Traffic Data

Cross-Sectional Data

Introduction

Cross-sectional data is captured by stationary induction loops, radar, or infrared sensors. The collected information is provided either directly as single vehicle data or aggregated into macroscopic quantities. In this lecture we define the measurable and derived quantities characterizing both data formats, with special attention on the difference between temporal and spatial averages. Traffic density, a spatially defined quantity, cannot be directly measured using cross-sectional detectors, but several estimation methods are presented and discussed. Speed estimation methods are introduced to overcome the inability of single-loop detectors to directly measure vehicle speed.

Microscopic Measurement: Single-Vehicle Data

Cross-sectional data, measured at a fixed cross-section on the road, can be captured by laying pneumatic tubes across the road, by radar, or optically with infrared sensors or light barriers. Most commonly, however, induction loops are installed beneath the road surface. They detect whether a metallic object (such as a car) is above them (Fig.1). A single-loop detector can directly measure (only) the following quantities:

- The time $t^{\alpha} = t_{\alpha}^{0}$ at which the front of vehicle α passes the detector (voltage drop in Fig.1).
- The time t_{α}^{1} at which the rear end of the vehicle passes the detector (voltage rise in Fig.1).

It is impossible for single-loop detectors to measure vehicle speed, but we can obtain an estimate in the case of relatively uniform speed values by assuming an average vehicle length of l. However, this estimate is prone to large errors.

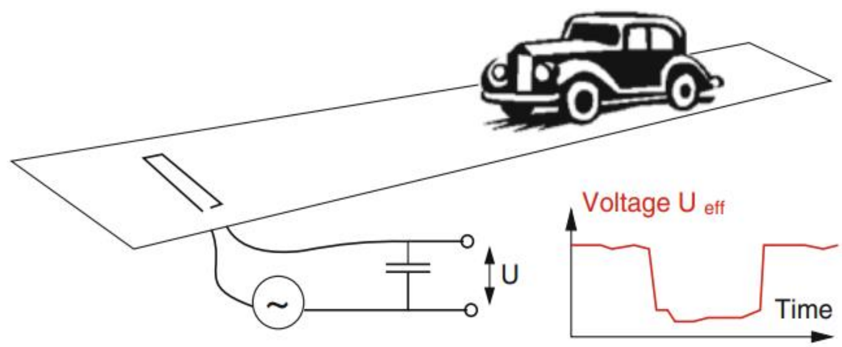

Fig. 1 the induction loop is part of an LC circuit (complemented by an external capacitor and an AC voltage source) tuned to be in resonance if the loop is "unoccupied", yielding a high voltage U_{eff}. The metallic parts of a vehicle will increase the inductance of the loop upon driving over it. This puts the circuit out of tune and decreases the voltage U_{eff}.

Double-loop detectors are composed of two (or more) induction loops separated by a fixed distance, e.g. 1 m. The time difference between passing the first and the second loop yields a direct measurement of the vehicle speed v_α. From these directly measured quantities, we can derive secondary microscopic quantities (cf. Fig.2):

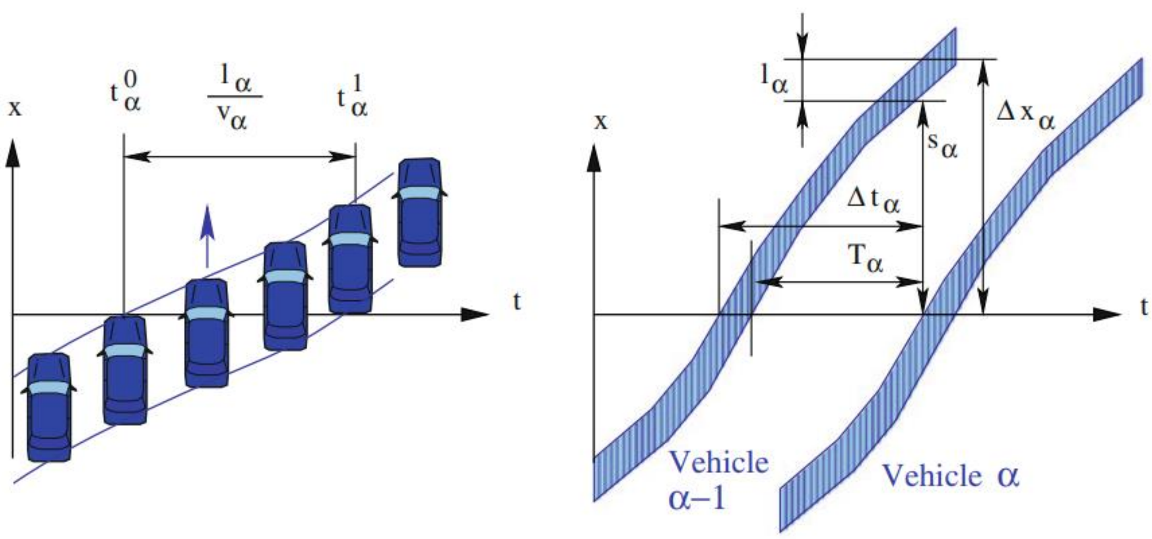

Fig. 2 Single-vehicle data as measured by an induction loop (or any other cross-sectional detector). The shaded area indicates the "detector occupancy" at different times.

• Length of each vehicle α,

$$l_\alpha = v_\alpha (t_\alpha^1 - t_\alpha^0) \quad 1$$

• Vehicle type (motorcycle, car, truck, etc.) by classifying the vehicle length,

• Time headway (sometimes also called simply headway) between the front bumpers of successive vehicles (the smaller index $\alpha - 1$ denotes the leading vehicle),

$$\Delta t_\alpha = t_\alpha^0 - t_{\alpha-1}^1 \quad 2$$

• Time gap between the rear and front bumpers

$$T_\alpha = t_\alpha^0 - t_{\alpha-1}^1 = \Delta t_\alpha - \frac{v_{\alpha-1}}{l_{\alpha-1}} \quad 3$$

• Distance headway

$$\mathrm{d}_\alpha = \mathrm{v}_{\alpha-1} \Delta \mathrm{t}_\alpha \quad 4$$

• and distance gap between the rear and front bumpers (sometimes denoted simply as gap)

$$\mathrm{s}_\alpha = \mathrm{d}_\alpha - l_{\alpha-1} \quad 5$$

All spatial quantities (vehicle length, distances) are only exact if the speed is constant during the measurement, which is a reasonable assumption.

Aggregated Data

Most detectors aggregate the microscopic single-vehicle data by averaging over fixed time intervals Δt and transmit only the macroscopic data (aggregated data) to the traffic control center. This saves both bandwidths in the transmission and disk space when archiving the data, but of course, all the microscopic information is lost. Time intervals vary between 20 s and 5 min, the most common being Δt = 60 s. Averages over a fixed number of vehicles (e.g. ΔN = 50 veh) are rarely used, even though they are statistically more meaningful. One or more of the following quantities are sent to the traffic control center:

Traffic flow. The traffic flow is defined as the number of vehicles ΔN passing the cross-section at location x within a time interval Δt:

$$Q(x,t) = \frac{\Delta N}{\Delta t} \qquad 6$$

It is usually given in units of vehicles per hour (veh/h) or vehicles per minute. In terms of the microscopic quantities, the traffic flow Q can be considered as the inverse of the time mean of the headways, $Q = 1/\langle \Delta t_\alpha \rangle$

The notation $\langle . \rangle$ is used for the arithmetic average in the context of measurements and for the expected value in the context of statistical considerations.

Sometimes, the inverse of the headway is called microscopic flow,

$$q_\alpha = \frac{1}{\Delta t_\alpha} \qquad 7$$

And the scatter plot of q_α versus v_α the microscopic flow-density diagram (Notice that the term microscopic fundamental diagram generally denotes the gap as a function of the speed for steady-state traffic flow as given by microscopic models).

We emphasize that the traffic flow Q can be considered as the harmonic mean of the microscopic flow:

$$Q = \frac{1}{\langle \Delta t_\alpha \rangle} = \frac{1}{\langle {}^{1}/_{q_\alpha} \rangle} \qquad 8$$

Generally, the harmonic mean of a series of values x_α is defined as the inverse of the arithmetic mean of the inverse, $X_H = \frac{1}{X_\alpha}$.

Occupancy. The dimensionless occupancy is the fraction of the aggregation interval during which the cross-section is occupied by a vehicle:

$$O(x,t) = \frac{1}{\Delta t} \sum_{\alpha=\alpha_0}^{\alpha_0+\Delta N-1} (t_\alpha^1 - t_\alpha^0) \qquad 9$$

Arithmetic means speed. The arithmetic mean speed is the average speed of the ΔN vehicles passing the cross-section during the aggregation interval:

$$V(x,t) = \langle v_\alpha \rangle = \frac{1}{\Delta N} \sum_{\alpha=\alpha_0}^{\alpha_0+\Delta N-1} v_\alpha \qquad 10$$

We use V for the macroscopic speed to distinguish it from the (microscopic) speed v_α of single vehicles. To emphasize that the speed is measured at a fixed location for a time interval, V is sometimes called time mean speed.

Harmonic mean speed. The harmonic mean speed is defined as:

$$V_H(x,t) = \frac{1}{\langle \frac{1}{v_\alpha} \rangle} = \frac{\Delta N}{\sum_{\alpha=\alpha_0}^{\alpha_0+\Delta N-1} \frac{1}{v_\alpha}} \qquad 11$$

When neglecting accelerations, V_H corresponds approximately to the (spatial) average of the speed at a fixed time instant. Therefore, V_H is sometimes called (not completely correctly) space mean speed. One can show that always $V_H \leq V$ where the equal sign only holds if all speeds are identical. The harmonic mean speed and the following two quantities are rarely available although they would be useful for a less biased traffic density estimate.

Arithmetic time means microscopic flow. The arithmetic time mean of microscopic flow is defined by:

$$Q^*(x,t) = \langle q_\alpha \rangle = \langle \frac{1}{\Delta t_\alpha} \rangle = \frac{1}{\Delta N} \sum_{\alpha=\alpha_0}^{\alpha_0+\Delta N-1} \frac{1}{\Delta t_\alpha} \qquad 12$$

This quantity is very useful in estimating the density when no microscopic data are available.

Speed variance. The speed variance:

$$Var(v) = \sigma_v^2(x,t) = \langle (v_\alpha - \langle v_\alpha \rangle)^2 \rangle = \langle v_\alpha^2 \rangle - \langle v_\alpha \rangle^2 \qquad 13$$

is a measure of the spread of the speed values within the aggregation interval. The spread is given by the standard deviation σ_v, the square root of the variance. The dimensionless coefficient of variation σ_v/V quantifies the relative spread of the speed values.
Considering the mean speed in a highly heterogeneous traffic flow, what is the advantage of averaging over a fixed number of vehicles instead of over fixed time intervals?

Estimating Spatial Quantities from Cross-Sectional Data

While the macroscopic quantities flow Q, occupancy O, and (in the case of double loop detectors) the arithmetic mean speed V are measured directly, other important quantities can only be estimated by making some assumptions. The traffic density is defined as a spatial average at a fixed time (the number of vehicles on a given road segment) but cross-sectional detectors can only measure temporal averages at a fixed location (the cross-section). Contrary to flow and density, the macroscopic speed can be defined both as a temporal and a spatial average. However, these two definitions are not equivalent.

Traffic Density

The traffic density $\rho(x, t)$ can be estimated using the hydrodynamic relation:

$$\rho(x,t) = \frac{Q(x,t)}{V(x,t)} = \frac{Flow}{Speed} \qquad 14$$

However, this equation implicitly assumes that the speed V is a spatial average (because the density is defined as a spatial quantity). Using the temporal averages obtained from cross-sectional detectors induces systematic errors: Faster vehicles are "seen" more frequently by detectors than slower vehicles, yielding a bias towards larger speed values.

Figure 3 shows a two-lane road where vehicles on the left lane drive twice as fast as vehicles on the right lane. The flow is equal on both lanes, thus the detector "sees" the same number of vehicles during the aggregation interval and reports the temporal mean speed $\langle v_\alpha \rangle$=108 km/hr. However, space means speed is:

(2/ 3) 72 km/h + (1/ 3) 144 km/h = 96 km/hr 15

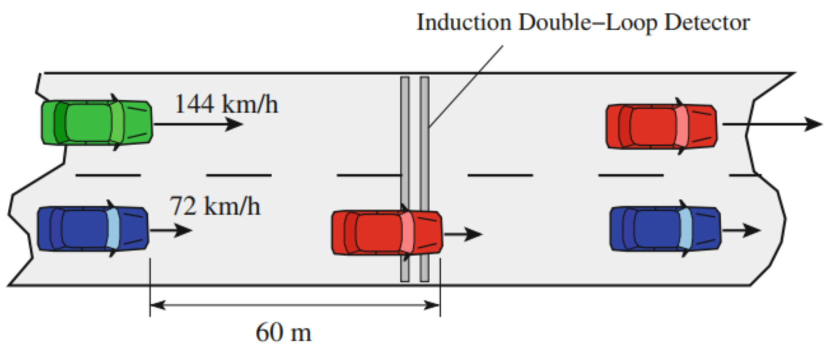

Fig. 3 Vehicles on the left lane drive twice as fast but at the same time headway.

Thus, the density as obtained by the hydrodynamic relation (14) underestimates the real density by a factor of 8/9.

We can obtain a better estimate for the density from its definition of "vehicles per distance", which can be expressed in terms of microscopic quantities as the inverse of the space mean of the distance headways,

$$\rho(x,t) = \frac{1}{\langle d_\alpha \rangle} = \frac{\Delta N}{\sum_\alpha d_\alpha} \quad 15$$

Similarly, the flow ("vehicles per time") can be written as the inverse of the time mean of the headways. For a given fixed time interval:

$$\Delta t = \sum_{\alpha=\alpha_0}^{\alpha_0+\Delta N-1} \Delta t_\alpha = \Delta N \langle \Delta t_\alpha \rangle$$

the flow is given by:

$$Q = \frac{\Delta N}{\Delta t} = \frac{1}{\langle \Delta t_\alpha \rangle} \quad 16$$

In the following section, we discuss two different ways of expressing the density in terms of the measurable quantities Δt_α and v_α.

Derivation from the Expected Value of Traffic Density

Inserting Eq. (4) into the definition of the expected density (15) yields:

$$\frac{1}{\rho} = \langle d_\alpha \rangle = \langle v_{\alpha-1}\Delta t_\alpha \rangle$$

$$\approx \langle v_\alpha \Delta t_\alpha \rangle$$

$$= \langle v_\alpha \rangle \langle \Delta t_\alpha \rangle + Cov(v_\alpha, \Delta t_\alpha)$$

$$= \frac{V}{Q} + Cov(v_\alpha, \Delta t_\alpha)$$

and solving for ρ gives us:

$$\rho = \frac{Q}{V}\left(\frac{1}{1+\frac{Q}{V}Cov(v_\alpha, \Delta t_\alpha)}\right) \tag{17}$$

Here Cov (·, ·) denotes the covariance, defined for two random variables x and y as:

$$Cov(x, y) = \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle = \langle xy \rangle - \langle x \rangle \langle y \rangle \tag{18}$$

The covariance is positive if both variables are positively correlated, i.e. larger values of x tend to be accompanied by proportionally larger values of y. The significance of such a linear relationship is quantified by the correlation coefficient.

$$r_{x,y} = \frac{Cov(x,y)}{\sigma_x \sigma_y} \tag{19}$$

For uncorrelated x and y (that is, the variables have no linear relationship) the coefficient is 0. Its value is bounded between−1 (x and y are perfectly anti-correlated, $x \propto -y$) and +1 (x and y are perfectly correlated, $x \propto y$). The correlation coefficient allows us to rewrite Eq. (17) as Wardrop's equation:

$$\rho = \frac{Q}{V}\left(\frac{1}{1+\frac{\sigma_V}{V}\frac{Q}{\sigma_Q} r_{v_\alpha,\Delta t_\alpha}}\right) \tag{20}$$

Thus, the real density equals the (widely used) estimate "flow divided by arithmetic mean speed" multiplied by a correction factor that captures the correlation between speed and headway, $r_{v_\alpha,\Delta t_\alpha}$, as well as the (relative) variance of vehicle speed and flow, σ_V /V and σ_Q/Q. In free traffic, $r_{v_\alpha,\Delta t_\alpha}$ is near zero since every driver is able to choose his or her speed independently. In congested traffic, however, the headway Δt_α usually increases with decreasing speed and tends to infinity as the speed approaches zero. Therefore, $r_{v_\alpha,\Delta t_\alpha}$ is negative in this case and the correction factor is greater than 1. Thus, the relation Q/V systematically underestimates the real density in congested.

Derivation from the Expected Value of Traffic Flow

A different approach to derive the density from measurable quantities combines the expected value of the flow (16) with Eq. (4):

$$\frac{1}{Q} = \langle \Delta t_\alpha \rangle = \langle \frac{d_\alpha}{v_{\alpha-1}} \rangle \qquad 21$$

$$\approx \langle \frac{d_\alpha}{v_\alpha} \rangle = \langle d_\alpha \rangle \langle \frac{1}{v_\alpha} \rangle + Cov(d_\alpha, \frac{1}{v_\alpha}) \qquad 22$$

$$= \frac{1}{\rho V_H} + Cov\left(d_\alpha, \frac{1}{v_\alpha}\right) \qquad 23$$

Again solving for ρ we obtain

$$\rho = \frac{Q}{V_H} \left(\frac{1}{1 - Q Cov(d_\alpha, \frac{1}{v_\alpha})}\right) \qquad 24$$

Where V_H is the harmonic mean speed (11) that gives stronger weight to small speed values. Since the distance headway d_α usually increases with v_α (and decreases with $1/v_\alpha$), Cov (d_α, $1/v_\alpha$) is negative and the correction factor is smaller than 1. Thus, Q/V_H generally overestimates the real density.

Discussion of the Two Approximations

In practice, the covariance's in Eqs. (20) and (24) are usually assumed to be zero and

$$\rho^{(1)} = \frac{Q}{V} \; or \; \rho^{(2)} = \frac{Q}{V_H} \qquad 25$$

is used to calculate the density (both relations can be applied to multi-lane traffic as well). The following statements help in assessing the errors of the two estimates:

1. If all vehicle speeds v_α are the same, then $V = V_H$ and thus $\rho = \rho^{(1)} = \rho^{(2)}$.

2. If all headways Δt_α are the same, then $Cov(v_\alpha, \Delta t_\alpha) = 0$, and thus $\rho = \rho^{(1)} = Q/V$ holds exactly. Otherwise, $\rho^{(1)}$ most likely underestimates the real density as $Cov(v_\alpha, \Delta t_\alpha)$ is usually negative.

3. If all distance headways d_α are the same, then Cov (d_α, 1/ v_α) = 0, and thus $\rho = \rho^{(2)} = Q/V_H$ holds exactly. Otherwise, $\rho^{(2)}$ most likely overestimates the real density since Cov (dα, 1/ v_α) is usually negative as well.

Space Mean Speed

The space mean speed (instantaneous mean) (V(t)) is the arithmetic mean of the speed of all vehicles within a given road segment at time t (in Fig. 4 this is a segment of length L around the detector),

$$\langle V(t)\rangle = \frac{1}{n(t)}\sum_{\alpha=1}^{n(t)} v_\alpha(t) \qquad 26$$

In general (that is, with multiple lanes and arbitrary speeds and accelerations), aggregated detector data is unsuitable for determining the space mean speed because the number and identities of vehicles used in the average in Eq. (26) change within the aggregation interval Δt. Also, it is possible that n(t) = 0.

We get a more suitable definition by averaging the instantaneous mean over the aggregation interval Δt. Furthermore, we choose the reference length L small enough so that no vehicles are on the reference road segment at time t or t + Δt and the vehicle speed does not change significantly during the time needed for passing the segment, $\tau_\alpha \approx$ L/v_α. Averaging Eq. (26) over time gives us:

$$\langle V\rangle = \frac{\int_t^{t+\Delta t} n(t)\langle V(\acute{t})\rangle d\acute{t}}{\int_t^{t+\Delta t} n(t)\acute{}d\acute{t}} = \frac{\sum_\alpha \int_{t_\alpha}^{t+t_\alpha} v_\alpha}{\sum_\alpha t_\alpha}$$

$$\approx \frac{\sum_\alpha \tau_\alpha v_\alpha}{\sum_\alpha \tau_\alpha} \approx \frac{nL}{\sum_\alpha L/v_\alpha}$$

$$= \frac{n}{\sum_{\alpha=1}^{n}\frac{1}{v_\alpha}}$$

Here n denotes the total number of vehicles that have passed the detector within the interval Δt (not to be confused with n(t), the number of vehicles on the referenced road segment). The speed values v_α are those obtained from the detector (i.e., measured at the same location but at different times, as opposed to measured simultaneously at different locations).

Thus, the time-averaged (over an aggregation interval) and space-averaged (over a road segment) speed is given by the harmonic mean,

$$\langle V\rangle = V_H \qquad 27$$

Although not exact, the harmonic mean V_H of temporal speed data obtained at a fixed location (stationary detectors) is often equated with the instantaneous mean, also called space mean speed in the literature:

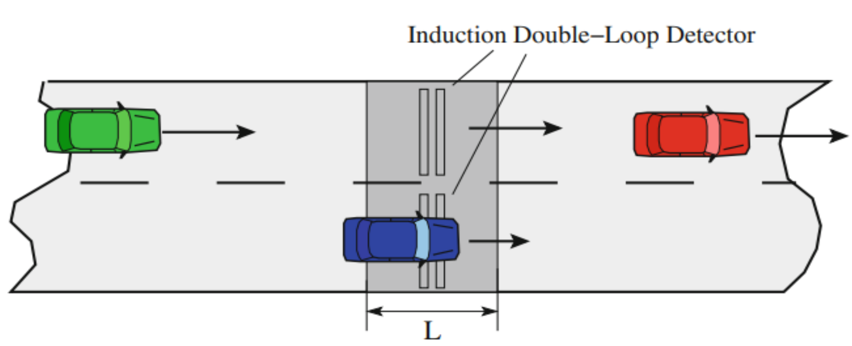

Fig. 4 Derivation of the space mean speed (27).

Determining Speed from Single-Loop Detectors

Single-loop detectors only measure the entry and exit times $t_\alpha^0$and t_α^1 of each vehicle α. If the vehicle length $l_\alpha$was known, we could obtain the speed from $v_\alpha = l_\alpha/(t_\alpha^1 - t_\alpha^0)$. However, single-loop detectors cannot measure vehicle length. Yet we can assume an average vehicle length $\langle l_\alpha \rangle$and use the definition of the occupancy (9) to derive an estimate of the average speed:

$$O = \frac{1}{\Delta t}\sum_\alpha (t_\alpha^1 - t_\alpha^0)$$

$$= \frac{1}{\Delta t}\sum_\alpha \frac{l_\alpha}{v_\alpha}$$

$$= \frac{n}{\Delta t}\left[\langle l_\alpha\rangle \langle\frac{1}{v_\alpha}\rangle + Cov\left(l_\alpha, \frac{1}{v_\alpha}\right)\right]$$

$$= Q\left[\langle l_\alpha\rangle \langle\frac{1}{v_\alpha}\rangle + Cov\left(l_\alpha, \frac{1}{v_\alpha}\right)\right]$$

Solving for $V_H = 1/\langle {}^1/_{v_\alpha}\rangle$we get:

$$V_H = \frac{Q\langle l_\alpha\rangle}{O\left[1 - \frac{Q}{O}Cov(l_\alpha, 1/v_\alpha)\right]} \quad 28$$

For large densities the covariance Cov(lα, $1/v_\alpha$) is nearly zero because all vehicles drive with approximately the same speed. Thus, the estimate for V_H simplifies to $Q\langle l_\alpha\rangle/O$. In free traffic, however, longer vehicles (trucks) usually drive more slowly than shorter vehicles (cars), thus $Cov(l_\alpha, 1/v_\alpha) > 0$.

In this case, $Q\langle l_\alpha\rangle/O$ systematically underestimates the harmonic mean of the speed values. However, since the harmonic mean is always less than the arithmetic means for any data with finite variance, $Q\langle l_\alpha\rangle/O$ may be a good estimate for the arithmetic mean. If all vehicle lengths are equal, the simple relation between occupancy and harmonic mean speed is exact (for arbitrary speed values). For all these cases the traffic density can be easily estimated as well, yielding:

$$V_H = \frac{Q}{\acute{\rho}} \text{ with } \acute{\rho} = \frac{O}{\langle l_\alpha\rangle}$$

To apply these equations only the average vehicle length $\langle l_\alpha\rangle$ must be known.

Example 1:

Consider the following 30 s excerpt from single-vehicle data of a cross-sectional detector:

Time (in s)	Speed (in m/s)	Lane (1 = right, 2 = left)	Vehicle length (in m)
2	26	1	5
7	24	1	12
7	32	2	4
10	32	2	5
12	29	1	4
18	28	1	4
20	34	2	5
21	22	1	15
25	26	1	3
29	38	2	5

1. Aggregate the data and calculate the macroscopic traffic flow and speed (arithmetic mean), separately for both lanes.

2. Calculate the traffic density in each lane assuming that speed and time headway of two succeeding vehicles are uncorrelated (which is realistic for free traffic).

3. Determine the flow, speed, and density of both lanes combined.

4. What percentage of the vehicles on the right lane (and in total) are trucks?

Solution:

1. Flow and speed: With an aggregation interval $\Delta t = 30$ s and $n_1 = 6$, $n_2 = 4$ measured vehicles on lanes 1 and 2, respectively, the flow and time mean speed on the two lanes are:

   $Q_1 = \frac{n_1}{\Delta t}$ = 0.2 vehicles/sec. = 720 vehicles/hr
   $Q_2 = \frac{n_2}{\Delta t}$ = 0.133 vehicles/sec. = 480 vehicles/hr

   $V_1 = \frac{1}{n_1} \sum_\alpha v_{1\alpha}$ = 25.8 m/sec.

   $V_2 = \frac{1}{n_2} \sum_\alpha v_{2\alpha}$ = 34.0 m/sec.

2. Density: When assuming zero correlations between speeds and time headways, the covariance $\mathrm{Cov}(v_\alpha, \Delta t_\alpha) = 0$.
   With Eq. (20), this means that calculating the true (spatial) densities by Q/V using the arithmetic (time) mean speed gives no bias:
   $\rho_1 = Q_1/V_1$ = 7.74 vehicles/km, $\rho_2 = Q_2/V_2$ = 3.92 vehicles/km.

3. Both lanes combined: Density and flow are extensive quantities increasing with the number of vehicles. Therefore, building the total quantities by simple summation over the lanes makes sense:
   $\rho_{tot} = \rho_1 + \rho_2 = 11.66$ vehicles/km, $Q_{tot} = Q_1 + Q_2 = 1{,}200$ vehicles/h.

Since speed is an intensive quantity (it does not increase with the vehicle number), summation over lanes makes no sense. Instead, we define the effective aggregated speed by requiring the hydrodynamic relation to be valid for total flow and total density as well:

$$V = \frac{Q_{tot}}{\rho_{tot}} = \frac{\rho_1 V_1}{\rho_2 V_2} = \frac{Q_{tot}}{\frac{Q_1}{V_1} + \frac{Q_2}{V_2}} = 28.5 \text{ m/sec.} = 102.9 \text{ km/hr}$$

4. Fraction of trucks: Two out of six (33 %) are in the right lane, none in the left, and two out of ten (20 %) total. Notice again that the given percentages are the fraction of trucks passing a fixed location (time mean). In the same situation, we expect the fraction of trucks observed by a "snapshot" of a road section at a fixed time (space mean). To be higher, at least if trucks are generally slower than cars