



Phase transition model of non-stationary traffic flow: Definition, properties and solution method



Sébastien Blandin ^{a,*}, Juan Argote ^b, Alexandre M. Bayen ^{b,c}, Daniel B. Work ^d

^a IBM Research Collaboratory - Singapore, Nanyang Technological University, N4-01a-0150, Nanyang Avenue, Singapore 639798, Singapore

^b Department of Civil and Environmental Engineering, University of California Berkeley, United States

^c Department of Electrical Engineering and Computer Sciences, University of California Berkeley, United States

^d Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign, United States

ARTICLE INFO

Article history:

Received 4 December 2011

Received in revised form 14 February 2013

Accepted 15 February 2013

Keywords:

Traffic flow

Macroscopic model

Driver heterogeneity

Cell-transmission model

Numerical method

Model benchmark

ABSTRACT

We consider the problem of modeling traffic phenomena at a macroscopic level. Increasing availability of streaming probe data allowing the observation of non-stationary traffic motivates the development of models capable of leveraging this information. We propose a phase transition model of non-stationary traffic in conservation form, capable of propagating joint measurements from fixed and mobile sensors, to model complex traffic phenomena such as hysteresis and phantom jams, and to account for forward propagation of information in congested traffic. The model is shown to reduce to the Lighthill–Whitham–Richards model within each traffic phase for the case of stationary states, and to have a physical mesoscopic interpretation in terms of drivers' behavior. A corresponding discrete formulation appropriate for practical implementation is shown to provide accurate numerical solution to the proposed model. The performance of the model introduced is assessed on benchmark cases and on experimental vehicle trajectories from the NGSIM datasets.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Motivation

The field of traffic modeling entails the design and analysis of mathematical tools for accurate representation of traffic dynamics. Road networks can be studied at different scales (Lesort et al., 2003). At a fine scale, the representation of traffic is closer to the true nature of the phenomena on the road. However, a detailed traffic model requires a large volume of accurate measurements for calibration and validation, and correspondingly large computation power. Consequently, different scales have been historically associated with different applications, depending on data availability and computational requirements.

At a *nanoscopic scale*, vehicles are considered to behave independently under the control of a driver who reacts to stimuli from neighboring vehicles according to a specific behavioral model. Traffic dynamics can be modeled as a set of coupled ordinary differential equations (ODE) with decision variables resulting from a demand model, see for instance Ahmed et al. (1996). At a *microscopic scale*, vehicles are considered to behave independently by reacting to stimuli from neighboring vehicles according to a dynamical model. Traffic dynamics can be modeled as a set of coupled ordinary differential equations

* Corresponding author.

E-mail addresses: sblandin@sg.ibm.com (S. Blandin), juan.argote@berkeley.edu (J. Argote), bayen@berkeley.edu (A.M. Bayen), dbwork@illinois.edu (D.B. Work).

(Brauer and Nohel, 1989). Nanoscopic and microscopic models have been used mostly for off-line non-real-time simulation and planning, see Fellendorf (1994) for VISSIM, and Cameron and Duncan (1996) for Quadstone Paramics.

At a *mesoscopic scale*, vehicles are considered as a large set of atomic elements with individual behavior according to macroscopic laws or relations. Traffic dynamics can be modeled using gas-kinetic models (Prigogine and Herman, 1971) or as cellular automata (Chowdhury et al., 2000). Mesoscopic models have been widely applied to real-time on-line and off-line network-wide control and dynamic traffic assignment, with significant prior data collection for calibration, see Mahmassani et al. (1994) for DYNASmart, and Ben-Akiva et al. (2002) for DynaMIT.

At a *macroscopic scale*, vehicles are considered to behave as a continuum medium. Traffic dynamics is modeled as a distributed system, using *partial differential equations* (PDE) inspired from hydrodynamics theory (Lighthill and Whitham, 1956; Richards, 1956; Garavello and Piccoli, 2006). Consequently, in this framework, the effect of network-wide route choices is not conveniently accounted for. One of the strengths of macroscopic models resides in the level of complexity they capture at a relatively low analytical and computational cost, and with limited data requirements for calibration. This has motivated the use of macroscopic models in particular for real-time on-line estimation and corridor management, see Messner and Papageorgiou (1990) for METANET, Chow et al. (2008) for TOPL, Bayen et al. (2011) for Mobile Millennium. Furthermore, the mathematical theory of hydrodynamics modeling brings a solid mathematical structure to macroscopic models. This theoretical strength can be leveraged for the development of mathematically sound finer-scale models, by equivalence, see Gazis et al. (1959), Helbing (2001), Hoogendoorn and Bovy (2001), and Lesort et al. (2003), and serve as an anchor for the development of consistent multi-scale modeling frameworks, such as the AIMSUN simulation software (Barcelo et al., 1998).

These properties of macroscopic models have motivated sustained research on extension of seminal models such as the *Lighthill–Whitham–Richards* (LWR) model, presented in the following section, with the goal of capturing complex observed phenomena missing from the LWR theory. Challenges in the design of so-called *higher-order models* relate to the development of sound physical understanding, well-defined analytical structure, and preserved computational tractability of the solution algorithm. In this article, we introduce and analyze a macroscopic flow model with phase transitions, posed as a perturbation around the LWR model. We show that allowing the speed function to take values around the classical stationary bivariate relation enables the simulation of higher-order traffic phenomena at the same computational cost as classical solution methods for the LWR model, and with a preserved physical interpretation.

1.2. First order scalar macroscopic models

The basis for the theory of macroscopic traffic modeling was set by the seminal articles from Lighthill and Whitham (1956) and Richards (1956), that introduced a PDE model describing the evolution of the density $k(t,x)$ of vehicles at time t and location x

$$\frac{\partial k(t,x)}{\partial t} + \frac{\partial Q(k(t,x))}{\partial x} = 0, \quad (1)$$

where the flow is expressed as a function of the density $Q(k(t,x))$. This so-called LWR model expresses the conservation of vehicles on the road. Different traffic models were later shown to be equivalent to the LWR formulation, in particular the Newell *car-following model* (Newell, 2002) and the *cell-transmission model* (CTM) (Daganzo, 1994), in the case of a triangular flux function $Q(\cdot)$ (Newell, 1993), also called *triangular fundamental diagram*:

$$Q(k) = \begin{cases} kv_{\max} & \text{if } k \leq k_c \\ (k - k_j)w & \text{if } k \geq k_c \end{cases} \quad (2)$$

where v_{\max} , k_c , k_j and w denote the *free-flow speed*, the *critical density*, the *jam density* and the *backward moving wave speed*, respectively, which are required to satisfy

$$k_c v_{\max} = (k_c - k_j)w \quad (3)$$

for continuity of the flow at the critical density. Density values below the critical density correspond to *free-flow* states and density values above the critical density correspond to *congestion* states. For non-vanishing values of the flow q and the speed v , the definition of the flux function $Q(\cdot)$ is equivalent to the definition of a speed function $V(\cdot)$ through the equality $q = vk$. The use of a single-valued fundamental diagram mapping a density value to a unique flow value dates back to the observations in Ohio by Greenshields (1935), in the Lincoln tunnel by Greenberg (1959), and in the Holland tunnel by Edie et al. (1963).

The LWR theory has been validated with experimental data for traffic modeling (see Nagel and Nelson, 2005 for instance), and in particular the triangular fundamental diagram from Eq. (2) has been shown to accurately model the *stationary relation* between density and flow, under proper time–space aggregation of traffic measurements (Castillo and Benitez, 1995; Cassidy and Coifman, 1997; Cassidy, 1998). One of the most desirable properties of the LWR model lies in its ability to capture the formation and growth of queues at bottlenecks. However the LWR model is known to be unable to reproduce more complex observed traffic phenomena such as *stop-and-go waves*, *traffic hysteresis*, and *phantom jams* (Hoogendoorn and Bovy, 2001).

1.3. Non-stationary traffic flow

The introduction of more complex macroscopic models for traffic flow dates back to the *Payne–Whitham* model (PW) (Payne, 1971), which consists in a 2×2 system¹ of PDEs. The first equation is the LWR PDE (see Eq. (1)) and the second equation models the acceleration of vehicles as resulting from a reaction to local traffic conditions and a relaxation around the stationary relation. In a discrete form with a minor modification at on-ramps (Papageorgiou et al., 1990), this model was later used extensively for estimation and control (Wang and Papageorgiou, 2005; Wang et al., 2007).

Several remarks were raised (Castillo et al., 1994; Daganzo, 1995) on the lack of physical consistency of so-called higher-order models, in which vehicles with negative speed, and *anisotropy* property, were shown to arise. Anisotropy characterizes the fact that modeled drivers react to stimuli from the front and from behind.² These considerations led to the development of improved non-scalar models that did not exhibit these flaws; using a convective derivative in Aw and Rascle (2000), and a so-called *non-equilibrium* model in Zhang (2002).

Independently, Kerner proposed the *three-phase theory* (Kerner and Rehborn, 1996, 1997; Kerner, 1998) postulating the existence of three phases of traffic; *free-flow*, *synchronized flow*, and *moving jam*, instead of the classical *free-flow* and *congestion*. So-called *third order* PDE models inspired from physical systems such as the model from Helbing (1995), introducing the *speed variance* as a traffic state, have also been proposed.

The relative strengths and weaknesses of first and second-order models have been outlined in several articles (Castillo et al., 1994; Daganzo, 1995; Papageorgiou, 1998; Lebacque and Lesort, 1999; Hoogendoorn and Bovy, 2001; Lesort et al., 2003), including discussions on the expected modeling abilities of extensions to state-of-the-art macroscopic traffic models. This coincides with the emergence of sustained research focused on the understanding and modeling of specific phenomena missing from the LWR theory; *stop-and-go* waves, *hysteresis* patterns, *capacity drop*, and the understanding, validation, and modeling of their candidate causes; *lane changes*, *heterogeneous drivers*, *bounded acceleration*, *acceleration and deceleration curves*. Complementary efforts have investigated the nature of these phenomena and proposed models to reproduce them, with specific emphasis on:

- Development of parsimonious models able to reproduce the capacity drop (Papageorgiou et al., 1990), hysteresis patterns (Zhang, 2002), stop-and-go waves (Li, 2005).
- Validation of the causality hypothesis and development of models for bounded acceleration (Lebacque, 1997), acceleration and deceleration waves (Yeo and Skabardonis, 2009), lane changes (Laval, 2005; Laval and Daganzo, 2006; Jin, 2010), heterogeneous drivers (Chiabaut et al., 2010; Laval and Leclercq, 2010).

From the perspective of intelligent transportation systems, a critical feature of traffic models consists in their ability to handle streaming measurements for accurate estimation of traffic conditions (Wang and Papageorgiou, 2005; Mihaylova et al., 2006; Work et al., 2010; Blandin et al., 2012). Observability of quantities required for on-line model calibration, ability to take advantage of measurements of various traffic quantities, computational tractability and model accuracy are essential properties for traffic monitoring.

In this work, we propose the analysis of the *phase transition model* (PTM), shown to be an extension of classical scalar first order macroscopic flow models derived from the LWR model described by Eq. (1). The PTM was introduced by Colombo (2002, 2003) following Kerner' remarks (Kerner and Rehborn, 1996, 1997). It was later extended in Blandin et al. (2011) to account for a constant speed in free-flow, and a variety of possible stationary fundamental diagrams in congestion. The corresponding system of PDEs has been shown to be well-posed and an efficient numerical scheme has been proposed to compute its numerical solution (Chalons and Goatin, 2008). In particular, the PTM does not lead to vehicles with negative speeds, and does not exhibit anisotropy.

The focus of the present work is on the assessment of accuracy and practicality of the PTM for traffic modeling and on-line traffic estimation. We show that, by allowing the classical state variable density k to be complemented in congestion by a perturbation variable p at the stationary state, the PTM is able to *propagate* the impact of non-stationary dynamics, rather than *predict* its emergence. It is observed that for time–space diagram reconstruction from initial and boundary conditions, significant accuracy is gained from the ability to account for joint observations of different traffic quantities. Specific attention is given to the physical mesoscopic interpretation of the model, and the practicality of implementation of the associated discrete solution algorithm. In particular the convexity of the state-space of the model, required for the use of finite volume schemes such as the Godunov scheme, is assessed. The performance of the proposed model and its ability to model complex traffic phenomena such as *hysteresis patterns*, *phantom jams*, forward-moving discontinuities in congestion, is assessed on benchmark cases and on experimental vehicle trajectories from the *Next Generation SIMulation* (NGSIM) datasets (NGSIM, 2006).

The main contributions of the present article are the following:

¹ A set of two equations for a 2-tuple of state variables.

² The justification for this appellation is the fact that in a so-called anisotropic model, we have $\lambda_1 \leq v \leq \lambda_2$, where v denotes the speed of vehicles, and $\lambda_{1,2}$ the characteristic speeds at which information propagate.

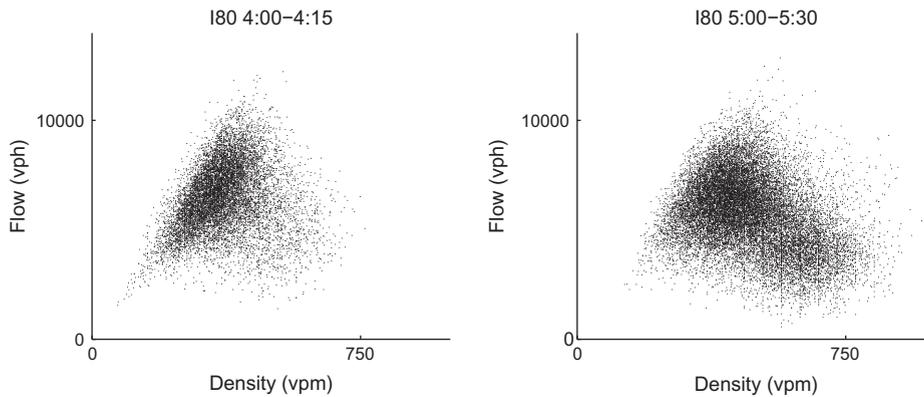


Fig. 1. Observed fundamental diagrams for NGSIM datasets. The mapping of density to flow, for quantities corresponding to the aggregation of all mainlines, is not single-valued in congestion. The datasets and numerical discretization parameters are described in Section 4.1.

1. *Extension of LWR theory:* we show that the PTM can be seen as a first order extension of the LWR theory at the fundamental diagram in the congestion phase. An appropriate discretization scheme, the *modified Godunov scheme*, is motivated by a physical interpretation, and described in details.
2. *Increased modeling capabilities:* we show that the PTM is able to model forward-moving shock waves and hysteresis phenomena in the congestion phase. We illustrate that the model can create phantom jams in specific scenarios.
3. *Performance assessment on real measurements:* we compare the performance and analyze the respective strengths of the calibrated CTM and PTM on the NGSIM datasets for I-80.

The remainder of the article is organized as follows. In Section 2 we introduce the phase transition model in the continuous setting. In Section 3 we describe an appropriate solution algorithm that can be viewed as a discrete version of the phase transition model. Section 4 describes the NGSIM datasets used for numerical illustrations and validation. Advanced modeling capabilities of the PTM are presented in Section 5, on synthetic benchmark cases and on experimental datasets. Calibration and analysis of the PTM for the NGSIM datasets is proposed in Section 6. Finally, Section 7 makes concluding remarks and presents related research issues on data fusion.

2. Continuous phase transition model

In this section, we present the continuous phase transition model (Blandin et al., 2011), derive its mathematical structure in the context of non-stationary traffic, and illustrate its applicability by solving the corresponding *Riemann problem* (Garavello and Piccoli, 2006).

2.1. Model definition

Similar to first-order models, the phase transition model consists of a PDE modeling the dynamics of traffic, and an empirical parametric fundamental diagram representing a relation from density to flow.

To account for the spread of the density-flow relation observed in congestion in experimental datasets (see Fig. 1), and in accordance with the definition of the fundamental diagram as a stationary relation, we consider a new speed function $V_{PTM}(\cdot)$ that can be considered to be a first-order extension of the classical density-speed relation in congestion:

$$V_{PTM}(u) \doteq \begin{cases} v_{\max} & \text{in free-flow} \\ V(k)(1+p) & \text{in congestion} \end{cases} \quad (4)$$

where $u \doteq (k, p)$ is a 2-tuple where k is the density and p is an additional real-valued variable, independent of the density, that models a perturbation around the stationary relation. The free-flow speed is again denoted v_{\max} , and $V(\cdot)$ is the stationary speed function associated with the flux function $Q(\cdot)$ in congestion. Note that once the stationary relation $V(\cdot)$ is calibrated, the perturbation p can be computed from joint observations of density and speed, or density and flow, according to relation (4). The flow of vehicles is naturally defined as the product of speed and density:

$$Q_{PTM}(u) \doteq kV_{PTM}(u) = \begin{cases} kv_{\max} & \text{in free-flow} \\ kV(k)(1+p) & \text{in congestion.} \end{cases} \quad (5)$$

When the perturbation vanishes ($p = 0$), the fundamental diagram defined in (4), or equivalently (5), reduces to the classical relation between stationary states. Values of speed higher or lower than the stationary fundamental diagram are admissible for non-stationary states ($p \neq 0$). Fig. 2 illustrates that the perturbation can be allowed to change sign (top and bottom rows) or be restricted to a constant sign (middle row for the case of a positive sign).

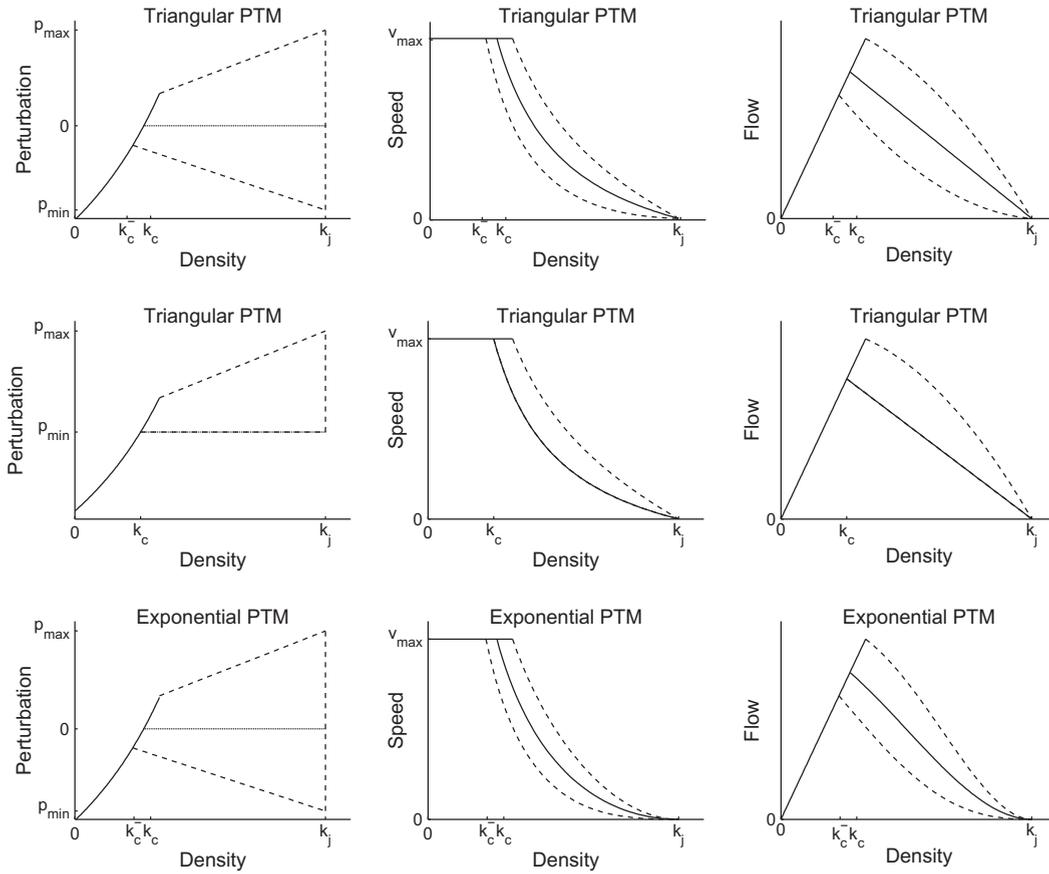


Fig. 2. PTM fundamental diagrams: in density-perturbation coordinates (left column), density-speed coordinates (center column), density-flow coordinates (right column) for the triangular stationary relation (top and center rows) and for the exponential stationary relation (bottom row). For low values of density, the solid line corresponds to the free-flow phase. For high values of density, the solid line corresponds to stationary congestion states. The congestion phase is the two-dimensional domain delimited by the dashed lines, around the stationary congestion states. The perturbation can take both positive and negative values (top and bottom rows), or have constant sign (center row).

The parameters of the fundamental diagram are the parameters of the stationary relation, and the bounds p_{\min} and p_{\max} on the perturbation (see Fig. 2). To guarantee the positivity of speeds, it is clear that it is necessary to have the following inequality:

$$p_{\min} > -1. \tag{6}$$

In particular, the PTM guarantees that vanishing speeds only arise at densities corresponding to vanishing stationary speed, which is the property P2 of Garavello and Piccoli (2009). Further details on the impact of the choice of the bounds p_{\min} and p_{\max} for the perturbation are discussed in Section 2.2 on model analysis and in Section 6 on model validation.

The PTM fundamental diagram is defined as an extension of the stationary diagram in congestion and is able to accommodate different stationary profiles. For illustration purposes, the triangular fundamental diagram from Eq. (2) is shown in Fig. 2, top and middle rows, and the case of the exponential fundamental diagram is illustrated in Fig. 2, bottom row.

Remark 1. The perturbation p defined in Eq. (4) accounts for variations of traffic speed around its stationary values. Note that no specific cause is attached to the value of the perturbation. This corresponds to the implicit modeling assumption made in this article that, at an aggregate level, the evolution of traffic flow is determined only by the current state (density, perturbation and associated values of flow, density), and by road properties captured by the set-valued PTM fundamental diagram, irrespective of other causal variables (lane changes, heterogeneous drivers, acceleration and deceleration waves).

The dynamics of the phase transition model is governed by different systems of PDEs for the free-flow phase and for the congestion phase

$$\begin{cases} \text{Free-flow : } \frac{\partial k}{\partial t} + \frac{\partial(kv)}{\partial x} = 0 \\ \text{Congestion : } \begin{cases} \frac{\partial k}{\partial t} + \frac{\partial(kv)}{\partial x} = 0 \\ \frac{\partial p}{\partial t} + \frac{\partial(pv)}{\partial x} = 0, \end{cases} \end{cases} \tag{7}$$

where in Eq. (7), the speed v is defined by $v \doteq V_{PTM}(k,p)$ according to Eq. (4), and thus has a different algebraic expression in free-flow and in congestion. In free-flow, the evolution of the density k satisfies the LWR model, described in Eq. (1). In congestion, the evolution of the state composed of the density k and the perturbation p is governed by two coupled conservation laws stating that these two quantities are conserved, and that they propagate at a speed v given by the fundamental diagram defined in Eq. (4). One might note that according to (7), the variable p denotes a *density* of perturbation. Similar to the quantity $k(t,x)$ corresponding to the total number of vehicles at time t in a neighborhood of x , averaged over the spatial size of the neighborhood, the quantity $p(t,x)$ can be understood as representing the sum of proportional speed deviations from the stationary speed, at the observed $k(t,x)$, due to the fact that all drivers in a neighborhood of x may not drive at the stationary speed $V(k(t,x))$. In that sense the quantity p can be computed as the sum of these deviations averaged over the considered neighborhood, and understood as a density of perturbation. For concision, we refer to p as a perturbation throughout the article.

The free-flow phase and the congestion phase are formally defined in the following section, in which we construct the solution to the PDE (7).

2.2. Model analysis

In this section, we present the properties of the solution to the *initial-boundary value problem* (IBVP) associated with the PTM, defined in Eq. (7). We remind the reader that the variable u denotes the 2-tuple (k,p) . For $(t,x) \in [0,T] \times [a,b]$, the IBVP reads

$$\left\{ \begin{array}{l} \text{PTM :} \\ \text{Fundamental diagram :} \\ \text{Initial condition :} \\ \text{Boundary condition :} \end{array} \right. \left\{ \begin{array}{l} \text{Free-flow :} \\ \text{Congestion :} \\ v = V_{PTM}(u) \\ u(0,x) = u_0(x) \\ \begin{cases} u(t,a) = u_a(t) \\ u(t,b) = u_b(t), \end{cases} \end{array} \right. \quad \left\{ \begin{array}{l} \frac{\partial k}{\partial t} + \frac{\partial(kv)}{\partial x} = 0 \\ \frac{\partial k}{\partial t} + \frac{\partial(kv)}{\partial x} = 0 \\ \frac{\partial p}{\partial t} + \frac{\partial(pv)}{\partial x} = 0 \end{array} \right. \quad (8)$$

where u_0, u_a, u_b denote the initial, upstream and downstream boundary conditions respectively. Solving the IBVP defined in (8) (on the time-space domain $[0,T] \times [a,b]$) consists in constructing the time-space diagram of traffic corresponding to the dynamics of the PTM, the fundamental diagram $V_{PTM}(\cdot)$, observations u_a and u_b from upstream and downstream sensors located at $x = a$ and $x = b$ respectively, and knowledge of the traffic state u_0 between sensors at some initial time 0. The construction of the solution to the IBVP is guided by the analysis of the system of PDEs ruling the evolution of the system in congestion. We refer the interested reader to Blandin et al. (2011) for specific details on the following derivations, and in particular for the result that the system of PDEs from Eq. (7) corresponding to the congested phase is hyperbolic.³ Note that the boundary conditions cannot be imposed in the strong sense as literally described in (8), but need to be imposed in the weak sense, see Strub and Bayen (2006) for the case of the LWR PDE.

The eigenvectors associated with the Jacobian of the congestion part of the system of Eq. (7) are

$$e_1(u) \doteq \begin{pmatrix} k \\ p \end{pmatrix}, \quad e_2(u) \doteq \begin{pmatrix} V(k) \\ -(1+p)V'(k) \end{pmatrix} \quad (9)$$

with the respective eigenvalues

$$\lambda_1(u) \doteq V_{PTM}(u) + \nabla V_{PTM}(u) \cdot e_1(u), \quad \lambda_2(u) \doteq V_{PTM}(u) \quad (10)$$

and with the associated Riemann invariants (Toro, 1997):

$$r_1(u) \doteq \frac{p}{k}, \quad r_2(u) \doteq V_{PTM}(u). \quad (11)$$

The details of the analysis leading to these results are available in Blandin et al. (2011), in particular the fact that the second term in the first eigenvalue in (10) is negative, which guarantees the anisotropy of the system. Eqs. (9) and (10) indicate that in congestion, the system can evolve either along trajectories tangent to the eigenvector e_1 , or along trajectories tangent to the eigenvector e_2 . Eq. (11) indicates that the quantity p/k is invariant along the trajectories of the first type (tangent to e_1), and the speed v is invariant along trajectories of the second type (tangent to e_2).

Remark 2. We remind the reader that the LWR theory describes the emergence of contact discontinuities, shock waves, and rarefaction waves, at spatial discontinuities of the solution to a scalar conservation law. With a concave diagram, a positive spatial gradient of density corresponds to a shock wave, and a negative spatial gradient of density corresponds to a

³ A system of conservation laws is called hyperbolic if the Jacobian of the system is diagonalizable with real eigenvalues. Most of the theory of systems of conservation laws is concerned with hyperbolic systems.

rarefaction wave (Ansoorge, 1990) (the converse is true for a convex diagram). In the specific case for which the fundamental diagram is affine, contact discontinuities arise, which model the propagation of a discontinuity without mixing of the upstream and downstream states, unlike a shock-wave.

First-type trajectories: since p , by definition, denotes a perturbation around the stationary state, the first invariant defined in Eq. (11) can be understood as the average perturbation per driver. A complete analysis (Blandin et al., 2011) shows that the first type of trajectories is *genuinely non-linear* (GNL) in general, hence can give rise to rarefaction waves or shock waves, as in the LWR theory. In particular, the trajectory of the first type for which $p/k = 0$ is conserved, i.e. for a vanishing perturbation, corresponds to the evolution of the LWR solution.

Second-type trajectories: the second type of trajectories, for which the speed v is conserved, is *linearly degenerate* (LD), and gives rise to contact discontinuities, which correspond to discontinuities propagating without mixing of the upstream and downstream state. This type of evolution does not exist for the congestion phase in the LWR model, but has been observed in practice (Cassidy and Windover, 1995), and is identical to the evolution of traffic in the free-flow phase of the LWR model with a triangular fundamental diagram. These discontinuities propagate downstream at the speed of vehicles.

For stability of the model, it is clear that a trajectory of the system should not exit the admissible state space ($[0, k_j]$ for a scalar model of density). This requires us to define the boundary of the admissible domain as the extremal admissible trajectories of the system. Consequently, we define the congestion phase T_c as a domain delimited by extremal values of the Riemann invariants (11), that are invariant along trajectories:

$$T_c \doteq \left\{ (k, p) \mid \frac{p_{\min}}{k_j} \leq \frac{p}{k} \leq \frac{p_{\max}}{k_j} \text{ and } 0 \leq v \leq v_{\max} \right\},$$

where p_{\min} and p_{\max} denote the minimal and maximal physically admissible values, respectively, for the perturbation variable, and k_j , v_{\max} denote the jam density and free-flow speed, respectively. The curves $p/k = p_{\min}/k_j$ and $p/k = p_{\max}/k_j$ correspond to the lower and upper envelope, respectively, of the congestion domain depicted in dashed line in Fig. 2 (with $p_{\max} > 0$ for the three rows, $p_{\min} < 0$ for the top and bottom rows, and $p_{\min} = 0$ for the middle row).

Remark 3. In order to enforce positivity of speed, the condition described in Eq. (6) for the lower bound p_{\min} must be satisfied. Characteristics speeds along first-type trajectories are negative only if the parameter p_{\max} is not too large, see Blandin et al. (2011) for analytical derivations for different models.

The free-flow phase T_f is defined as:

$$T_f \doteq \{k \mid 0 \leq k \leq k_c^-\} \quad (12)$$

where the point defined by $k = k_c^- \leq k_c$ corresponds to the intersection of the free-flow phase and the lower envelope of the congestion phase, see Fig. 2. For continuity of the flow at k_c between the free-flow phase and the stationary relation in congestion, the condition $V(k_c) = v_{\max}$ must be satisfied. This relation corresponds to Eq. (3) for the triangular diagram.

Remark 4. The PTM exhibits similarities with the non-equilibrium model (Zhang, 2002); in the congestion phase, both models have trajectories of the first type that are GNL in general and trajectories of the second type that are LD. Moreover, the second Riemann invariant, defined in Eq. (11) right for the PTM, is the speed for both the non-equilibrium model and the PTM. Aside from the existence of different dynamics for the free-flow phase and the congestion phase in the case of the PTM, with the use of a classical single-valued density-flow relation in free-flow according to empirical observations, the main difference between the two models lies in the nature of the first type of trajectories. For these trajectories, the Riemann invariant is $v - V(k)$ in the non-equilibrium model, whereas it is p/k in the PTM. Consequently, for the non-equilibrium model, the first type of trajectories can be obtained by a vertical translation of the equilibrium curve in (k, v) coordinates, whereas for the PTM, these trajectories converge toward a vanishing speed at the jam density. This is illustrated for the triangular fundamental diagram in Fig. 3.

The interested reader is referred to Colombo (2003) and Blandin et al. (2011) for results of well-posedness of the Cauchy problem associated with the PTM, to Colombo et al. (2007) for the well-posedness of the IBVP associated with the PTM, to Dafermos (2010) for more details on conservation laws in general, and to Garavello and Piccoli (2006) for more details on conservation laws in the context of traffic flow on networks. In the next section, we present the Riemann problem and a Riemann solver for the PTM.

2.3. Riemann problem

A critical property of a traffic model is its ability to provide simple, physically correct solutions to elementary benchmark scenarios of interest for traffic (Lebacque and Lesort, 1999). For hyperbolic conservation laws, a classical benchmark scenario is a Riemann datum, which is a centered, piecewise constant, initial condition:

$$u(t = 0, x) = \begin{cases} u_{\text{up}} & \text{if } x < 0 \\ u_{\text{down}} & \text{if } x > 0. \end{cases} \quad (13)$$

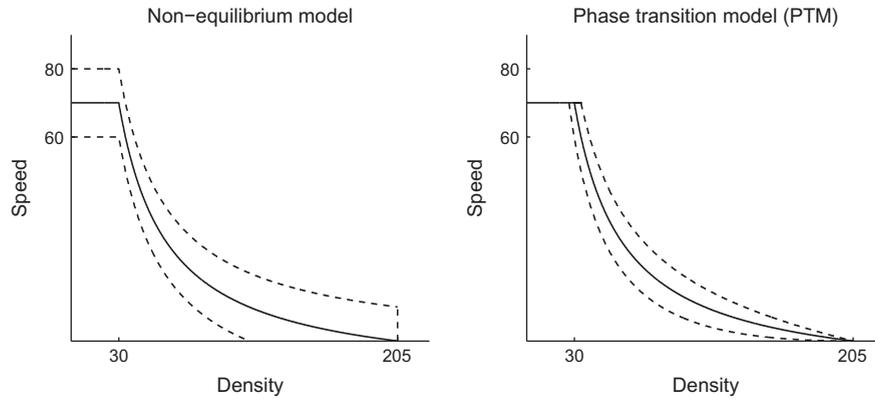


Fig. 3. Non-stationary density speed relations: for the non-equilibrium model (Zhang, 2002), left, and the phase transition model, right. The spread in the allowable speed is 10 mph above and below the classical relation at the critical density for the non-equilibrium model, and 10 mph below the classical relation at the critical density for the PTM. The non-equilibrium model captures a set-valued density speed relation in congestion and in free-flow, delimited by the dashed lines. The PTM captures a set-valued density speed relation in congestion, delimited by the dashed lines, and a single-valued density speed relation in free-flow, indicated by the solid line for low values of density.

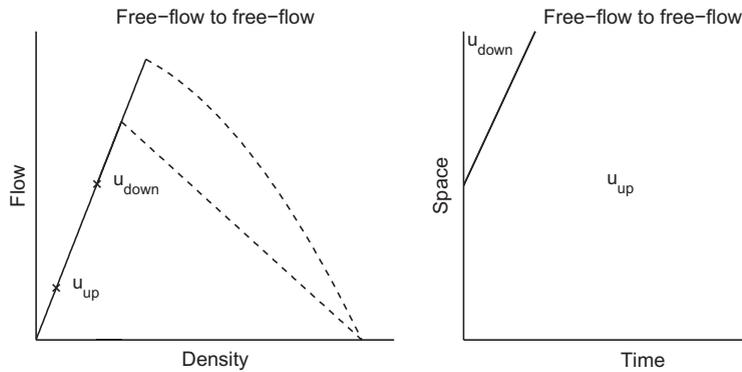


Fig. 4. Riemann problem between free-flow states: a contact discontinuity propagating at speed v_{max} connects the upstream state u_{up} and the downstream state u_{down} .

A Riemann problem is defined as a Cauchy problem for an initial datum of the type (13). The Riemann solver is the mathematical method constructed to provide the solution to the Riemann problem. In other words the Riemann solver produces the evolution in time of the initial traffic conditions defined by the Riemann datum (13). In the case of the PTM, $u \doteq (k, p)$ in congestion and $u \doteq k$ in free-flow. The Riemann solver defined in Blandin et al. (2011), in the case of system trajectories of the first type with concavity of constant sign, is the following^{4,5}:

1. If $(u_{up}, u_{down}) \in T_f \times T_f$: the solution to the Riemann problem for the PTM coincides with the solution to the Riemann problem for the LWR model. It consists of a contact discontinuity⁶ from u_{up} to u_{down} . This is illustrated in Fig. 4.
2. If $(u_{up}, u_{down}) \in T_c \times T_c$: we denote the intermediary state by $u_m \doteq (k_m, p_m)$, defined as the point of the congestion phase T_c satisfying

$$\begin{cases} \frac{p_m}{k_m} = \frac{p_{up}}{k_{up}} \\ V_{PTM}(u_m) = V_{PTM}(u_{down}). \end{cases}$$

The solution to the Riemann problem consists of a shock wave or rarefaction wave along a curve of the first type from u_{up} to u_m , followed by a contact discontinuity along a curve of the second type from u_m to u_{down} . This is illustrated in Fig. 5, for the

⁴ It is clear that along trajectories of the first type, shock waves or rarefaction waves can occur. In the interest of space, in this section we provide graphical illustrations for only one of the two types of waves for each scenario.

⁵ The illustrations in this section are provided for the case of a triangular stationary relation, with positive perturbation ($p_{min} = 0$) for simplicity.

⁶ A contact discontinuity is a wave occurring for a characteristic field with constant characteristic speed. The speed of the discontinuity corresponds to the left and right identical characteristic speeds.

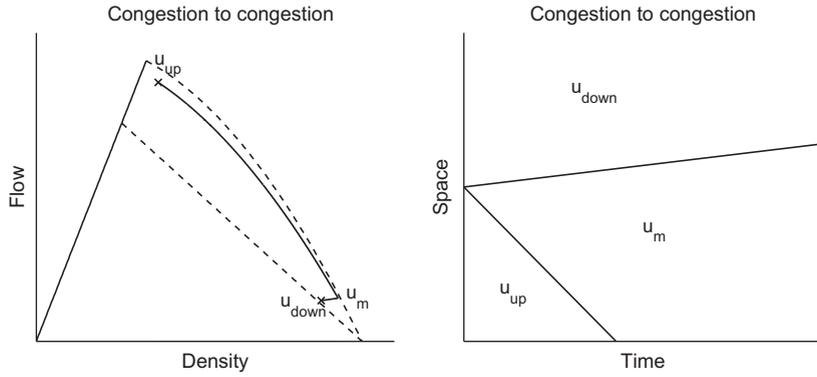


Fig. 5. Riemann problem between states in congestion: an intermediary state u_m arises between the upstream state u_{up} and the downstream state u_{down} .

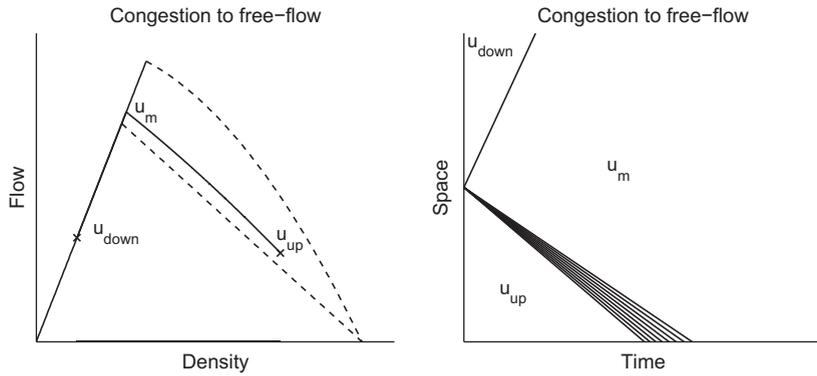


Fig. 6. Riemann problem between upstream state in congestion and downstream free-flow state: an intermediary state u_m in the free-flow phase arises between the upstream state u_{up} and the downstream state u_{down} .

case in which a shock wave, arises from u_{up} to u_m .

3. If $(u_{up}, u_{down}) \in T_c \times T_f$: we note $u_m \doteq (k_m, p_m)$ the point of the congestion phase T_c defined by:

$$\begin{cases} \frac{p_m}{k_m} = \frac{p_{up}}{k_{up}} \\ V_{PTM}(u_m) = V_{PTM}(u_{down}) = v_{max}. \end{cases}$$

The solution to the Riemann problem consists of a shock wave or rarefaction wave along a curve of the first type from u_{up} to u_m , followed by a contact discontinuity along a curve of the second type from u_m to u_{down} . This is illustrated in Fig. 6, for the case in which a rarefaction wave, represented by a fan on the right sub figure, arises from u_{up} to u_m .

4. If $(u_{up}, u_{down}) \in T_f \times T_c$: we note $u_m \doteq (k_m, p_m)$ the point of the congestion phase T_c defined by

$$\begin{cases} \frac{p_m}{k_m} = \frac{p_{min}}{k_j} \\ V_{PTM}(u_m) = V_{PTM}(u_{down}), \end{cases}$$

and we note $Q_{PTM}(\cdot)$ the flux function defined in Eq. (5). We assume that the entropy condition

$$\frac{Q_{PTM}(u_{up}) - Q_{PTM}(u_m)}{k_{up} - k_m} \geq \lambda_1(u_m),$$

which guarantees that the shock-like phase transition is entropic, is satisfied⁷. The solution to the Riemann problem consists of a shock-like phase transition from u_{up} to u_m , followed by a contact discontinuity along a curve of the second type from u_m to u_{down} . This is illustrated in Fig. 7 for the case of a phase transition with negative speed.

⁷ The entropy condition states that the slope, in density-flow coordinates, of the line connecting u_{up} and u_m , is greater than the first characteristic speed at u_m (with notations of Eq. (10)).

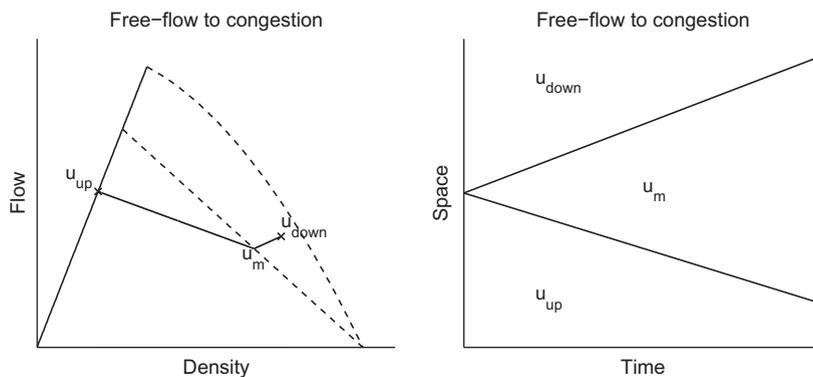


Fig. 7. Riemann problem between upstream free-flow and downstream state in congestion: an intermediary state u_m in the congestion phase arises between the upstream state u_{up} and the downstream state u_{down} .

In the following section we leverage the definition of the Riemann solver from this section to introduce a discretized model for the PTM, derived as a modification of the seminal Godunov scheme (Godunov, 1959).

3. Discretized phase transition model

In this section, we present a formulation of the PTM in discrete time and discrete space, which has been shown (Chalons and Goatin, 2008; Blandin et al., 2011) to provide an accurate numerical approximation to the solution of the PTM in a continuous setting (7). This so-called *modified Godunov scheme* is derived from the seminal Godunov scheme, and we present a physical interpretation from a mesoscopic perspective.

3.1. Discrete model definition

We consider a discretization of time and space in intervals of size Δt , Δx , respectively, and note $u_j^n = (k_j^n, p_j^n)$, $j = 1, \dots, J$, $n = 1, \dots, N$, the value of the numerical solution in the discretization cell centered at $j\Delta x$, $(n + 1/2)\Delta t$.

The classical Godunov scheme consists of two steps, generally summarized by a conservative formulation, involving a numerical Godunov flux, for the dynamics of the discrete solution. First, the solution to the Riemann problems generated at time $n\Delta t$ by the couples (u_j^n, u_{j+1}^n) is computed on the time interval $[n\Delta t, (n + 1)\Delta t]$. Second, the solution to the Riemann problems is averaged at time $(n + 1)\Delta t$ on each cell $[(j - 1/2)\Delta x, (j + 1/2)\Delta x]$.

The main modification to the original Godunov scheme is due to the existence of phase transitions, which can yield a state outside the admissible state-space after the averaging step of the classical scheme. This can be prevented by considering that the discrete solution at $j\Delta x$, $n\Delta t$, applies to a space extent $\Delta \bar{x}_j^n$ different from Δx , and which belongs to a single phase. We note \bar{u}_j^n the discrete solution on the modified cells. The modified Godunov scheme consists of the following sequential steps:

- *Forward propagation*: computation of the discrete solution \bar{u}_j^{n+1} on the modified grid with cells of size $\Delta \bar{x}_j^n$. This step consists of computing the average of the solutions to the Riemann problems between neighboring (u_j^n, u_{j+1}^n) , on modified cells containing a single phase, free-flow or congestion.
- *Sampling*: computation of the discrete solution u_j^{n+1} on the regular grid by sampling from the discrete solution \bar{u}_j^{n+1} on the modified grid.
- *Projection*: projection of the solution onto the congested phase, in the case of a cell boundary between congestion and congestion or congestion and free-flow.

These steps are described in more detail in the following sections.

3.1.1. Forward propagation

The discretized model used for propagating the state forward is defined as follows:

$$\Delta \bar{x}_j^{n+1} \bar{u}_j^{n+1} = \Delta x u_j^n - \Delta t \left[q_{MG}^{\text{down}}(u_j^n, u_{j+1}^n) - q_{MG}^{\text{up}}(u_{j-1}^n, u_j^n) \right] \quad (14)$$

where $q_{MG}^{\text{up}}(u_a, u_b)$ and $q_{MG}^{\text{down}}(u_a, u_b)$ denote the upstream and downstream numerical modified Godunov flux, respectively, between the state u_a upstream and the state u_b downstream. The numerical modified Godunov flux is defined as follows:

$$\left\{ \begin{array}{l}
 \text{If } (u_a, u_b) \in T_f \times T_f \\
 \quad q_{MG}^{up}(u_a, u_b) = q_{MG}^{down}(u_a, u_b) = k_a v_{max} \\
 \text{If } (u_a, u_b) \in T_c \times T_c \\
 \quad q_{MG}^{up}(u_a, u_b) = q_{MG}^{down}(u_a, u_b) = (k_m V_{PTM}(u_m), p_m V_{PTM}(u_m)) \\
 \text{If } (u_a, u_b) \in T_c \times T_f \\
 \quad q_{MG}^{up}(u_a, u_b) = k_b v_{max} - k_b \sigma_{PT}(u_m, u_b) \\
 \quad q_{MG}^{down}(u_a, u_b) = (k_m V_{PTM}(u_m), p_m V_{PTM}(u_m)) - u_m \sigma_{PT}(u_m, u_b) \\
 \text{If } (u_a, u_b) \in T_f \times T_c \\
 \quad q_{MG}^{up}(u_a, u_b) = (k_m V_{PTM}(u_m), p_m V_{PTM}(u_m)) - u_m \sigma_{PT}(u_a, u_m) \\
 \quad q_{MG}^{down}(u_a, u_b) = k_a v_{max} - u_a \sigma_{PT}(u_a, u_m)
 \end{array} \right. \quad (15)$$

where u_m denotes the intermediary point, defined in Section 2.3, arising between the upstream state u_a and downstream state u_b . The velocity function $V_{PTM}(\cdot)$ is introduced in Eq. (4). The last right-hand-side term in the third and fourth cases of Eq. (15) denotes the speed $\sigma_{PT}(u_a, u_b)$ of the phase transition arising between the upstream state u_a and the downstream state u_b . The phase transition speed can be computed using the Rankine–Hugoniot relation between the states across the phase transition

$$\sigma_{PT}(u_a, u_b) \doteq \begin{cases} \frac{Q_{PTM}(u_a) - Q_{PTM}(u_b)}{k_a - k_b} & \text{if } u_a, u_b \text{ belong to different phases} \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

where $Q_{PTM}(\cdot)$, introduced in Eq. (5), denotes the flow of vehicles. One may note that the Rankine–Hugoniot speed in the first line of Eq. (16) expresses the conservation of the density k of vehicles at discontinuities. This guarantees the conservation of vehicles across phase transitions in the third and fourth cases of Eq. (15). The quantity $\Delta \bar{x}_j^{n+1}$ is defined as:

$$\Delta \bar{x}_j^{n+1} \doteq \Delta x + \Delta t \left(\sigma_{PT}(u_j^n, u_{j+1}^n) - \sigma_{PT}(u_{j-1}^n, u_j^n) \right) \quad (17)$$

Note that $\Delta \bar{x}_j^{n+1}$ may be different from Δx only if a phase transition arises at time $n\Delta t$ at the upstream or downstream boundary of the cell centered at $j\Delta x$. The perturbation term p is independent from the density only in congestion. In free-flow, the perturbation term is uniquely determined by the value of the density (see Fig. 2), hence it does not appear in the definition of the free-flow phase (12), does not have specific dynamics (7) in free-flow, and its value is not meaningful in free-flow. This explains why it does not have to be conserved at phase transitions (third and fourth cases of Eq. (15)), since it is a meaningful quantity only on one side of the phase transition. Within the congestion phase, it is clearly conserved (second case of Eq. (15)).

3.1.2. Sampling

In the case of neighboring cells belonging to different phases (third and fourth cases of Eq. (15)) the solution of the Riemann problem must be integrated on a domain that belongs to the same phase rather than on each cell. This modification to the original Godunov scheme is required due to the non-convexity of the union of the congestion and free-flow phases $T_f \cup T_c$, which implies that the averaging step of the classical Godunov scheme might produce a state outside of the admissible domains. This explains the additional last term in the third and fourth cases of (15), compared to the classical modified Godunov scheme. The additional terms in the third and fourth cases of Eq. (15) specify that the averaging of the solution to the Riemann problem is done only within each phase.

The discrete solution u_j^{n+1} on the regular grid is computed by sampling uniformly in space between the values $\bar{u}_{j-1}^{n+1}, \bar{u}_j^{n+1}, \bar{u}_{j+1}^{n+1}$ on the spatial domain corresponding to the cell j . In the case represented in Fig. 8, u_j^{n+1} is sampled from \bar{u}_{j-1}^{n+1} and \bar{u}_j^{n+1} , which cover the cell j at time $n + 1$, according to their rate of presence in the cell. The value u_{j-1}^{n+1} is taken equal to \bar{u}_{j-1}^{n+1} since the whole spatial domain corresponding to cell j at time $n + 1$ belongs to the domain of existence of \bar{u}_{j-1}^{n+1} . The value u_{j+1}^{n+1} is sampled from \bar{u}_j^{n+1} and \bar{u}_{j+1}^{n+1} , which cover the cell $j + 1$ at time $n + 1$, according to their rate of presence in the cell. It has been shown that the Van der Corput sequence, known to be uniformly distributed, is appropriate for this procedure (Chalons and Goatin, 2008).

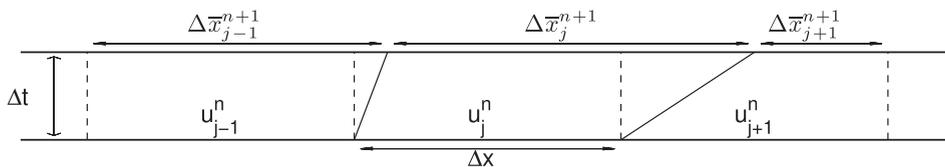


Fig. 8. Phase transitions: arising between neighboring cells are depicted in oblique solid lines. The boundaries between regular cells are depicted in vertical dashed lines. The solution u_j^{n+1} at time $t^{n+1} = (n+1)\Delta t$ is defined on cells of size Δx_j^{n+1} delimited by the trajectory of phase transitions. This figure corresponds to the case where only two phase transitions arise at the 4 cell boundaries represented.

3.1.3. Projection

The definition of the modified Godunov flux (15) guarantees that no averaging on the solution to the Riemann problem between states belonging to different phases occurs. However, the congestion phase T_c itself is not convex at the intersection with the free-flow phase (see Fig. 2 left column). To account for this non-convexity, a projection step has been proposed in Blandin et al. (2011) and shown to give accurate results, due to the low curvature of the phase boundary at this location. This step consists in projecting any non-admissible value resulting from the modified Godunov scheme onto the congestion phase, along the first trajectories (that preserve the first Riemann invariants). This process is illustrated in Fig. 11, domain B.

Remark 5. The congestion phase T_c of the PTM is convex at each boundary except at the boundary corresponding to the free-flow speed (see Fig. 2). This limited non-convexity can be corrected by a third step in the modified Godunov scheme (see boxed algorithm below). In the case of the non-equilibrium model, the non-convexity of the congestion domain ranges from the critical density to the jam density (see Fig. 3 left), unless a Greenshields (linear) equilibrium velocity function is used. Consequently the use of a similar modified Godunov scheme would yield a more significant error in general than in the case of the PTM.

In summary, at each time $(n+1)\Delta t$, the discrete solution $u_j^{n+1}, j = 1, \dots, J$ to the PTM can be computed from the discrete solution $u_j^n, j = 1, \dots, J$ at the previous time $n\Delta t$ as follows.

1. Forward propagation:

- compute $\Delta \bar{x}_j^{n+1}$ according to Eq. (17) for cells j for which a phase transition arises at a boundary.
 - compute $\bar{u}_j^{n+1}, j = 1, \dots, J$, using the discrete model dynamics described in Eqs. (14) and (15).
2. *Sampling:* compute the value $u_j^{n+1}, j = 1, \dots, J$ by sampling the candidate values $\bar{u}_{j-1}^{n+1}, \bar{u}_j^{n+1}, \bar{u}_{j+1}^{n+1}$ on the domain corresponding to the cell j .
3. *Projection:* if a state with velocity greater than the free-flow speed v_{\max} and admissible first Riemann invariant arise, project along first Riemann invariant on the congestion phase (see Fig. 11, domain B).

A formulation similar to the supply–demand for scalar conservation laws, which is known to be equivalent to the Godunov scheme, can also be derived for the above algorithm within each phase. In the following section, we present a physical perspective on the PTM dynamics.

3.2. Mesoscopic interpretation

In this section we show that a meaningful mesoscopic interpretation, outlined in Blandin et al. (2010), can be associated with the discrete numerical method presented in the previous section and with the solution to the Riemann problem in Section 2.3, see Daganzo et al. (1999) for a related discussion on a *Markovian* perspective for modeling of complex traffic phenomena. In particular, we provide a physical interpretation for the numerical flux between two cells that depends on their corresponding phases.

First, we note that according to its definition in Eq. (4), the perturbation variable p indicates how much the speed of the associated traffic state deviates from the stationary state. Hence from the mesoscopic perspective it can be viewed as an indication of the aggressiveness of the corresponding element of flow. It follows that the quantity p/k is the average aggressiveness per driver; it is positive if the observed speed is higher than the stationary speed, and negative otherwise.

Free-flow upstream to free-flow downstream: the flux between two cells in the free-flow phase T_f is the flow from the upstream cell (first case of Eq. (15)). This is in accordance with the definition of the free-flow phase as a phase where the traffic demand is lower than the traffic supply, hence the demand from the upstream cell can be accommodated by the downstream cell, and the corresponding upstream flow can be realized between the two cells.

Congestion upstream to congestion or free-flow downstream: the flux between two cells in the congestion phase is the flow of the intermediary state u_m (second and third case of Eq. (15)). In this case, the middle state u_m has same value of p/k as the upstream state u_{up} , and same speed v as downstream state u_{down} (see second and third case of Section 2.3). The fact that the numerical flux between two cells is defined by the average aggressiveness of the upstream state u_{up} and the speed of the downstream state u_{down} can be understood as follows: drivers from the upstream cell adapt their speed to the speed of the drivers from the downstream cell. They adapt their speed and pick their spacing according to their natural driving behavior, represented here by the average aggressiveness of the element of flow to which they belong. Hence the numerical flux between the two cells is the flow that corresponds to a personalized modification, by the upstream drivers from the state u_{up} , of their speed, in order to match the speed of the downstream drivers.

Free-flow upstream to congestion downstream: the flux between two cells in the congestion phase is the flow of the intermediary state u_m (fourth case of Eq. (15)). In this case, the middle state u_m has the lowest possible average aggressiveness p/k compatible with the speed v of the downstream state u_{down} (see fourth case of Section 2.3). This can be understood as follows; in the free-flow phase, the demand is not constrained by the supply and drivers are free to drive at the free-flow speed, hence the traffic state is entirely defined by its density. In particular, the aggressiveness of drivers is uniquely determined from the value of the density. When the drivers from the free-flow phase enter the congestion phase, they switch from a phase in which the demand is the defining constraint to a phase in which the supply is the defining constraint. This naturally

yields a degree of aggressiveness that represents their driving speed compared to the stationary speed. Since the free-flow drivers emerge from a free-flow phase, the average aggressiveness of their corresponding element of flow is as low as possible. However, they have to adapt their speed according to the speed of the downstream traffic. They modify their speed to reach the speed of the downstream drivers, which is the second defining element for the intermediary state. In that sense the aggressiveness of drivers is created by the interaction of a free-flow phase upstream and a congestion phase downstream. This is similar to the fact that congestion in the LWR model arises only from junctions.

4. Vehicle trajectories datasets

In this section we present the datasets used for numerical illustrations of the PTM physical properties in Section 5.

4.1. Datasets specifications

We consider the following NGSIM datasets (NGSIM, 2006) from highway I-80:

1. *I-80 4:00–4:15*: vehicles trajectories are recorded at a 0.1 s resolution for a 0.34 miles stretch of freeway I-80 Northbound at Powell Street, Emeryville, CA, with six lanes including a HOV lane, during 15 min from 4 pm to 4:15 pm (2052 vehicles). We consider the middle stretch from relative mile 0.03 to relative mile 0.3, which is a straight line. We remove the first and last 100 s to avoid boundary effect. We discretize the dataset into 14 cells and 749 time-steps.
2. *I-80 5:00–5:30*: vehicles trajectories are recorded at a 0.1 s resolution for a 0.34 miles stretch of freeway I-80 Northbound at Powell Street, Emeryville, CA, with six lanes including a HOV lane, during 30 min from 5 pm to 5:30 pm (3626 vehicles). We consider the middle stretch from relative mile 0.03 to relative mile 0.3, which is a straight line. We remove the first 100 and last 250 s to avoid boundary effect. We discretize the dataset into 14 cells and 1104 time-steps.

The choice of the number of cells for discretization of the dataset is driven by the sampling frequency of the collected data for accurate computation of flow, density and speed according to Edie's generalized definition (Edie, 1963). Under this definition, the parameters that characterize the traffic conditions in a time–space domain A can be obtained as:

$$k(A) = t(A)/|A| \quad (18a)$$

$$Q(A) = d(A)/|A| \quad (18b)$$

$$V(A) = d(A)/t(A), \quad (18c)$$

where $k(A)$, $Q(A)$, and $V(A)$ are the density, flow, and speed in domain A respectively, $|A|$ is the domain area (in units of time–distance), $t(A)$ is the total time spent in A by all vehicles that cross the domain, and $d(A)$ is the total distance traveled by those vehicles.

We want to have a time-step “large” compared to the sampling period and a cell size “large” compared to the distance traveled by the vehicles during a sampling period, see Lesort et al. (2003) for a detailed discussion on the notion of scale in traffic modeling. Here we consider a time-step of 10 times the sampling period, and a cell size of 10 times the distance traveled at the free-flow speed (taken as 70 mph) between two consecutive reports. The average point speed of the vehicles is 17 mph and 12 mph for I-80 4:00–4:15 and I-80 5:00–5:30, respectively. The discretization parameters for each dataset are given in Table 1.

In the following analysis, we consider the computed values of density, flow and speed according to Edie's generalized Definition 18, for the discretization grid associated with the parameters from Table 1, as *ground-truth* traffic state for benchmarking the performances of the PTM and the CTM.

4.2. Dataset properties

The density time–space diagrams of the NGSIM datasets used for model testing and validation are reproduced in Fig. 9. In this figure, the five leftmost mainline lanes are considered for computation of the density, flow, and speed according to Edie's generalized definition, with discretization parameters described in the previous sub section.

Several heavy congestion patterns can be observed propagating upstream at similar speeds. Heavy congestion can also be observed propagating downstream in the I80 5–5:30 dataset, starting from the upstream boundary at around 750 s in relative time.

Table 1

Discretization parameters: the time-step and cell size are chosen as a function of the sampling period.

	Sampling period	Time-step	Cell size	Average # points per cell
I-80 4–4:15	0.1 s	1 s	102 feet	71 points
I-80 5–5:30	0.1 s	1 s	102 feet	96 points

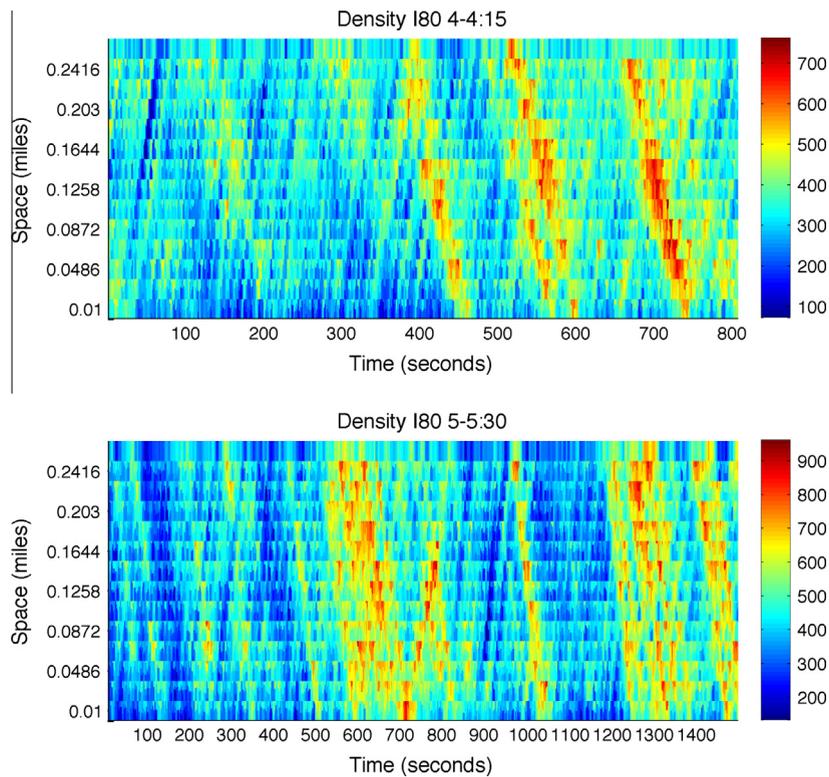


Fig. 9. Density (vpm) time–space diagrams for I80 4–4:15 (top) and I80 5–5:30 (bottom).

One may note that the datasets considered exhibit only congestion phenomena, hence only allow to validate the PTM in the congestion phase. This is a clear limitation of the experimental results presented in this article, that should be complemented in the future by an analysis of the performance of the PTM on datasets more representative of general traffic properties.

5. Model properties

In this section we present the properties of the PTM on benchmark tests and on the experimental vehicle trajectories described in Section 4.

5.1. Set-valued fundamental diagram

In this section, we quantify how the set-valued nature of the congestion phase of the PTM fundamental diagram increases the accuracy with which non-stationary traffic observations are accounted for by the admissible domain of the model. We propose a method for projecting state values falling outside of the state space back onto the state-space for observations either erroneous or not predicted by the model, and compare its performance with a similar method for the LWR model. In an estimation setting, the figures presented in this section indicate appropriate range of values for the error statistics associated with the boundary of the state-space.

5.1.1. PTM theory

A common issue faced by applied researchers lies in the possible incompatibility of real measurements and state-space values of the theoretical model. In this section, we propose a simple method for projecting measurements that do not belong to the model state-space onto its boundary, and we illustrate the associated error for both the PTM and the LWR model.

The PTM state variables include the density k and an additional perturbation variable p in congestion that captures non-stationary traffic states observed in practice (see Fig. 10). However, for measurements outside of the union $T_f \cup T_c$ of the free-flow phase and the congestion phase, a method has to be designed to project the observations onto the admissible domain.

For the LWR model, this can be accomplished by assuming that measurements of density outside of the range $[0, k_j]$ are not valid or correspond to the endpoints $\{0, k_j\}$, that measurements of speed outside of the range $[0, v_{\max}]$ are not valid or

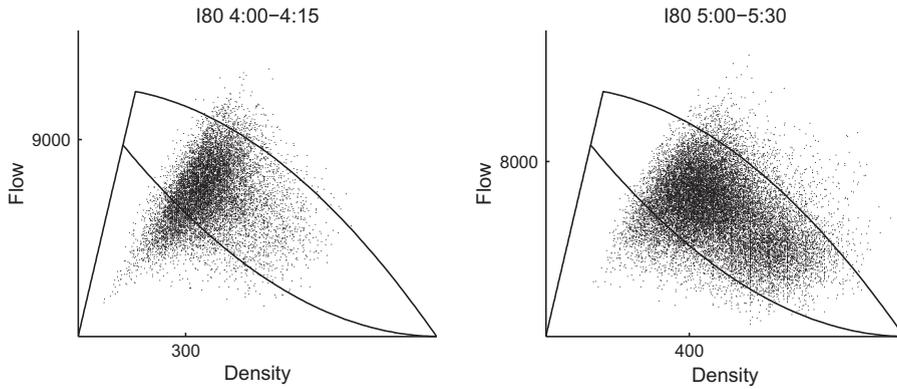


Fig. 10. Set-valued congestion phase: the PTM triangular fundamental diagram is able to capture to some extent the cloud of points observed in the congestion phase. The figures correspond to a visual fit of a triangular stationary relation with positive and negative perturbation, and a free-flow speed $v_{\max} = 70$ mph. Note that the apparent linear edge on the left of the cloud of points, does not correspond to the free-flow speed v_{\max} , but to the maximal speed of vehicles in the conditions considered, approximately 30 mph and 20 mph for the left and right sub figure respectively.

correspond to the endpoints $\{0, v_{\max}\}$, and that a measurement of a given quantity automatically yields another quantity using the stationary relation.

In the case of the PTM, for traffic observations falling outside of the domain $T_f \cup T_c$, we propose to project along the eigen-trajectories of the system, i.e. along the curves of constant average aggressiveness p/k or constant speed v . Here we address the case of perfect measurements. The projection method is illustrated in Fig. 11.

Remark 6. In the case of a joint measurement from different traffic quantities, for instance loop counts and loop occupancies, or probe speeds and loop counts, the set-valued nature of the PTM fundamental diagram in congestion explicitly captures a two-dimensional admissible domain for this joint traffic measurement. In the case of measurement of a single quantity, according to the PTM, in congestion, the measurement corresponds to a set of admissible states. Additional assumptions can be made to restrict this set to a single value, for instance the assumption of stationary state, used in the LWR model.

5.1.2. Empirical validation

In this section we illustrate the average absolute error associated with the process of projection onto the fundamental diagram for the CTM and the PTM. We consider the discretized values of flow, density, and speed from the NGSIM datasets, as measurements of the ground-truth traffic state, and we assume that the associated measurement error (due to sensing and processing error) can be neglected in the analysis presented in this section.

For points falling outside of the fundamental diagram of the PTM, we compute the average absolute error between the measurement, and its projection onto the diagram using the method described in the previous section. For this illustration,

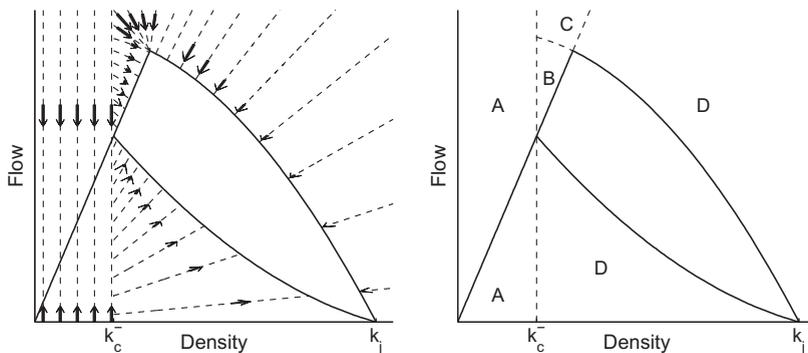


Fig. 11. Projection onto PTM diagram: outliers with a density corresponding to the free-flow phase (domain A in right figure) are projected along iso-density curves (vertical dashed lines in left figure). Outliers with higher density than the maximal free-flow density (k_c) and higher speed than the free-flow speed (domain B in right figure) are projected along iso-average aggressiveness curves, or trajectories of the first type (dashed curves in left figure), except for outliers with higher speed than the free-flow speed and higher average aggressiveness than its maximum allowed value (domain C in right figure), which are projected onto the maximal capacity point of the diagram (dashed curves in left figure). Outliers with higher density than the maximal free-flow density and lower speed than the free-flow speed (domain D in right figure) are projected along iso-speed curves, or trajectories of the second type (dashed lines emanating from the origin in left figure).

the fundamental diagram parameters chosen are the optimal parameters computed in the following section using density error metric (see Tables 3 and 4, first line, in the case of the PTM). The associated density and speed errors are represented in the second and last columns of Table 2, respectively.

We also represent the error associated with using the stationary relation in the CTM. Specifically, in the first column of Table 2, we present the error on density when observing speed and computing the density from the stationary relation. In the third column of Table 2, we present the error on speed when observing density and computing the speed from the stationary relation.

It is clear from Table 2 that inferring the density from speed measurements and the stationary relation yields a very high error (first column). One may also note that for the PTM, since the projection is often realized along iso-speed trajectories, there is a very low resulting error on speed.

The respective accuracy of the dynamics of the two models is presented in Section 6.2.

5.2. Forward-moving discontinuity in congestion phase

One of the specific features of the stationary bivariate relation for traffic is that information propagates downstream in free-flow, and upstream in congestion. The PTM has a two-dimensional state in congestion that allows for two speeds of propagation of information. The PTM is usually calibrated to have the first speed of propagation negative (see Section 6.1) according to the stationary state theory. The second speed of propagation is always positive (see Section 2.2), and corresponds to the speed of vehicles at the downstream state. In this section we study the increased modeling capabilities brought by this feature.

5.2.1. PTM theory

In the congestion phase, two types of waves can arise in the PTM solution (see Section 2.2). Waves of the second type are contact-discontinuities; they connect two states $u_{up} = (k_{up}, p_{up})$ and $u_{down} = (k_{down}, p_{down})$ with identical traffic speed $v = V_{PTM}(u_{up}) = V_{PTM}(u_{down})$, and move at the speed v of traffic. This phenomenon is illustrated in Fig. 12.

This type of waves models the propagation without dispersion of traffic phases with different density and flow, but identical speed. Since the discontinuity propagates at the speed of traffic, this phenomenon does not depend on the sign of the flow difference between the upstream and the downstream phase, and, unlike for a shockwave, no mixing is introduced between the two states on each side of the discontinuity. In particular, the interface between two groups of vehicles with different densities and flow but identical speed is preserved with time.

5.2.2. Empirical validation

The aggregate time–space diagrams for density and speed in the I80 5–5:30 dataset are presented in Fig. 13. In addition to several backward moving shock waves spanning the entire domain, we observe a clear forward-moving discontinuity in the time–space diagram for density, emanating from the upstream of the section around time 750 s, while no forward-moving discontinuity appears in the time–space diagram for speed. A number of forward-moving discontinuities are observed during the episodes of light congestion (blue in Fig. 13). One may note that the forward-moving discontinuities for the episodes of light congestion travel faster than the forward moving discontinuity corresponding to heavy congestion (red in Fig. 13), as the PTM would predict. Finally the comparison of the top and bottom plots of Fig. 13 illustrates that discontinuities in density can propagate forward or backward, whereas discontinuities in speed only propagate backward, which corresponds to the constitutive properties of the PTM detailed in Section 2.2. The ability of the model to reproduce this phenomenon is illustrated in Section 6.2.

5.3. Hysteresis phenomenon

The phenomenon of hysteresis has been studied with much attention by the transportation community (Newell, 1962; Treiterer and Myers, 1974; Zhang, 1999), with different candidate explanations; acceleration and deceleration waves, heterogeneous drivers, lane changes. Recent results (Laval, 2011) seem to discard the hypothesis of acceleration and deceleration waves in favor of the hypothesis of heterogeneous drivers. The model proposed in this article is consistent with this theory in the sense that loops can form in both directions, irrespective of the speed gradient.

Table 2

Average projection error: average absolute error associated with the stationary state hypothesis for the CTM and with the projection on the set-valued fundamental diagram for the PTM. The first column corresponds to the computation of density from speed observation using the stationary relation, and the converse for the third column (LWR model in both cases). The second column corresponds to the error in density due to projection, and similarly for the speed in the fourth column (PTM in both cases). For each dataset, the error is averaged for all values of density, speed, and flow represented in Fig. 10, obtained from vehicle trajectories according to the method described in Section 4.1 to handle NGSIM data.

	Density CTM error (vpm)	Density PTM error (vpm)	Speed CTM error (mph)	Speed PTM error (mph)
I-80 4–4:15	58.9	8.5	5.4	0.1
I-80 5–5:30	61.0	5.2	3.4	0.0

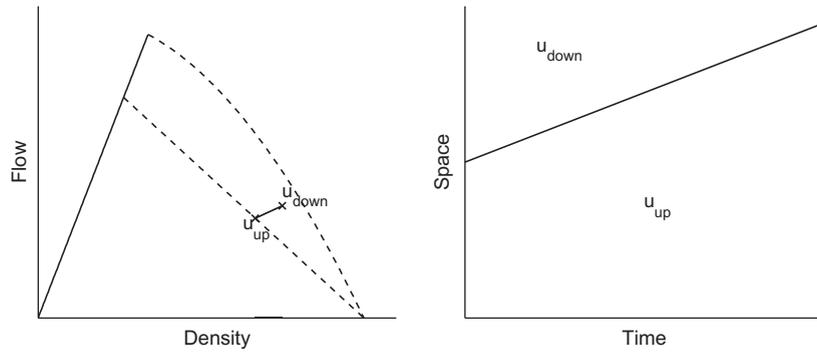


Fig. 12. Forward-moving discontinuity in congestion: the interaction between states u_{up} and u_{down} with the same speed, yields a forward-moving contact discontinuity in the congestion phase. Inverting the location of u_{up} and u_{down} on the road also yields a forward-moving contact discontinuity.

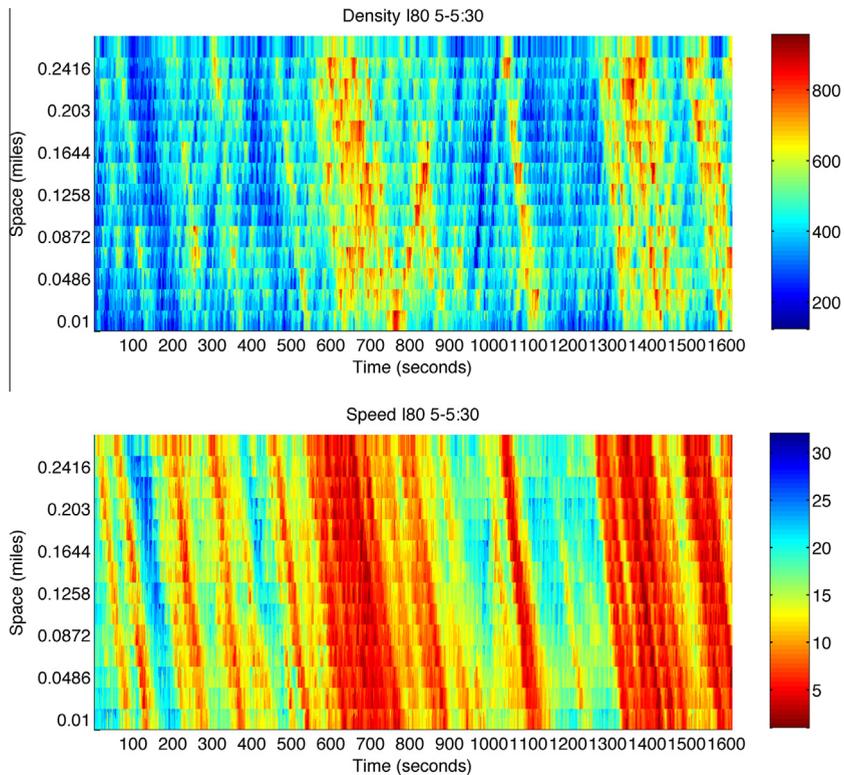


Fig. 13. I80 5–5:30 time–space diagrams: density in vpm (top) and speed in mph (bottom). A forward-moving discontinuity corresponding to heavy congestion (in red) arises from the upstream boundary around time 750 s. No forward-moving discontinuity in speed is observed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The existence of hysteresis loops in density–flow coordinates can be traced back to the solution to the Riemann problem described in Section 2.3 (second case of u_{up} , u_{down} , in the congestion phase T_c). An intermediary state u_m arises between the initial upstream state u_{up} and the initial downstream state u_{down} . The upstream state u_{up} and the intermediary state u_m have the same value of the average aggressiveness per driver p/k . The downstream state u_{down} and the intermediary state u_m have the same value of speed v . The relative value of these two quantities for the upstream and downstream states u_{up} and u_{down} impacts the orientation of the loop (see Fig. 14).

Only two configurations are possible in congestion for which no hysteresis loop arises.

- In the case of an upstream state u_{up} and a downstream state u_{down} with the same value of the average aggressiveness p/k , no hysteresis occurs, and the only difference introduced by switching the two states is the change of nature of the connecting wave (shock or rarefaction).

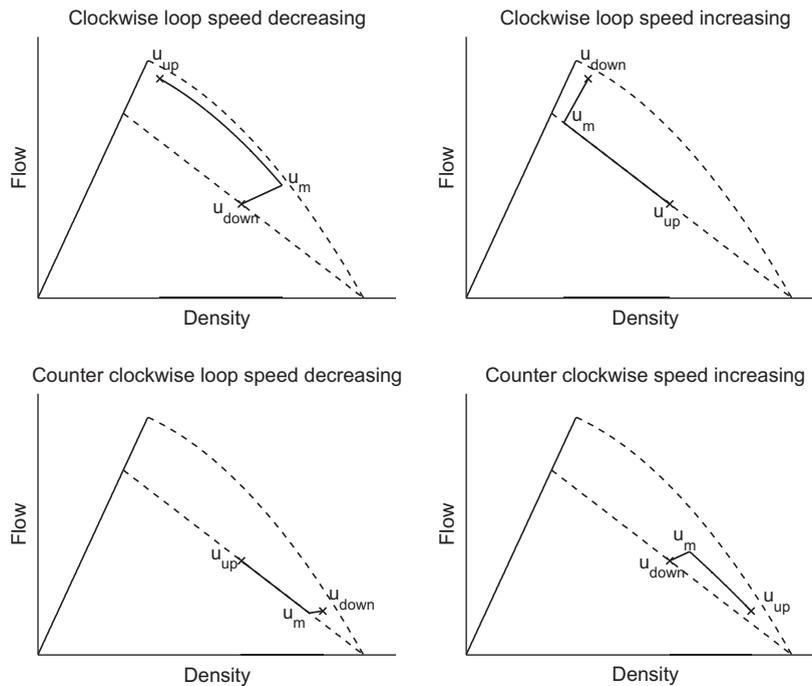


Fig. 14. Hysteresis patterns: between upstream and downstream states, u_{up} and u_{down} , respectively. Clockwise loops (top row) introduce values of the density outside of the interval defined by the initial densities of the upstream and downstream states. Counterclockwise loops (bottom row) introduce values of the flow outside of the interval defined by the flows of the initial upstream and downstream states.

- In the case of an upstream state u_{up} and a downstream state u_{down} with same value of speed v , the connecting wave is a contact discontinuity for the two permutations of the initial setting, and corresponds to the forward-moving discontinuity described in Section 5.2.

The type of hysteresis phenomena modeled by the PTM is similar to the hysteresis described in Zhang (2002), and can be physically explained using the mesoscopic interpretation presented in Section 3.2. The NGSIM dataset used in this study unfortunately did not enable us to observe any of these phenomena directly (the dataset is very limited in time and space). Note that these hysteresis phenomena are for macroscopic quantities, not for trajectories, see Laval (2011) for instance.

5.4. Phantom jam

The emergence and propagation of traffic disturbances causing drivers to decrease their speed for no clear reason in congestion, or so-called phantom jams, is well-known to most commuters and the topic of active research, with recent explanations tracing its cause back to the heterogeneity of driving behaviors (Laval and Leclercq, 2010). It is clear from the solution to the Riemann problem defined in Section 2.3 that the PTM is not able to model the emergence of extreme⁸ values of speed. However, in this section we show that the PTM is able to model the emergence of extreme values of density and flow.

We motivate the subsequent macroscopic description by the mesoscopic model described in Section 3.2, in which drivers from an upstream phase adjust their speed to the speed of the drivers from the downstream phase, according to their own driving behavior. One might note that this mesoscopic model can be obtained similarly by expressing the fact that upstream drivers maximize their speed under the constraint of their driving behavior, and the speed of the downstream drivers, that they cannot exceed without creating an accident.

In the congestion phase, the solution to the Riemann problem associated with the PTM exhibits different types of hysteresis loops (see previous section). It is clear that the intermediary state arising always has a speed in the interval defined by the speeds of the upstream and downstream states. However, different loop behaviors can arise (see Fig. 14):

1. Clockwise loops are the cause of the emergence of intermediary states with extreme values of density (top row). In the case of a transition from a high speed to a low speed, a higher value of density arises (top left case) whereas in the case of a transition from a low speed to a high speed a lower value of density arises (top right case).

⁸ In the context of a Riemann problem, by extreme value of speed we mean a value of speed outside of the interval defined by the speeds of the upstream and downstream states.

2. Counter-clockwise loops are the cause of the emergence of intermediary states with extreme values of flow (bottom row). In the case of a transition from a high speed to a low speed, a lower value of flow arises (bottom left case) whereas in the case of a transition from a low speed to a high speed a higher value of flow arises (top right case).

In the four cases described above, from a mesoscopic perspective, the difference in flow and density between the downstream state u_{down} and the middle emerging state u_m is explained by the fact that the upstream and downstream drivers have different levels of aggressiveness, translating into different preferred spacing for the same speed.

This feature of the PTM seems appropriate for modeling the propagation of disturbances in traffic. In Fig. 14, one can note that if u_{down} is viewed as a disturbance of u_{up} , then the intermediary state u_m arising is always an amplified disturbance, in density or flow, and this amplified disturbance travels upstream (negative wave speed connecting u_{up} and u_m). Note that this phenomenon is a locally convected phenomenon, i.e. it results from the emergence of u_m from u_{up} and u_{down} . It is very different from convective instabilities commonly observed in fluids, characterized by a dispersion relation (Saffman, 1992), that amplify over time.

6. Model validation

In this section we present a method for calibrating the parameters of the PTM, and numerical results obtained on the datasets described in Section 4.

6.1. Model calibration

The parameters of the PTM consist of the free-flow speed, the parameters of the stationary relation used in the congestion phase (4), and the parameters p_{min} and p_{max} , specifying the maximal admissible range around the stationary relation for the fundamental diagram in congestion. The parameters are depicted for several different fundamental diagrams in Fig. 2.

Several algorithms can be found in the literature for calibration of a single model parameter, in particular in the case of the CTM. The jam density k_j and the free-flow speed v_{max} are usually assumed to be known, and the congestion wave speed w can be estimated using different methods. An algorithm for the estimation of the congestion wave speed based on vehicle trajectories can be found in Lu and Skabardonis (2007). This parameter can similarly be estimated considering a Lagrangian approach as described in Chiabaut et al. (2009).

In this section, we consider a methodology similar to the methodology described in Cremer and Papageorgiou (1981), for joint estimation of all parameters of the model from macroscopic quantities.

6.1.1. Methodology and results

We propose to compare the time space diagrams reconstructed by the PTM and the CTM from the knowledge of initial and boundary conditions. Given a training data set of macroscopic measurements, the method consists of the definition of a cost function, the direct computation of the initial and boundary condition, and of the identification of the set of parameters at which the minimum of the cost function is attained for the reconstructed time space diagrams. We consider the L^1 metric:

$$L^1(u_{\text{PTM}}) = \frac{\sum_n \sum_j |u_{\text{PTM}}(n\Delta t, j\Delta x) - u_{\text{NGSIM}}(n\Delta t, j\Delta x)|}{\sum_n \sum_j |u_{\text{NGSIM}}(n\Delta t, j\Delta x)|}, \quad (19)$$

in which we note $u_{\text{PTM}}(n\Delta t, j\Delta x)$ the value of the quantity u , at the time-step indexed by n , at the cell indexed by j , computed using the discrete PTM on a grid with discretization parameters Δt and Δx . We use a similar notation for the ground-truth u_{NGSIM} obtained by discretizing the NGSIM vehicle trajectories.

The quantities u_{NGSIM} and u_{PTM} must be defined on the same discretization grid. The reference u_{NGSIM} is defined on the physical grid described in Section 4.1, chosen according to data availability. The discrete solution u_{PTM} to the PTM can be computed on an arbitrarily refined numerical grid. The discrete solution converges toward the analytical solution when the refinement of the numerical grid increases. The level of refinement of the grid is guided by the numerical benchmarks from Blandin et al. (2011) and Chalons and Goatin (2008), which provide empirical results on the distance to the analytical solution of the PDE as a function of the refinement level. The values of the reference u_{NGSIM} on the refined grid can be obtained from its values on the coarse physical grid in a straightforward manner.

In order to assess the importance of the error variable used for calibration, we propose to compute the error metric (19) when u is successively defined as k , q and v . Due to the non-linearity of the models investigated, and the consequent non-convexity of the optimization problem considered, the cost function is optimized by exhaustive enumeration on a grid with parameters $\Delta k = 10$ vpm, $\Delta v = 5$ mph, $\Delta w = 0.5$ mph, $\Delta p_{\text{min}} = 0.1$, $\Delta p_{\text{max}} = 0.1$. The optimal parameters for I80 4:00–4:15 and I80 5:00–5:30 are presented in Tables 3 and 4 respectively. In the interest of space, we only consider the instantiation of the PTM for a triangular stationary relation, with potentially positive and negative perturbation, as depicted in Fig. 2, top row.

The optimal parameters obtained fall into the range of plausible values from a physical perspective. There is significant dependency of the optimal parameters to the error variable (which arises for the CTM as well, see Table 5). One might note that the optimal value of the congestion wave speed obtained for the five aggregated lanes is greater than the typical value

Table 3

Optimal parameters for I80, 4:00–4:15: PTM parameters that minimize the L^1 error metric, obtained from the ground truth field for density (k), flow (q) and speed (v).

Error variable	Optimal parameters					L^1 Error		
	k_j^*	v_{\max}^*	w^*	p_{\min}^*	p_{\max}^*	k	q	v
Density (k)	160	50	14	−0.01	0.94	0.152	0.185	0.164
Flow (q)	150	40	15.5	−0.21	0.77	0.164	0.176	0.157
Speed (v)	120	40	13.5	−0.99	0.99	0.160	0.190	0.137

Table 4

Optimal parameters for I80, 5:00–5:30: PTM parameters that minimize the L^1 error metric, obtained from the ground truth field for density (k), flow (q) and speed (v).

Error variable	Optimal parameters					L^1 error		
	k_j^*	v_{\max}^*	w^*	p_{\min}^*	p_{\max}^*	k	q	v
Density (k)	190	55	13	−0.25	0.95	0.130	0.170	0.165
Flow (q)	180	55	13.5	−0.75	0.45	0.141	0.164	0.170
Speed (v)	150	55	13	−0.05	0.35	0.145	0.172	0.161

for a single lane (around 11.5 mph), which might be due to tentative lane changes. The optimal value of the congestion wave speed is fairly stable across error variables, however the jam density exhibits large relative variations, in particular for the I80 datasets in which the congestion level is more important.

The optimal parameters obtained for different error variables can vary significantly for a given dataset, in particular for the perturbation parameters. In the case of the density error variable, one might note that the parameter p_{\max} often takes larger values than the parameter p_{\min} in absolute value. This corresponds to a fundamental diagram with a wider domain in congestion above the stationary relation than below, and using the mesoscopic interpretation from Section 3.2, to a greater number of aggressive drivers than non-aggressive drivers.

One might note that a natural calibration of the PTM, which would consist in using the parameters of a calibrated CTM for the stationary relation of the PTM, and second in maximizing the spread between p_{\min} and p_{\max} in order to minimize the projection error described in Section 5.1, is not obtained as a result of the optimization procedure. The optimal values of the classical parameters k_j and w of the PTM, are very similar to the optimal values of the parameters of the CTM (see Table 5), however the parameters p_{\min} and p_{\max} are not often set to their extremal values. We recall that the parameter p_{\min} is lower bounded by -1 to guarantee positivity of speed (6), and that large values of the parameter p_{\max} correspond to positive first characteristic speed in congestion, which is not desirable for physical reasons. This result illustrates that the set-valued diagram of the PTM is valuable for accurate modeling of traffic state, but that there is a trade-off between a wide congestion phase and congestion dynamics close to the dynamics associated with the stationary relation.

In the next section, we analyze the sensitivity of the model to the parameters, in the case of density as error variable.

6.1.2. Model sensitivity

The sensitivity of the error metric (19) to the parameters around the optimum indicates the order of magnitude of the error likely to occur in a practical setting, where uncertainties arise in the calibration procedure. These uncertainties can be due to inherent measurement noise, to numerical error in the optimization routine, or to the fact that the parameters are calibrated from a visual fit. Thus the ability of the model to guarantee good performances for parameters in a neighborhood of the optimum value is an important factor contributing to the model choice.

We propose to assess the sensitivity of the error metric to the parameters in two different ways. We compute the *partial variation* of the error metric around the optimum, i.e. the variation of the error metric when a single parameter varies around the optimum. We also compute the *total variation* of the error, i.e. the variation of the error metric when a single parameter varies around the optimum, and consequently the other parameters are re-calibrated according to this change.

Fig. 15 presents the results for the congestion wave speed w , and Fig. 16 presents the results for the perturbation parameters p_{\min} and p_{\max} . One may note that for both models, the *total variation* of the error metric is relatively low for variations around the optimal value, i.e. re-calibration of the other parameters of the model is able to account for a lack of optimality of the considered parameter.

The results for the *partial variation* of the congestion wave speed differ significantly for the CTM and the PTM. It is clear that lower values of the congestion wave speed have a relatively low impact on the error metric, for both models. However, the CTM error is highly sensitive to higher values of the congestion wave speed, which can be explained by noting that the optimal free-flow speed is low, and that subsequently, modifying the value of the congestion wave speed toward higher values impacts the phase to which observations around the optimal critical density belong to. This is not the case when modifying the value of the congestion wave speed toward lower values. This high sensitivity, in the case of higher values than the optimal congestion wave speed, does not occur with the PTM, which may be due to the two-dimensional nature of the con-

Table 5

Optimal parameters for I80, 4:00–4:15: CTM parameters that minimize the L^1 error metric, obtained from the ground truth field for density (k), flow (q) and speed (v).

Error variable	Optimal parameters			L^1 error		
	k_j^*	v_{\max}^*	w^*	k	q	v
Density (k)	160	25	13.7	0.158	0.198	0.180
Flow (q)	170	28	13.4	0.159	0.184	0.162
Speed (v)	170	31	13.7	0.161	0.185	0.159

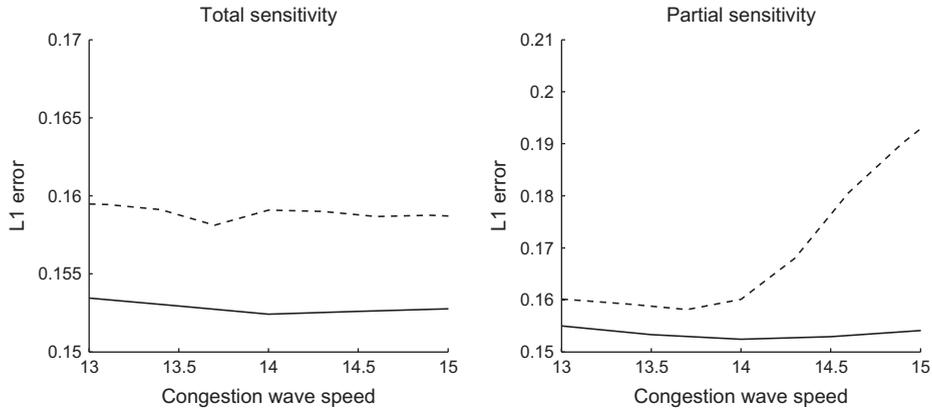


Fig. 15. Sensitivity to congestion wave speed: for the PTM (solid line) and the CTM (dashed line), for the I80 4:00–4:15 dataset, using the density error. The left sub figure denotes the sensitivity of the *global optimum* to the value of the congestion wave speed. The right sub figure denotes the sensitivity of the error metric *along the dimension* corresponding to the congestion wave speed.

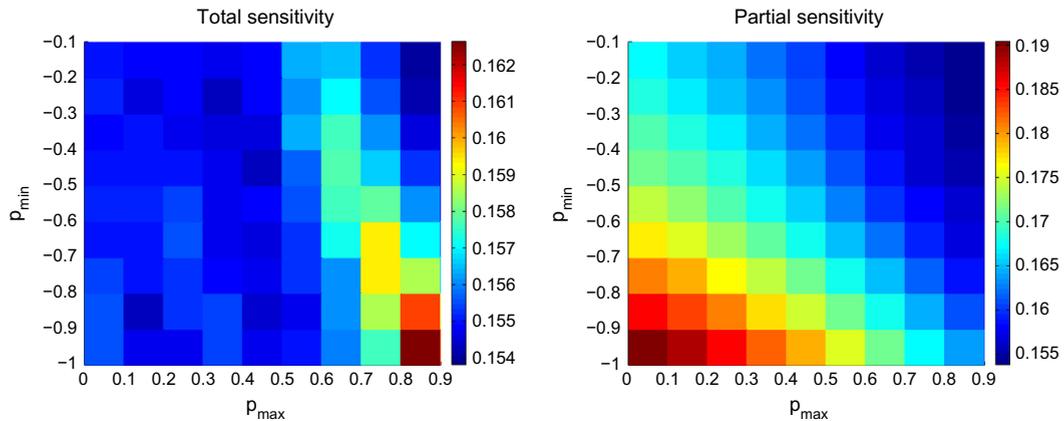


Fig. 16. Sensitivity to perturbation parameters: for the I80 4:00–4:15 dataset, using the density error. The left sub figure denotes the sensitivity of the *global optimum* to the value of the perturbation parameters. The right sub figure denotes the sensitivity of the error metric *on the plane* corresponding to the two perturbation parameters.

gestion phase, and the corresponding choice of the perturbation bounds p_{\min} and p_{\max} , which illustrates the higher robustness of the PTM to calibration error for this parameter.

In the case of the perturbation parameters, it is observed in the right subfigure of Fig. 16 that both minimal and maximal perturbation parameters variations have comparable influence on the error, when other parameters are not re-calibrated. On the other hand, the left subfigure of Fig. 16 illustrates that although the error does not vary widely on the grid proposed, specific coordinated values of the perturbation parameters increase the error. The numerical results illustrate that it can be useful to have a large value for the maximal perturbation, only if the absolute value of the minimal perturbation is not too large. This can be understood as the fact that imposing a specifying ratio between the absolute values of the perturbation parameters de facto restricts the number of points considered for calibrating the other parameters, and can lead to a bias in the estimated shock wave speed. The left subfigure of Fig. 16 also illustrates that having large absolute values for the

Table 6

Model accuracy: L^1 error on the I80 5:00–5:30 testing dataset, for parameters obtained following the training procedure using (a) density, (b) flow, and (c) speed as the error variable on the I80 4:00–4:15 training dataset, between the reconstructed profile and the ground-truth profile for density, flow, and speed, for the CTM and the PTM.

	Density	Flow	Speed
<i>(a) Optimal density parameters</i>			
PTM	0.139	0.167	0.165
CTM	0.146	0.242	0.227
<i>(b) Optimal flow parameters</i>			
PTM	0.141	0.173	0.163
CTM	0.146	0.195	0.191
<i>(c) Optimal speed parameters</i>			
PTM	0.142	0.171	0.165
CTM	0.147	0.190	0.189

two perturbation parameters, resulting in a large congestion phase, yields a large error. This can be understood as a trade-off between having a large congestion phase, allowing to capture more non-stationary points, and having non-stationary waves with speeds comparable to waves between corresponding stationary states.

6.2. Model comparison

In this section we propose a comparative quantitative analysis of the ability of each model to reconstruct the time–space diagram from the knowledge of initial and boundary conditions.

6.2.1. Methodology and quantitative results

In order to assess the performance of the PTM, we propose to compare the model with a classical well-known discrete model from the literature, for which implementation details and calibration procedure are well documented: the CTM.

We assume that boundary conditions are known upstream and downstream, as well as an initial condition. These terms are computed explicitly from the spatio-temporal discretization of vehicle trajectories (see Section 4.1). This corresponds to the computation of the solution to the IBVP defined in Eq. (8) and in practical terms to the case of measurements from sensors (loops, radars, probes, etc.) available at given locations on the freeway as boundary conditions, and a spatial profile assumed to be known at some instant (cameras, satellite, empty road, etc.) as initial condition. The traffic profile on the stretch of road between the sensors, from the time at which the spatial profile is known, is computed by running the models forward in time. For initial or boundary conditions falling outside of the fundamental diagram, we use the projection methods described in Section 5.1.1 for both the CTM and the PTM.⁹

We compare the solution to the initial-boundary value problem for the CTM and the PTM as follows:

- *Training procedure:* we calibrate the model parameters on the I80 4:00–4:15 dataset, for a given error variable (e.g. density, flow or speed).
- *Testing procedure:* we compare the models on the I80 5:00–5:30 dataset on the reconstructed density, flow, and speed fields. We use the set of parameters obtained from the training procedure.

The optimal parameters for the PTM are included in Table 3. The CTM is also calibrated using the procedure described in the previous section and the optimal parameters are included in Table 5.

Remark 7. The values of the free-flow speed v_{\max} obtained after calibration can be significantly lower than classical values for both models (e.g. for the CTM in the case of the I80 4:00–4:15 dataset, with values of 25 mph). It is clear that the cost function does not depend on the free-flow speed for sufficiently high values of the free-flow speed. Since the NGSIM datasets exhibit heavy to moderate congestion with maximal speeds between 20 mph and 30 mph (see Fig. 10), it is expected that optimal values of the free-flow speed fall above this range. Further analysis show that for all datasets, the cost function increases by about 1% per mph for free-flow speed values under the optimal free-flow speed, however the variations of the cost function stay within a 0.5% range for values of the free-flow speed higher than its optimal value. Since the I80 5:00–5:30 dataset exhibits heavier congestion than the I80 4:00–4:15 dataset, this remark legitimates the use of the free-flow speed calibrated on the I80 4:00–4:15 dataset for testing on the I80 5:00–5:30 dataset.

We compute the L^1 error (19) for the three error variables density, flow, and speed, for the three sets of optimal parameters based on the error variable. The results for the CTM and the PTM are presented in Table 6.

⁹ For completeness, different projections have been tested for the PTM, in particular iso-density in the congestion phase. The results obtained are not significantly different from the ones presented in this section and are omitted in the interest of space.

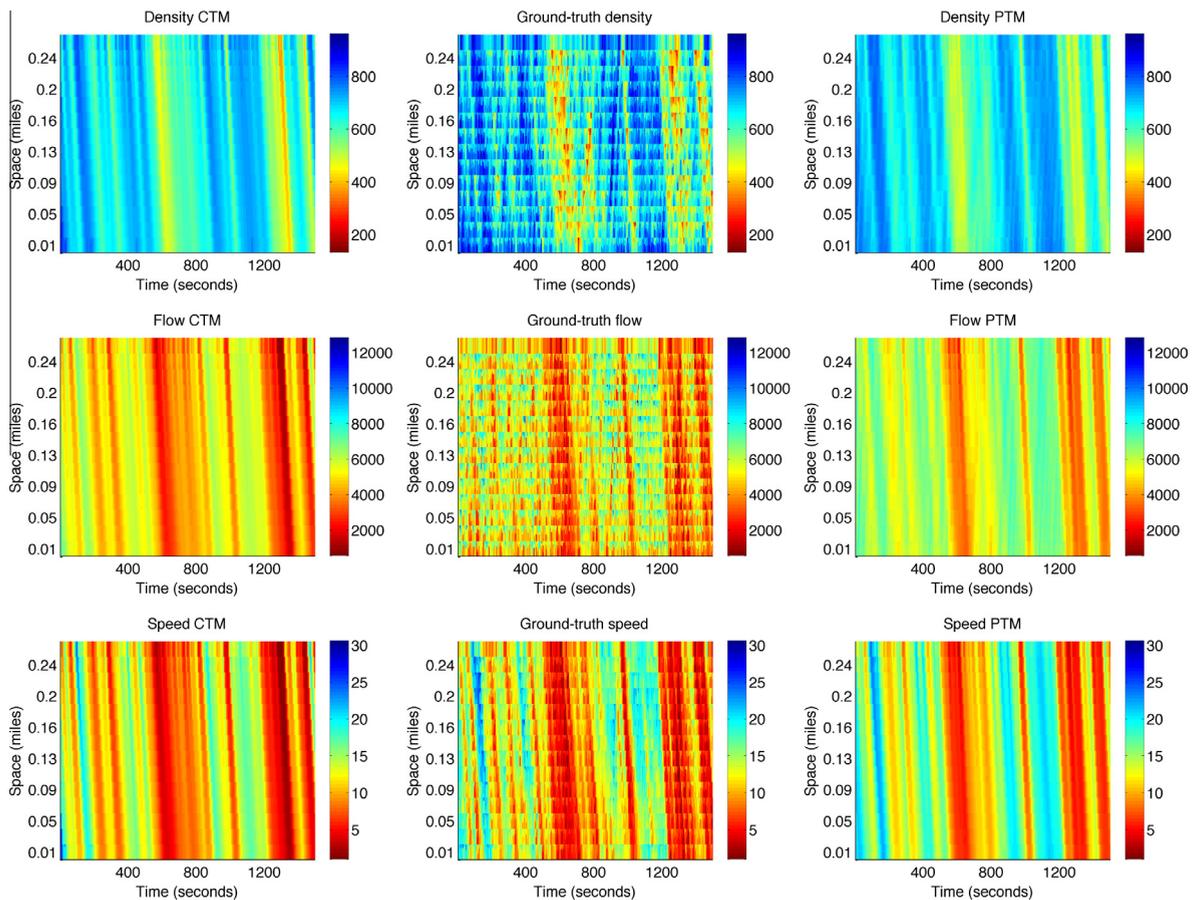


Fig. 17. Time–space diagrams for I80 5–5:30, for density (top row), flow (center row), speed (bottom row), for the CTM (left column), the PTM (right column) and the ground-truth profile (center column).

The results from Table 6 illustrate that the CTM and PTM have relatively similar performances in the estimation of density irrespective of the error variable considered in the parameter calibration. This is consistent with the fact that the PTM is a direct extension of the LWR model in congestion, and specifically of the CTM for the numerical results presented in this section. Thus, even though the PTM has a larger state-space, the corresponding dynamics do not necessarily provide significant added value compared to the CTM when estimating density conditions, regardless of the error variable used for calibration.

However, the PTM shows a clear superiority for the other two relevant traffic variables (flow and speed in that case). For both the variable flow and speed, the error is reduced by 10–25% by the PTM compared to the CTM. Moreover, the error results for the PTM remain fairly constant for all the estimated traffic variables, regardless the parameter calibration approach. This illustrates that the PTM calibration is more robust in the sense that optimal parameters for a given error variable yield a low error for other error variables as well and that errors in the parameter estimation lead to small perturbations in the final results. In the following section, we study the difference in the reconstructed time–space profiles from a qualitative perspective.

6.2.2. Qualitative analysis

The time–space diagrams for the dataset I80 5:00–5:30 with the discretization parameters detailed in Section 4.1 are represented in Fig. 17 for the CTM, the PTM and the ground-truth profile obtained directly by discretization of the vehicles trajectories. One may note that flow and speed are computed a posteriori from the density field in the case of the CTM, whereas in the case of the PTM, speed is obtained from the density k and perturbation p fields, and the flow is obtained classically as a product of density and speed.

The comparison of the reconstructed density fields for the CTM and the PTM shows that the PTM is able to some extent to propagate forward observations gathered at the upstream boundary. The capability of the PTM for hysteresis modeling is illustrated for the density variable in the second part of the time domain for the upstream part of the section, where interactions between forward moving discontinuities and backward moving shock waves yield curved propagation of congestion

waves. However, one might note that the impact of these more complex phenomena is not clear overall since the L^1 error for density is similar for the two models (Table 6, left column).

The comparison of the reconstructed flow fields for the CTM and the PTM shows that the PTM allows the propagation of forward moving discontinuities in flow, and density, within the backward moving congestion phase around the center of the time period. The comparison of the reconstructed speed fields for the CTM and the PTM shows that the PTM captures more extensively the high speed waves at the beginning of the time period. Similarly, high speed values around the center of the time period, and the downstream end of the section, are more accurately captured by the PTM.

7. Discussion

In this article, we analyzed the phase transition model, which consists of an extension of the Lighthill–Whitham–Richards model to capture non-stationary congestion phenomena. We presented the structure of the solution to the continuous model, and an appropriate discretization scheme. We proposed a physically motivated mesoscopic interpretation related to the behavior of heterogeneous drivers, and showed that the solution to the Riemann problem exhibits a simple structure convenient for analytical and algebraic investigations.

We studied the performances and features of the phase transition model on benchmark cases and on the Next Generation SIMulation trajectories datasets, and illustrated that the model is able to capture several complex traffic phenomena not accounted for by the Lighthill–Whitham–Richards model: set-valued fundamental diagram, hysteresis patterns, forward moving discontinuity in congestion, and amplification of disturbances.

The ability of the phase transition model to represent more complex traffic phenomena and account for observations of different traffic quantities jointly or independently is promising for future research. In particular, the model proposes a robust framework for traffic estimation with fixed sensors and probe data, with no requirement to convert measured traffic quantities, which is known to degrade the validity of typical assumptions on the observation error.

Acknowledgments

The authors are grateful to the NGSIM project for open on-line availability of the accurate vehicle trajectories collected by the project. The authors would like to acknowledge a motivating and instructive discussion on the topic with Professor Ludovic Leclercq. The first author was funded during part of this work by the *University of California Transportation Center* as part of the *UCTC Doctoral Dissertation Research Grant 2011*. The second author would like to acknowledge the financial support of “la Caixa” graduate fellowship program.

References

- Ahmed, K., Ben-Akiva, M., Koutsopoulos, H., Mishalani, R., 1996. Models of freeway lane changing and gap acceptance behavior. In: Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Lyon, France, pp. 501–515.
- Ansorge, R., 1990. What does the entropy condition mean in traffic flow theory? *Transportation Research Part B* 24 (2), 133–143.
- Aw, A., Rascle, M., 2000. Resurrection of “second order” models of traffic flow. *SIAM Journal on Applied Mathematics* 60 (3), 916–938.
- Barcelo, J., Casas, J., Ferrer, J., Garcia, D., 1998. Modelling advanced transport telematic applications with microscopic simulators: the case with AIMSUN2. In: Proceedings of the 10th European Simulation Symposium, Nottingham, UK, pp. 362–367.
- Bayen, A., Butler, J., Patire, A., 2011. Mobile Millennium Final Report. Tech. rep., California Center for Innovative Transportation, Institute of Transportation Studies, University of California, Berkeley, Research Report, UCB-ITS-CWP-2011-6.
- Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H., Mishalani, R., 2002. Real-time simulation of traffic demand-supply interactions within DynaMIT. *Transportation and Network Analysis: Current Trends: Miscellanea in Honor of Michael Florian* 63, 19–36.
- Blandin, S., Couque, A., Bayen, A., Work, D., 2012. On sequential data assimilation for scalar macroscopic traffic flow models. *Physica D* 241, 1421–1440.
- Blandin, S., Work, D., Goatin, P., Piccoli, B., Bayen, A., 2010. A class of perturbed cell transmission models to account for traffic variability. In: 89th Transportation Research Board Annual Meeting, Washington, DC, January 10–14.
- Blandin, S., Work, D., Goatin, P., Piccoli, B., Bayen, A., 2011. A general phase transition model for vehicular traffic. *SIAM Journal on Applied Mathematics* 71 (1), 107–127.
- Brauer, F., Nohel, J., 1989. *The Qualitative Theory of Ordinary Differential Equations: An Introduction*. Dover, Mineola, NY.
- Cameron, G., Duncan, G., 1996. PARAMICS—Parallel microscopic simulation of road traffic. *The Journal of Supercomputing* 10 (1), 25–53.
- Cassidy, M., 1998. Bivariate relations in nearly stationary highway traffic. *Transportation Research Part B* 32 (1), 49–59.
- Cassidy, M., Coifman, B., 1997. Relation among average speed, flow, and density and analogous relation between density and occupancy. *Transportation Research Record: Journal of the Transportation Research Board* 1591, 1–6.
- Cassidy, M., Windover, J., 1995. Methodology for assessing dynamics of freeway traffic flow. *Transportation Research Record* 1484, 73–79.
- Castillo, J.D., Benitez, F., 1995. On the functional form of the speed-density relationship. Part II: Empirical investigation. *Transportation Research Part B* 29 (5), 391–406.
- Castillo, J.D., Pintado, P., Benitez, F., 1994. The reaction time of drivers and the stability of traffic flow. *Transportation Research Part B* 28 (1), 35–60.
- Chalons, C., Goatin, P., 2008. Godunov scheme and sampling technique for computing phase transitions in traffic flow modeling. *Interfaces and Free Boundaries* 10 (2), 195–219.
- Chiabaut, N., Buisson, C., Leclercq, L., 2009. Fundamental diagram estimation through passing rate measurements in congestion. *IEEE Transactions on Intelligent Transportation Systems* 10 (2), 355–359.
- Chiabaut, N., Leclercq, L., Buisson, C., 2010. From heterogeneous drivers to macroscopic patterns in congestion. *Transportation Research Part B* 44 (2), 299–308.
- Chow, A., Dadok, V., Dervisoglu, G., Gomes, G., Horowitz, R., Kurzhanski, A., Kwon, J., Lu, X., 2008. Topl: tools for operational planning of transportation networks. In: Proceedings of the 2008 Dynamic Systems and Control Conference, Ann Arbor, Michigan, pp. 20–22.
- Chowdhury, D., Santen, L., Schadschneider, A., 2000. Statistical physics of vehicular traffic and some related systems. *Physics Reports* 329 (4–6), 199–329.
- Colombo, R., 2002. On a 2×2 hyperbolic traffic flow model. *Mathematical and Computer Modelling* 35 (5–6), 683–688.
- Colombo, R., 2003. Hyperbolic phase transitions in traffic flow. *SIAM Journal on Applied Mathematics* 63 (2), 708–721.

- Colombo, R., Goatin, P., Priuli, F., 2007. Global well-posedness of traffic flow models with phase transitions. *Nonlinear Analysis* 66 (11), 2413–2426.
- Cremer, M., Papageorgiou, M., 1981. Parameter identification for a traffic flow model. *Automatica* 17 (6), 837–843.
- Dafermos, C., 2010. *Hyperbolic Conservation Laws in Continuum Physics*. Springer-Verlag, Berlin, Germany.
- Daganzo, C., 1994. The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B* 28 (4), 269–287.
- Daganzo, C., 1995. Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B* 29 (4), 277–286.
- Daganzo, C., Cassidy, M., Bertini, R., 1999. Possible explanations of phase transitions in highway traffic. *Transportation Research Part A* 33 (5), 365–379.
- Edie, L., 1963. Discussion of traffic stream measurements and definitions. In: *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, London, UK, pp. 139–154.
- Edie, L., Foote, R., Herman, R., Rothery, R., 1963. Analysis of single lane traffic flow. *Traffic Engineering* 33 (4), 21–27.
- Fellendorf, M., 1994. VISSIM: A microscopic simulation tool to evaluate actuated signal control including bus priority. In: *64th Institute of Transportation Engineers Annual Meeting*, Dallas, TX.
- Garavello, M., Piccoli, B., 2006. *Traffic Flow on Networks*. American Institute of Mathematical Sciences, Springfield, MO.
- Garavello, M., Piccoli, B., 2009. On fluid-dynamic models for urban traffic. *Network and Heterogeneous Media* 4 (1), 107–126.
- Gazis, D., Herman, R., Potts, R., 1959. Car-following theory of steady-state traffic flow. *Operations Research* 7 (4), 499–505.
- Godunov, S., 1959. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik* 89 (3), 271–306.
- Greenberg, H., 1959. An analysis of traffic flow. *Operations Research* 7 (1), 79–85.
- Greenshields, B., 1935. A study of traffic capacity. *Proceedings of the Highway Research Board* 14 (1), 448–477.
- Helbing, D., 1995. Improved fluid-dynamic model for vehicular traffic. *Physical Review E* 51 (4), 3164–3169.
- Helbing, D., 2001. Traffic and related self-driven many-particle systems. *Reviews of Modern Physics* 73 (4), 1067–1141.
- Hoogendoorn, S., Bovy, P., 2001. State-of-the-art of vehicular traffic flow modelling. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, Special Issue on Road Traffic Modelling and Control* 215, 283–303.
- Jin, W., 2010. A kinematic wave theory of lane-changing traffic flow. *Transportation Research Part B* 44 (8–9), 1001–1021.
- Kerner, B., 1998. Experimental features of self-organization in traffic flow. *Physical Review Letters* 81 (17), 3797–3800.
- Kerner, B., Rehborn, H., 1996. Experimental properties of complexity in traffic flow. *Physical Review E* 53 (5), 4275–4278.
- Kerner, B., Rehborn, H., 1997. Experimental properties of phase transitions in traffic flow. *Physical Review Letters* 79 (20), 4030–4033.
- Laval, J., 2005. Linking synchronized flow and kinematic waves. In: *Proceedings of Traffic and Granular Flow'05*, Berlin, Germany, pp. 521–526.
- Laval, J., 2011. Hysteresis in traffic flow revisited: an improved measurement method. *Transportation Research Part B* 45 (2), 385–391.
- Laval, J., Daganzo, C., 2006. Lane-changing in traffic streams. *Transportation Research Part B* 40 (3), 251–264.
- Laval, J., Leclercq, L., 2010. A mechanism to describe the formation and propagation of stop-and-go waves in congested freeway traffic. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 368 (1928), 4519–4541.
- Lebacque, J.-P., 1997. A finite acceleration scheme for first order macroscopic traffic flow models. In: *Proceedings of the 8th IFAC Symposium on Transportation Systems*, pp. 815–820.
- Lebacque, J.-P., Lesort, J.-B., 1999. Macroscopic traffic flow models: a question of order. In: *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Jerusalem, Israel.
- Lesort, J.-B., Bourrel, E., Henn, V., 2003. Various scales for traffic flow representation: some reflections. In: *Proceedings of Traffic and Granular Flow'03*, Delft, The Netherlands, pp. 125–139.
- Li, T., 2005. Nonlinear dynamics of traffic jams. *Physica D: Nonlinear Phenomena* 207 (1–2), 41–51.
- Lighthill, M., Whitham, G., 1956. On kinematic waves. Part II: A theory of traffic flow on long crowded roads. *Proceedings of the Royal Society of London* 229 (1178), 317–345.
- Lu, X., Skabardonis, A., 2007. Freeway traffic shockwave analysis: exploring NGSIM trajectory data. In: *86th Transportation Research Board Annual Meeting*, Washington, DC, January 21–25.
- Mahmassani, H., Hu, T., Peeta, S., Ziliaskopoulos, A., 1994. Development and Testing of Dynamic Traffic Assignment and Simulation Procedures for ATIS/ATMS Applications. Tech. rep., Technical Report DTFH61-90-R-0074-FG, Center for Transportation Research, University of Texas at Austin.
- Messner, A., Papageorgiou, M., 1990. METANET: a macroscopic simulation program for motorway networks. *Traffic Engineering and Control* 31 (8–9), 466–470.
- Mihaylova, L., Boel, R., Hegyi, A., 2006. An unscented Kalman filter for freeway traffic estimation. In: *Proceedings of the IFAC Symposium on Control in Transportation Systems*, Delft, the Netherlands, pp. 29–31.
- Nagel, K., Nelson, P., 2005. A critical comparison of the kinematic-wave model with observational data. In: *Proceedings of the 16th International Symposium on Transportation and Traffic Theory*, College Park, MD, pp. 145–164.
- Newell, G., 1962. Theories of instability in dense highway traffic. *Journal of the Operations Research Society of Japan* 5, 9–54.
- Newell, G., 1993. A simplified theory of kinematic waves in highway traffic. Part II: Queueing at freeway bottlenecks. *Transportation Research Part B* 27 (4), 289–303.
- Newell, G., 2002. A simplified car-following theory: a lower order model. *Transportation Research Part B* 36 (3), 195–205.
- NGSIM, 2006. Next Generation Simulation. <<http://ngsim-community.org/>>.
- Papageorgiou, M., 1998. Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A* 32 (5), 323–329.
- Papageorgiou, M., Blasseville, J.-M., Hadj-Salem, H., 1990. Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris. Part I: Modelling. *Transportation Research Part A* 24 (5), 345–359.
- Payne, H., 1971. Models of freeway traffic and control. *Mathematical Models of Public Systems* 1 (1), 51–61.
- Prigogine, I., Herman, R., 1971. *Kinetic Theory of Vehicular Traffic*. Elsevier, New-York, NY.
- Richards, P., 1956. Shock waves on the highway. *Operations Research* 4 (1), 42–51.
- Saffman, P., 1992. *Vortex Dynamics*. Cambridge University Press, Cambridge, UK.
- Strub, I., Bayen, A., 2006. Weak formulation of boundary conditions for scalar conservation laws: an application to highway traffic modelling. *International Journal of Robust Nonlinear Control* 16 (16), 733–748.
- Toro, E., 1997. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer, New-York, NY.
- Treiterer, J., Myers, J., 1974. The hysteresis phenomenon in traffic flow. In: *Proceedings of the 6th International Symposium on Transportation and Traffic Theory*, Sydney, Australia.
- Wang, Y., Papageorgiou, M., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. *Transportation Research Part B* 39 (2), 141–167.
- Wang, Y., Papageorgiou, M., Messmer, A., 2007. Real-time freeway traffic state estimation based on extended Kalman filter: a case study. *Transportation Science* 41 (2), 167–181.
- Work, D., Blandin, S., Tossavainen, O.-P., Piccoli, B., Bayen, A., 2010. A traffic model for velocity data assimilation. *Applied Mathematics Research eXpress* 1, 1–35.
- Yeo, H., Skabardonis, A., 2009. Understanding stop-and-go traffic in view of asymmetric traffic theory. In: *Proceedings of the 18th International Symposium on Transportation and Traffic Theory*, Hong-Kong, China, pp. 99–115.
- Zhang, H.M., 1999. A mathematical theory of traffic hysteresis. *Transportation Research Part B* 33 (1), 1–23.
- Zhang, H.M., 2002. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B* 36 (3), 275–290.