock Waves on the Highway. *Operations Research*, 4(1):42–51, Feb. 1956. ISSN 0030-364X. doi: 10.1287/opre.4.1.42.
- F. Rozet and G. Louppe. Score-based Data Assimilation. *Advances in Neural Information Processing Systems*, 36:40521–40541, Dec. 2023.
- T. Seo, T. Kusakabe, and Y. Asakura. Traffic State Estimation with the Advanced Probe Vehicles Using Data Assimilation. Sept. 2015. doi: 10.1109/ITSC.2015.139.
- R. Shi, Z. Mo, and X. Di. Physics-Informed Deep Learning for Traffic State Estimation: A Hybrid Paradigm Informed By Second-Order Traffic Models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1):540–547, May 2021. ISSN 2374-3468. doi: 10.1609/aaai.v35i1.16132.
- R. Shi, Z. Mo, K. Huang, X. Di, and Q. Du. A Physics-Informed Deep Learning Paradigm for Traffic State and Fundamental Diagram Estimation. *IEEE Transactions on Intelligent Transportation Systems*, 23(8):11688–11698, Aug. 2022. ISSN 1558-0016. doi: 10.1109/TITS.2021.3106259.
- A. Singh, R. A. Borsoi, D. Erdogmus, and T. Imbiriba. Learning semilinear neural operators: A unified recursive framework for prediction and data assimilation. In *The Twelfth International Conference on Learning Representations*, Oct. 2023.
- Z. Sun, Y. Yang, and S. Yoo. A neural PDE solver with temporal stencil modeling. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *ICML’23*, pages 33135–33155, Honolulu, Hawaii, USA, July 2023. JMLR.org.
- B. T. Thodi, Z. S. Khan, S. E. Jabari, and M. Menéndez. Incorporating Kinematic Wave Theory Into a Deep Learning Method for High-Resolution Traffic Speed Estimation. *IEEE Transactions on Intelligent Transportation Systems*, 23(10):17849–17862, Oct. 2022. ISSN 1558-0016. doi: 10.1109/TITS.2022.3157439.
- B. T. Thodi, S. V. R. Ambadipudi, and S. E. Jabari. Fourier neural operator for learning solutions to macroscopic traffic flow models: Application to the forward and inverse problems. *Transportation Research Part C: Emerging Technologies*, 160:104500, Mar. 2024. ISSN 0968-090X. doi: 10.1016/j.trc.2024.104500.
- U. USDT. Next generation simulation (ngsim) vehicle trajectories and supporting data. Dataset, ITS DataHub via Data.transportation.gov, 2016. URL <https://doi.org/10.21949/1504477>.
- Y. Wang and M. Papageorgiou. Real-time freeway traffic state estimation based on extended Kalman filter: A general approach. *Transportation Research Part B: Methodological*, 39(2):141–167, Feb. 2005. ISSN 0191-2615. doi: 10.1016/j.trb.2004.03.003.
- D. B. Work, S. Blandin, O.-P. Tossavainen, B. Piccoli, and A. M. Bayen. A Traffic Model for Velocity Data Assimilation. *Applied Mathematics Research eXpress*, 2010(1):1–35, Jan. 2010. ISSN 1687-1200. doi: 10.1093/amrx/abq002.

H. Yang, Jin, and Ran. Freeway traffic state estimation: A Lagrangian-space Kalman filter approach. *Journal of Intelligent Transportation Systems*, 23(6):525–540, Jan. 2019. ISSN 1547-2450. doi: 10.1080/15472450.2018.1476147.

A Fluid Mechanics Solver

We provide more background on applying Godunov's scheme to a velocity-based LWR.

The macroscopic process \mathcal{F}_θ takes velocities \mathbf{v}_t as input and translates them to densities ρ_t using the fundamental diagram. For Greenshields this is simply:

$$\rho_t = R_e(\mathbf{v}_t) = \rho_{\max}(1 - \mathbf{v}_t/V_{\max}) \quad (3)$$

The Godunov step then updates each cell k given its neighbors $k - 1$ and $k + 1$:

$$\rho_{t+1}^k = \rho_t^k + \frac{\Delta t}{\Delta x} (Q_{\text{boundary}}(\rho_t^{k-1}, \rho_t^k) - Q_{\text{boundary}}(\rho_t^k, \rho_t^{k+1})) \quad (4)$$

Q_{boundary} is the flow at the cell boundaries and is given by the solution to Riemann's problem [LeVeque, 2002]:

$$Q_{\text{boundary}}(\rho_\ell, \rho_r) := \begin{cases} \min [Q_e(\rho_\ell), Q_e(\rho_r)] & \text{if } \rho_\ell \leq \rho_r \\ \max [Q_e(\rho_\ell), Q_e(\rho_r), Q_e(\rho_c)] & \text{if } \rho_\ell \geq \rho_r \end{cases} \quad (5)$$

Where ρ_c is the critical density where flow is maximal, it is the threshold between the congestion regime ($\rho \geq \rho_c$) and free flow ($\rho \leq \rho_c$). For more insight into the Riemann problem, refer to Appendix B.

Then the Fundamental Diagram is used again to convert ρ_{t+1} to \mathbf{v}_{t+1} .

$$\mathbf{v}_{t+1} = V_e(\rho_{t+1}) = V_{\max}(1 - \rho_{t+1}/\rho_{\max}) \quad (6)$$

The solver can thus be defined as the composition of (3), (4) and (6).

$$\mathcal{F}_\theta = V_e \circ \text{Godunov} \circ R_e \quad (7)$$

Note that the relations can be greatly simplified for specific Fundamental Diagrams. The interested reader can refer to Work et al. [2010] for Greenshields and Herrera and Bayen [2010] for the Triangular flow.

B Solution to the Riemann Problem for LWR

The Riemann problem studies the evolution of a conservation law under a piecewise-constant initial value with a discontinuity at the middle. The initial density on the left and right sides is defined as:

$$\forall x \in \mathbb{R} \setminus \{0\}, \quad \rho(x, t=0) = \begin{cases} \rho_\ell & \text{if } x < 0 \\ \rho_r & \text{if } x > 0 \end{cases}$$

For a 1D Hyperbolic conservation law of the form $\partial_t \rho + \partial_x Q(\rho) = 0$ and strictly concave flow function $Q \in C^1(\mathbb{R})$, a general solution can be derived analytically. Figure 3 summarizes the solution dynamics depending on the values of ρ_ℓ and ρ_r .

The problem is solved by a similarity solution, i.e. it only depends on x/t .

If $\rho_\ell > \rho_r$ the solution is a rarefaction wave or "fan":

$$\forall x \in \mathbb{R}, \forall t > 0 \quad \rho(x, t) = \begin{cases} \rho_\ell & \text{if } x/t < Q'(\rho_\ell) \\ Q'^{-1}(x/t) & \text{if } Q'(\rho_\ell) \leq x/t \leq Q'(\rho_r) \\ \rho_r & \text{if } x/t > Q'(\rho_r) \end{cases}$$

When the wave speeds verify $Q'(\rho_\ell) < 0 < Q'(\rho_r)$ the density spreads in the two directions, this is a two-sided rarefaction. Here we assumed that the derivative Q' of the flow is invertible which is true

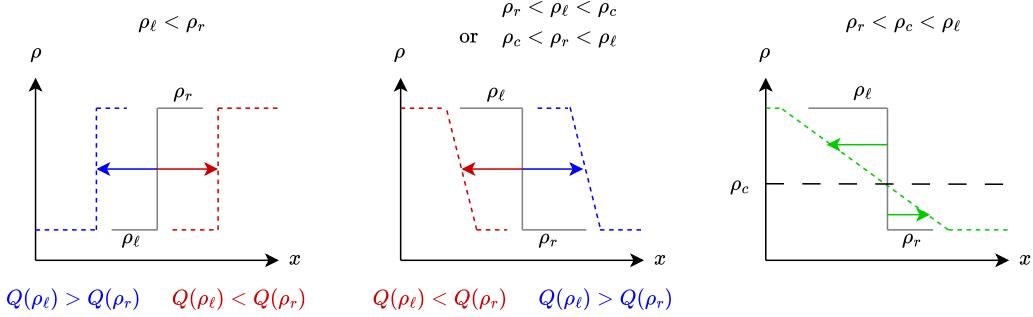


Figure 3: Solution to the LWR-Greenhields PDE with Riemann initial value (solid line) at $t = 0$. The dashed line represents the density value for a given $t > 0$. The wave direction varies depending on the sign of $Q(\rho_\ell) - Q(\rho_r)$ i.e. if the inflow is greater (blue) or lower (red) than the outflow.
Left: Shockwave propagation. *Center:* Rarefaction wave. *Right:* Two-sided rarefaction.

only for strictly concave Q . In the case of Triangular flow, the two-sided rarefaction can instead be solved with $\rho(x, t) = \rho_c$ if $Q'(\rho_\ell) \leq x/t \leq Q'(\rho_r)$.

If $\rho_\ell < \rho_r$ the solution is a discontinuous shockwave:

$$\forall x \in \mathbb{R}, \forall t > 0 \quad \rho(x, t) = \begin{cases} \rho_\ell & \text{if } x/t < c \\ \rho_r & \text{if } x/t > c \end{cases} \quad \text{Where } c = \frac{Q(\rho_\ell) - Q(\rho_r)}{\rho_\ell - \rho_r}$$

This setting is the basis for solving finite volume schemes where density is piecewise constant with a discontinuity at the cell boundaries. Here the left and right "cells" are of infinite length. To apply this scheme in discretized space and time, the CFL condition $|Q'| \leq \Delta x \Delta t$ must hold. It ensures that the waves don't travel past the cell boundaries during one time step.

Note that the value of $Q(\rho(x = 0, t))$ is constant for $t > 0$. This boundary flow is the one used in Godunov's scheme (5). The flow Q_{boundary} is either equal to the left or right side depending on the direction of the wave, or to the critical value $Q(\rho_c) = \max_{0 \leq \rho \leq \rho_{\max}} Q(\rho)$ when there is a two-sided rarefaction. Because the flow is concave, the value of $\rho(x = 0, t)$ for a two-sided rarefaction is $Q'^{-1}(0) = \rho_c$, when Q reaches its maximum.

C Jacobian of scalar Godunov scheme

It is possible to rewrite the Godunov step (4) in terms of supply s and demand d :

$$\begin{aligned} \rho_{t+1}^k &= \rho_t^k + \frac{\Delta t}{\Delta x} (\min(d_t^{k-1}, s_t^k) - \min(d_t^k, s_t^{k+1})) \\ d_t^k &= \begin{cases} Q_e(\rho_t^k) & \text{if } \rho_t^k < \rho_c \\ Q_e(\rho_c) & \text{if } \rho_t^k \geq \rho_c \end{cases} \quad s_t^k = \begin{cases} Q_e(\rho_t^k) & \text{if } \rho_t^k > \rho_c \\ Q_e(\rho_c) & \text{if } \rho_t^k \leq \rho_c \end{cases} \end{aligned}$$

Let us introduce the following scalar derivatives:

$$\begin{aligned} A_t^k &= \left. \frac{\partial}{\partial \rho_t^k} \min(d_t^{k-1}, s_t^k) \right|_{\rho_t^k} = Q'_e(\rho_t^k) \cdot \mathbb{1}(\rho_t^k > \rho_c) \mathbb{1}(d_t^{k-1} > s_t^k) \\ B_t^k &= \left. \frac{\partial}{\partial \rho_t^k} \min(d_t^k, s_t^{k+1}) \right|_{\rho_t^k} = Q'_e(\rho_t^k) \cdot \mathbb{1}(\rho_t^k < \rho_c) \cdot \mathbb{1}(d_t^k < s_t^{k+1}) \\ C_t^{k-1} &= \left. \frac{\partial}{\partial \rho_t^{k-1}} \min(d_t^{k-1}, s_t^k) \right|_{\rho_t^k} = Q'_e(\rho_t^{k-1}) \cdot \mathbb{1}(\rho_t^{k-1} < \rho_c) \cdot \mathbb{1}(d_t^{k-1} < s_t^k) \\ D_t^{k+1} &= \left. \frac{\partial}{\partial \rho_t^{k+1}} \min(d_t^k, s_t^{k+1}) \right|_{\rho_t^k} = Q'_e(\rho_t^{k+1}) \cdot \mathbb{1}(\rho_t^{k+1} > \rho_c) \cdot \mathbb{1}(d_t^k > s_t^{k+1}) \end{aligned}$$

The Jacobian matrix $\mathbf{J} \in \mathbb{R}^{C \times C}$ is tridiagonal. The diagonal terms are:

$$J_{k,k} = \frac{\partial \rho_{t+1}^k}{\partial \rho_t^k} \Big|_{\rho_t} = 1 + \frac{\Delta t}{\Delta x} (A_t^k - B_t^k)$$

Then the lower and upper diagonal terms are:

$$\begin{aligned} J_{k,k-1} &= \frac{\partial \rho_{t+1}^k}{\partial \rho_t^{k-1}} \Big|_{\rho_t} = \frac{\Delta t}{\Delta x} C_t^{k-1} \\ J_{k,k+1} &= \frac{\partial \rho_{t+1}^k}{\partial \rho_t^{k+1}} \Big|_{\rho_t} = -\frac{\Delta t}{\Delta x} D_t^{k+1} \end{aligned}$$

Finally, the complete solver Jacobian is:

$$\mathbf{F}_t = \frac{\partial \mathcal{F}_\theta}{\partial \mathbf{v}} \Big|_{\mathbf{v}_t} = \text{diag}[V'_e(\boldsymbol{\rho}_{t+1})] \mathbf{J} \text{ diag}[R'_e(\mathbf{v}_t)]$$

In the specific case of Greenshields, this simplifies to:

$$\mathbf{F}_t = \mathbf{J}$$