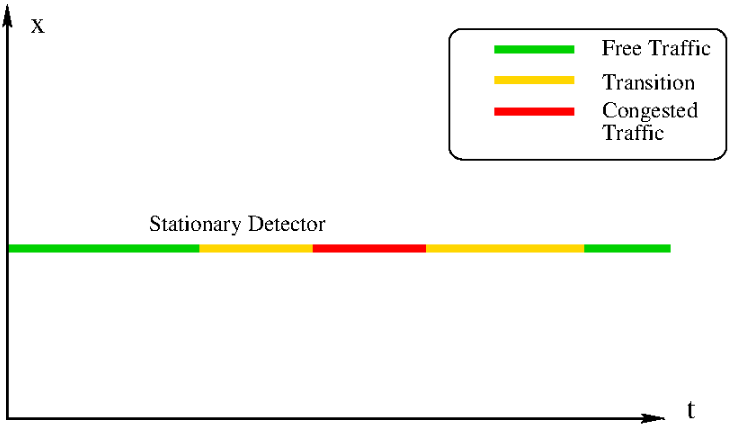
Lecture 4: Data Fusion

- 4.1. Data Fusion: Problem Statement
- 4.2. Data Fusion “by Hand”
- 4.3. Reliability Weighting
- 4.4. Adaptive Smoothing Method

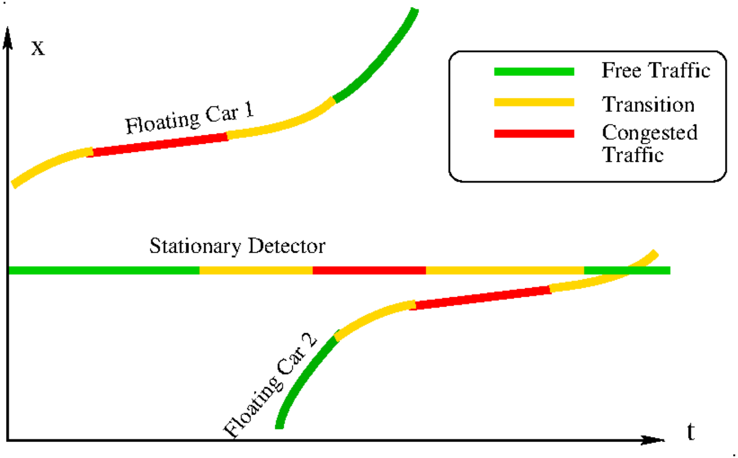
4.1. Data Fusion: Problem Statement



Traffic flow data may come from several sources:

- stationary detectors

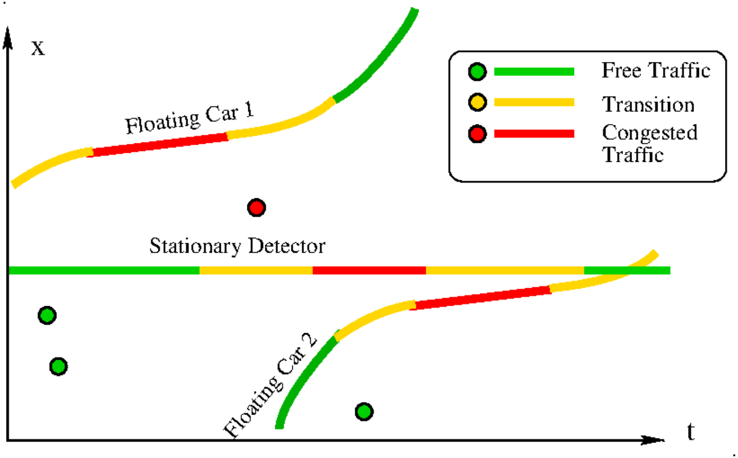
4.1. Data Fusion: Problem Statement



Traffic flow data may come from several sources:

- stationary detectors
- floating cars

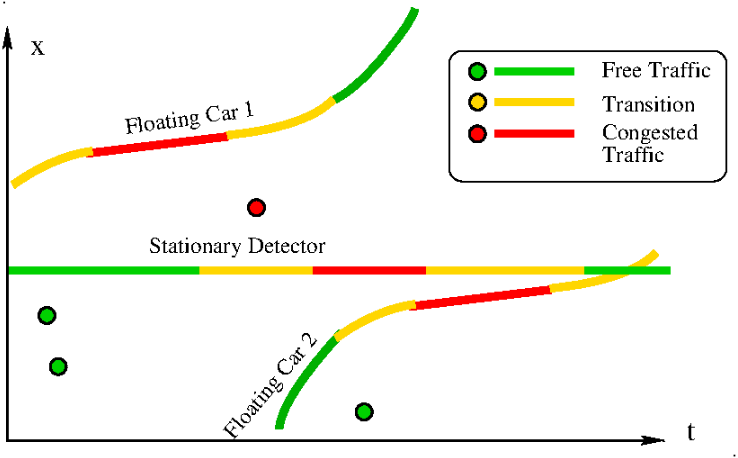
4.1. Data Fusion: Problem Statement



Traffic flow data may come from several sources:

- stationary detectors
- floating cars
- point observations by drivers ("jam reporter") or authorities

4.1. Data Fusion: Problem Statement



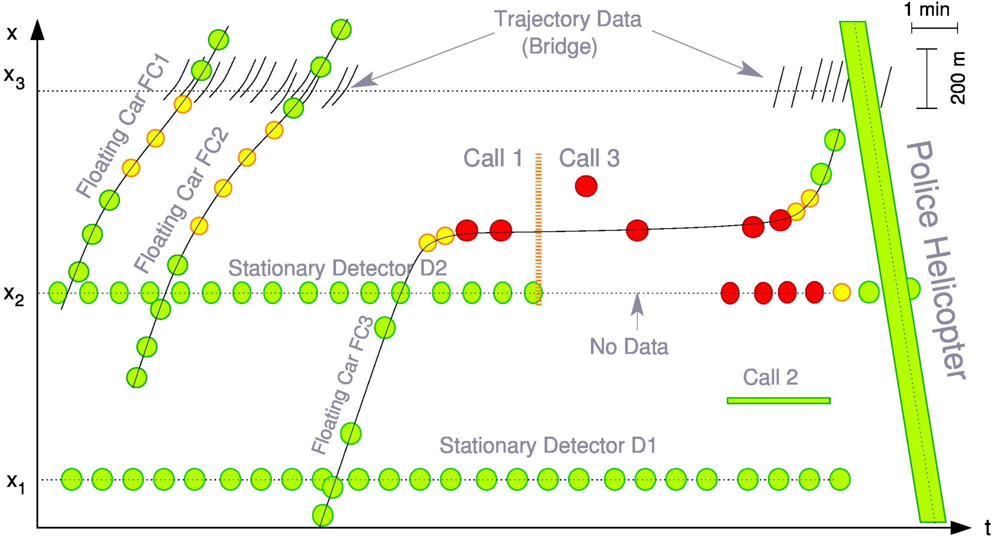
Traffic flow data may come from several sources:

- stationary detectors
- floating cars
- point observations by drivers ("jam reporter") or authorities

They may also of different reliability and even contradictory **(spot such an inconsistency above!)**.

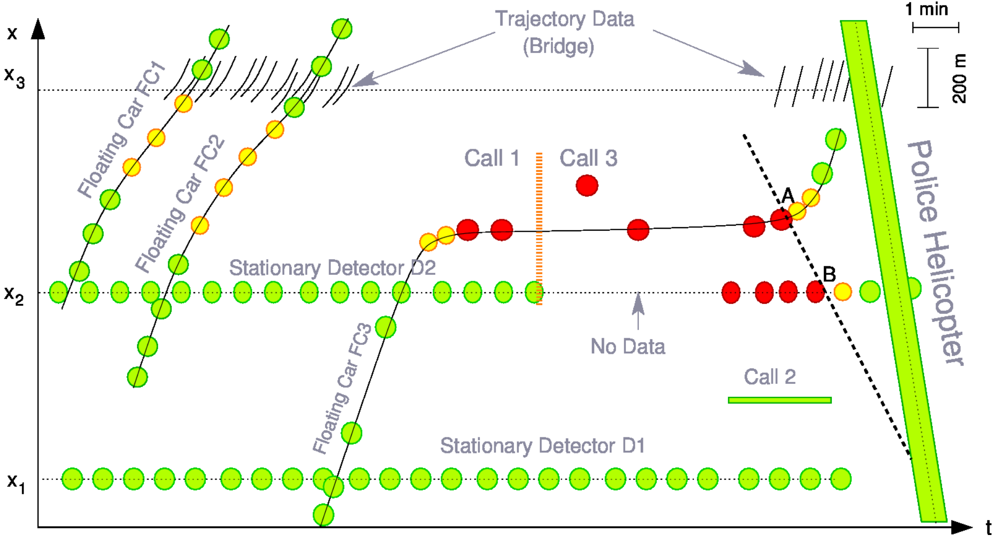
→ back reliability weighting

4.2. Data Fusion "by Hand"



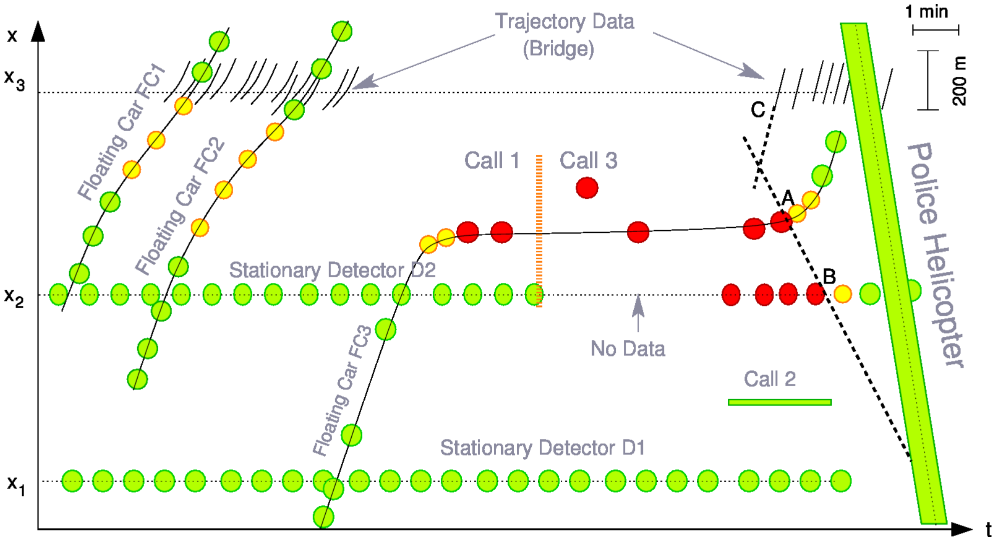
An accident happened: When and where?

Solution



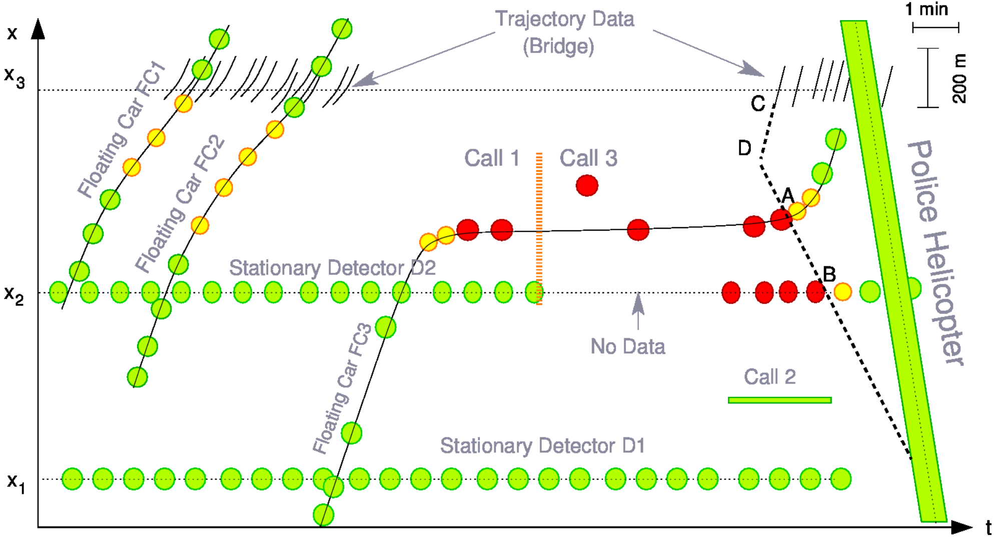
A moving downstream jam front always has the wave speed $w \Rightarrow$ straight line connecting A and B

Solution



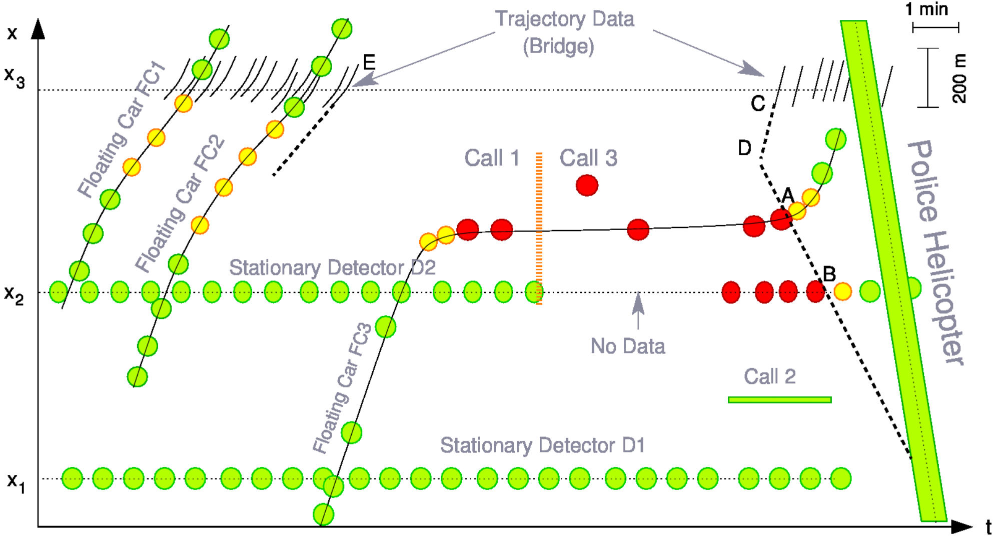
Extrapolate to the past the trajectory C of the first vehicle observed from the bridge after the blocking

Solution



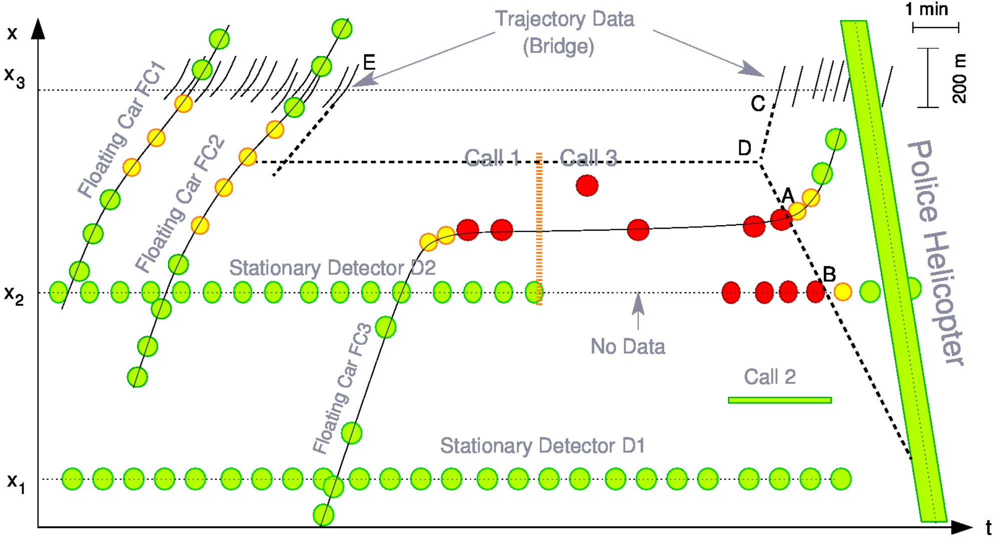
The intersection of both lines gives the location x_{crash} of the accident and the time the road block is lifted (D)

Solution



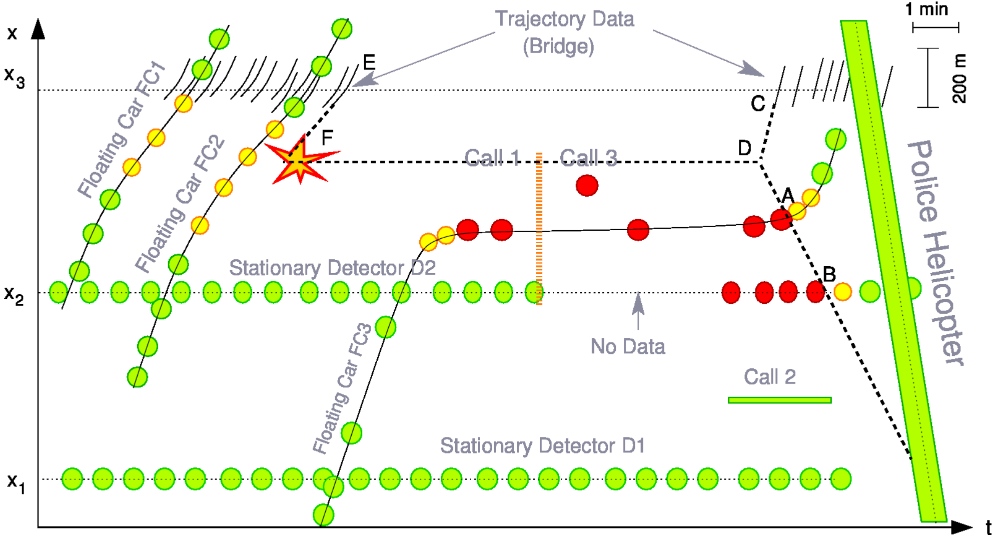
Extrapolate to the past the trajectory E of the last vehicle that made it through the future accident location

Solution



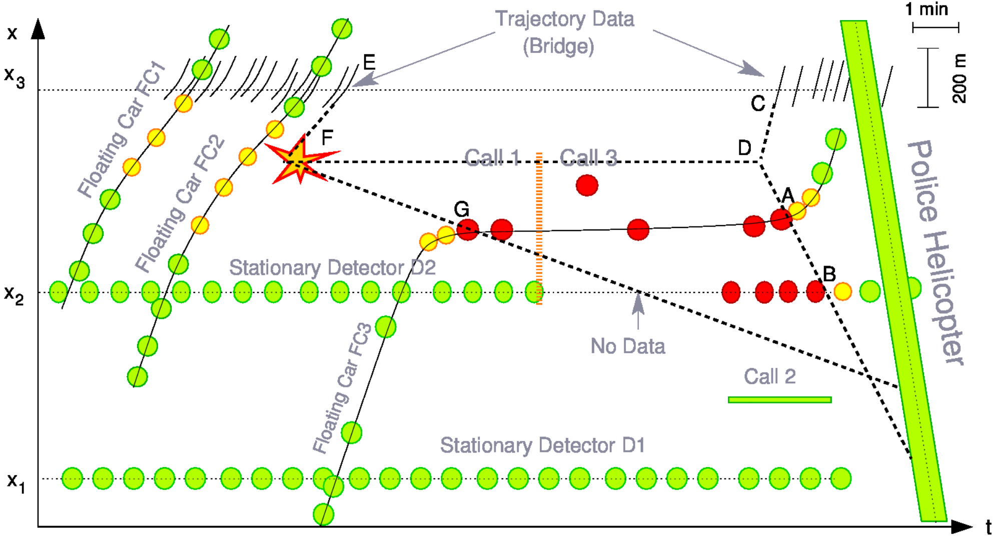
Since the accident is not moving, the intersection of trajectory E with the line $x = x_{\text{crash}}$ gives the time of the accident

Solution



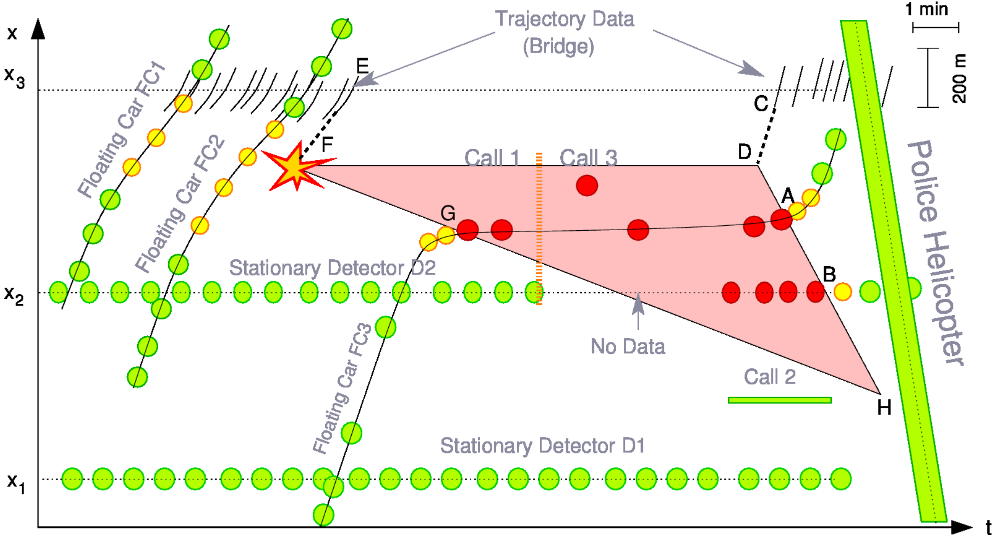
The accident happened at the spatiotemporal point F

Solution



Assuming a constant upstream traffic demand the upstream jam front propagates at a constant velocity ⇒ line FG

Solution



The location of the accident, the time of lifting the block, and the spatiotemporal dynamics of the jam is revealed!

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma^2_{\hat{Y}}(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma_{\hat{Y}}^2(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma_{\hat{Y}}^2(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

  of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma_{\hat{Y}}^2(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma^2_{\hat{Y}}(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \overset{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma^2_{\hat{Y}}(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

4.3 Reliability weighting

Not all data sources are equally reliable. And may contradict each other. How to weight them optimally, i.e., find optimal weights for $\hat{Y} = \sum_m r_m Y_m$?

- Assume M independent and unbiased measurements $Y_m, m = 1, ..., M$ with error variances σ_m^2. From the unbiasedness and the general variance rule $V(aY_1 + bY_2) = a^2 V(Y_1) + b^2 V(Y_2)$
- $\Rightarrow$ Optimization problem: find the reliability weightings r_i such that the variance

$$\sigma_{\hat{Y}}^2(\boldsymbol{r}) = \sum_m r_m^2 \sigma_m^2 \stackrel{!}{=} \min_{\boldsymbol{r}} \quad \Big| \quad \sum_m r_m = 1$$

of the weighted average $\hat{Y} = \sum_m r_m y_m$ is minimized subject to the normalisation condition.

? Why we need independence when using this formula? Is it practically fulfilled?

Otherwise, the variance formula will contain additional covariance terms. Independency generally fulfilled.

? Why we need the restraint $\sum_m r_m = 1$?

Otherwise, the estimator is no longer unbiased: $E(\hat{Y}) = \sum_m r_m E(Y) \neq E(Y)$

Solving the restrained minimization problem

The method of **Lagrange multipliers** does the magic! With a single restraint (a generalisation is straightforward, see Tutorial 04), do the following:

1. Formulate the restraint as an "=0" equation: $g(\boldsymbol{r}) = \sum_m r_m - 1 = 0$
2. Define the **Lagrange function** by adding to the function f to be minimized the restraint multiplied by a *Lagrange multiplier* λ:

$$L(\boldsymbol{r}) = f(\boldsymbol{r}) - \lambda g(\boldsymbol{r}) = \sum_{m'} r_{m'}^2 \sigma_{m'}^2 - \lambda \left(\sum_{m'} r'_m - 1 \right)$$

3. Minimize L unconditionally:

$$\frac{\partial L}{\partial r_m} = 2 r_m \sigma_m^2 - \lambda \stackrel{!}{=} 0$$

4. Calculate λ by inserting the result into the restraint. Here, we have $r_m = \lambda / (2\sigma_m^2) \Rightarrow$

$$r_m = \frac{\sigma_m^{-2}}{\sum_{m'} \sigma_{m'}^{-2}}$$

Solving the restrained minimization problem

The method of **Lagrange multipliers** does the magic! With a single restraint (a generalisation is straightforward, see Tutorial 04), do the following:

1. Formulate the restraint as an "=0" equation: $g(\boldsymbol{r}) = \sum_m r_m - 1 = 0$

2. Define the **Lagrange function** by adding to the function f to be minimized the restraint multiplied by a *Lagrange multiplier* λ:

$$L(\boldsymbol{r}) = f(\boldsymbol{r}) - \lambda g(\boldsymbol{r}) = \sum_{m'} r_{m'}^2 \sigma_{m'}^2 - \lambda \left(\sum_{m'} r'_m - 1 \right)$$

3. Minimize L unconditionally:

$$\frac{\partial L}{\partial r_m} = 2 r_m \sigma_m^2 - \lambda \stackrel{!}{=} 0$$

4. Calculate λ by inserting the result into the restraint: Here, we have $r_m = \lambda / (2\sigma_m^2) \Rightarrow$

$$r_m = \frac{\sigma_m^{-2}}{\sum_{m'} \sigma_{m'}^{-2}}$$

Solving the restrained minimization problem

The method of **Lagrange multipliers** does the magic! With a single restraint (a generalisation is straightforward, see Tutorial 04), do the following:

1. Formulate the restraint as an "=0" equation: $g(\boldsymbol{r}) = \sum_m r_m - 1 = 0$

2. Define the **Lagrange function** by adding to the function f to be minimized the restraint multiplied by a *Lagrange multiplier* λ:

$$L(\boldsymbol{r}) = f(\boldsymbol{r}) - \lambda g(\boldsymbol{r}) = \sum_{m'} r_{m'}^2 \sigma_{m'}^2 - \lambda \left(\sum_{m'} r'_m - 1 \right)$$

3. Minimize L unconditionally:

$$\frac{\partial L}{\partial r_m} = 2 r_m \sigma_m^2 - \lambda \stackrel{!}{=} 0$$

4. Calculate λ by inserting the result into the restraint: Here, we have $r_m = \lambda/(2\sigma_m^2) \Rightarrow$

$$r_m = \frac{\sigma_m^{-2}}{\sum_{m'} \sigma_{m'}^{-2}}$$

Solving the restrained minimization problem

The method of **Lagrange multipliers** does the magic! With a single restraint (a generalisation is straightforward, see Tutorial 04), do the following:

1. Formulate the restraint as an "=0" equation: $g(\boldsymbol{r}) = \sum_m r_m - 1 = 0$

2. Define the **Lagrange function** by adding to the function f to be minimized the restraint multiplied by a *Lagrange multiplier* λ:

$$L(\boldsymbol{r}) = f(\boldsymbol{r}) - \lambda g(\boldsymbol{r}) = \sum_{m'} r_{m'}^2 \sigma_{m'}^2 - \lambda \left(\sum_{m'} r'_m - 1 \right)$$

3. Minimize L unconditionally:

$$\frac{\partial L}{\partial r_m} = 2 r_m \sigma_m^2 - \lambda \stackrel{!}{=} 0$$

4. Calculate λ by inserting the result into the restraint: Here, we have $r_m = \lambda/(2\sigma_m^2) \Rightarrow$

$$r_m = \frac{\sigma_m^{-2}}{\sum_{m'} \sigma_{m'}^{-2}}$$

Solving the restrained minimization problem

The method of **Lagrange multipliers** does the magic! With a single restraint (a generalisation is straightforward, see Tutorial 04), do the following:

1. Formulate the restraint as an "=0" equation: $g(\boldsymbol{r}) = \sum_m r_m - 1 = 0$

2. Define the **Lagrange function** by adding to the function f to be minimized the restraint multiplied by a *Lagrange multiplier* λ:

$$L(\boldsymbol{r}) = f(\boldsymbol{r}) - \lambda g(\boldsymbol{r}) = \sum_{m'} r_{m'}^2 \sigma_{m'}^2 - \lambda \left(\sum_{m'} r'_m - 1 \right)$$

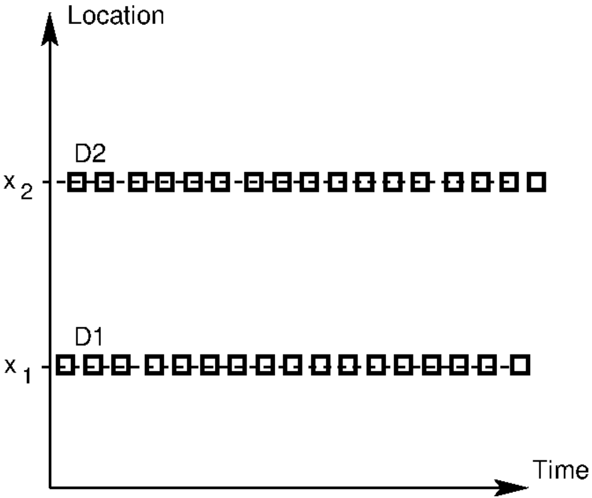
3. Minimize L unconditionally:

$$\frac{\partial L}{\partial r_m} = 2 r_m \sigma_m^2 - \lambda \stackrel{!}{=} 0$$

4. Calculate λ by inserting the result into the restraint: Here, we have $r_m = \lambda/(2\sigma_m^2) \;\Rightarrow$

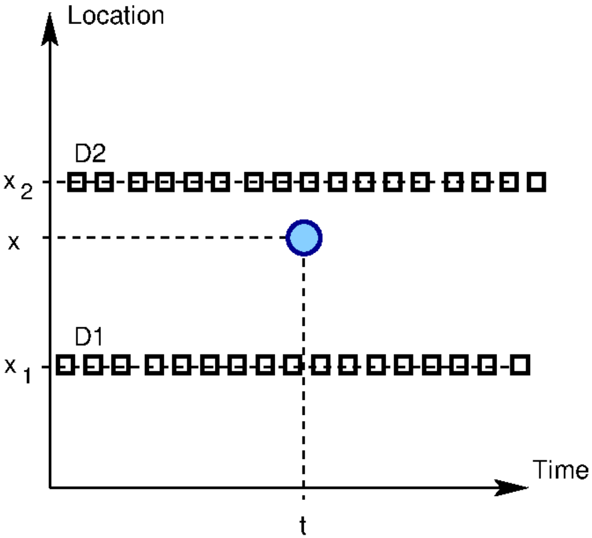
$$r_m = \frac{\sigma_m^{-2}}{\sum_{m'} \sigma_{m'}^{-2}}$$

4.4. Adaptive Smoothing Method (ASM)
1. isotropic smoothing



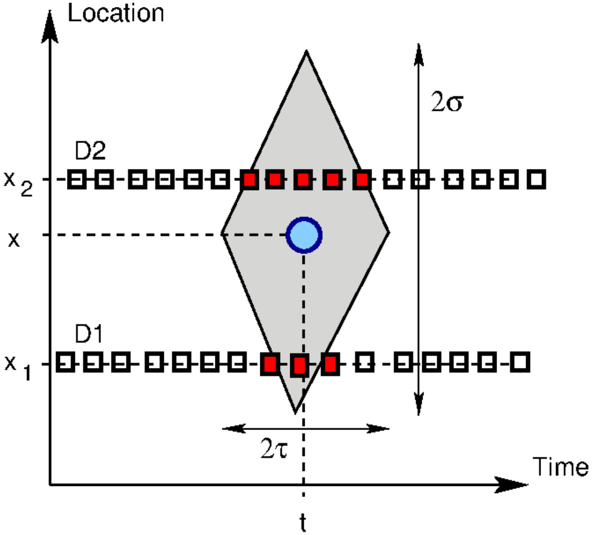
- **Given:** data points $\{(y_i, x_i, t_i)\}$ of quantity Y at the spatiotemporal points (x_i, t_i)

4.4. Adaptive Smoothing Method (ASM)
1. isotropic smoothing



- **Given:** data points $\{(y_i, x_i, t_i)\}$ of quantity Y at the spatiotemporal points (x_i, t_i)
- **Wanted:** Estimate $y(x, t)$ everywhere

4.4. Adaptive Smoothing Method (ASM)
1. isotropic smoothing



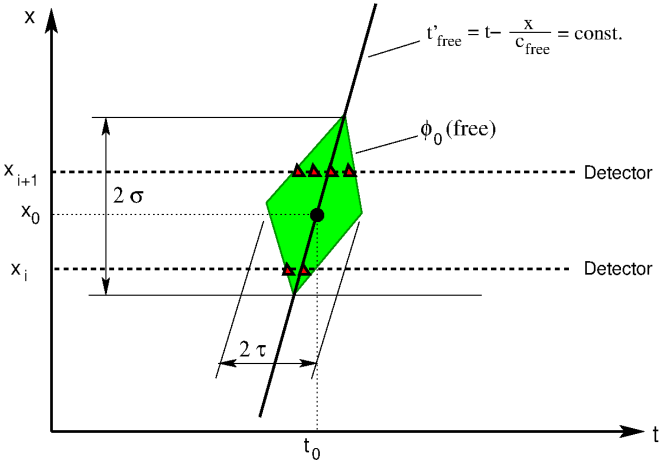
- **Given:** data points $\{(y_i, x_i, t_i)\}$ of quantity Y at the spatiotemporal points (x_i, t_i)
- **Wanted:** Estimate $y(x,t)$ everywhere
- **Isotropic solution:**

  $y(x,t) = \sum_i w_i y_i$ with
  $w_i \propto \phi_0(x - x_i, t - t_i)$ and

$$\phi_0(x,t) = \exp\left[-\left(\frac{|x|}{\sigma} + \frac{|t|}{\tau}\right)\right]$$

Adaptive Smoothing Method
2. anisotropic smoothing

Use smoothing kernels with skewed time axis representing the wave velocities



- "Free" filter with c_{free} near v_0:

$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{free}}}\right)$$

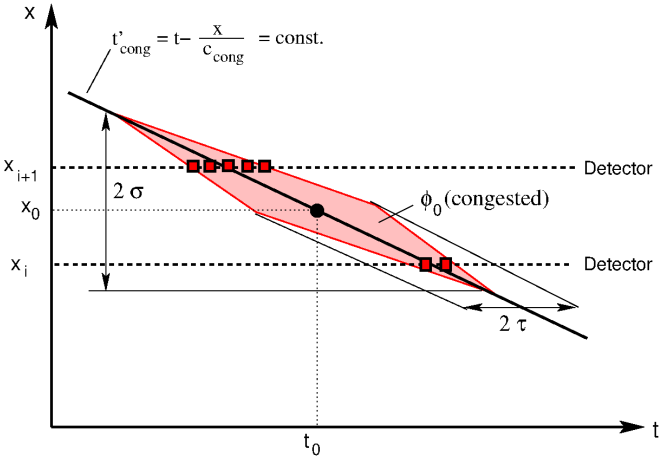
- "Congested" filter with $c_{\text{cong}} \approx -15\,\text{km/h}$:

$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{cong}}}\right)$$

- Weighting of the filters according to the "congested" predictor

Adaptive Smoothing Method
2. anisotropic smoothing

Use smoothing kernels with skewed time axis representing the wave velocities



- "Free" filter with c_{free} near v_0:

$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{free}}}\right)$$

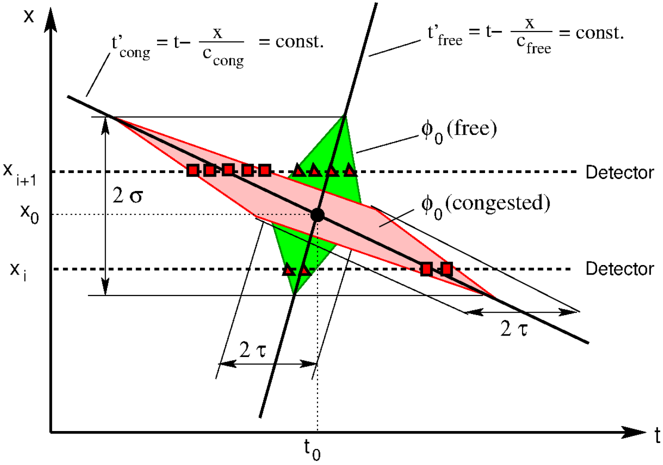
- "Congested" filter with $c_{\text{cong}} \approx -15\,\text{km/h}$:

$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{cong}}}\right)$$

- Weighting of the filters according to the "congested" predictor

Adaptive Smoothing Method
2. anisotropic smoothing

Use smoothing kernels with skewed time axis representing the wave velocities



- "Free" filter with c_{free} near v_0:

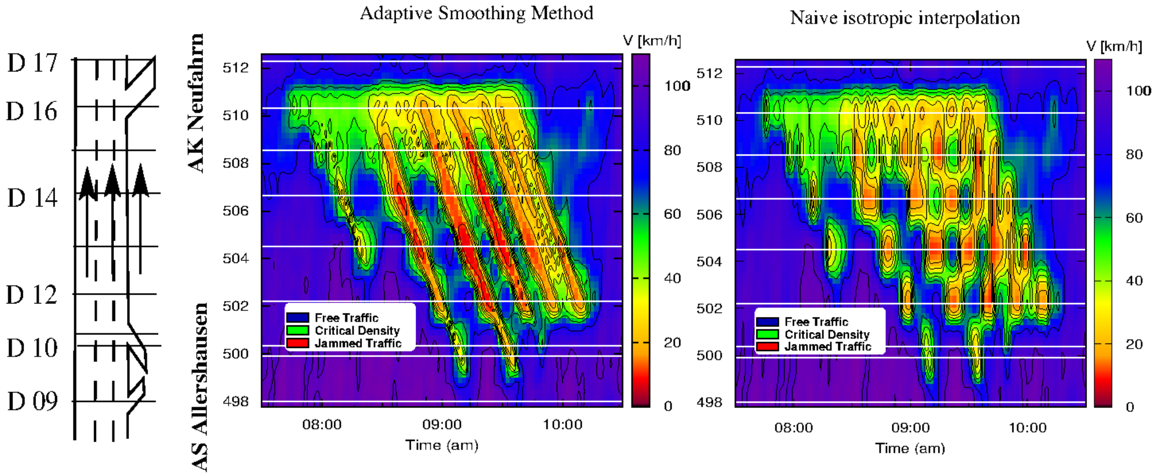
$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{free}}}\right)$$

- "Congested" filter with $c_{\text{cong}} \approx -15\,\text{km/h}$:

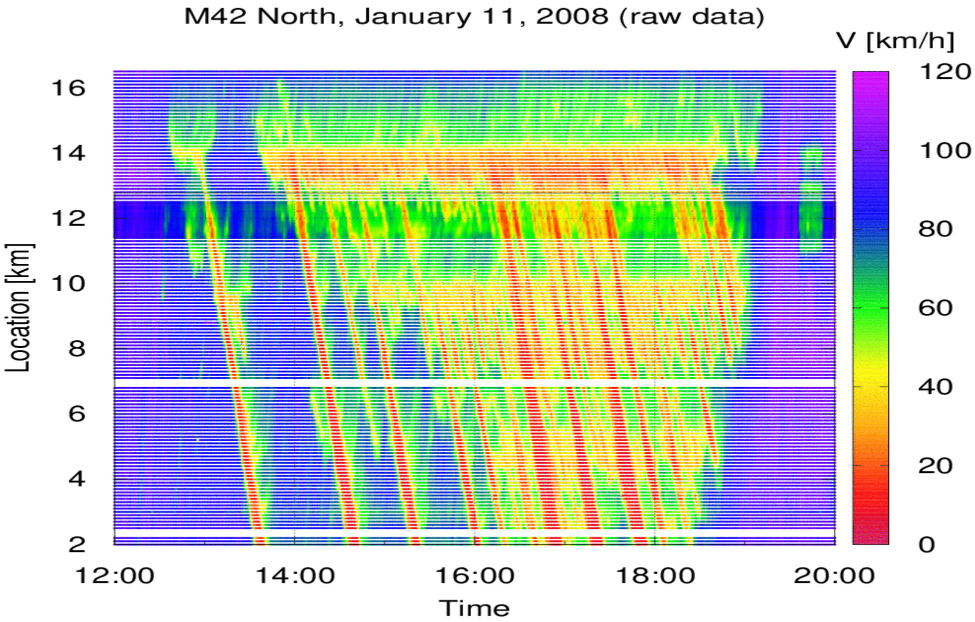
$$w_i \propto \phi_0\left(x - x_i, t - t_i - \frac{x - x_i}{c_{\text{cong}}}\right)$$

- Weighting of the filters according to the "congested" predictor

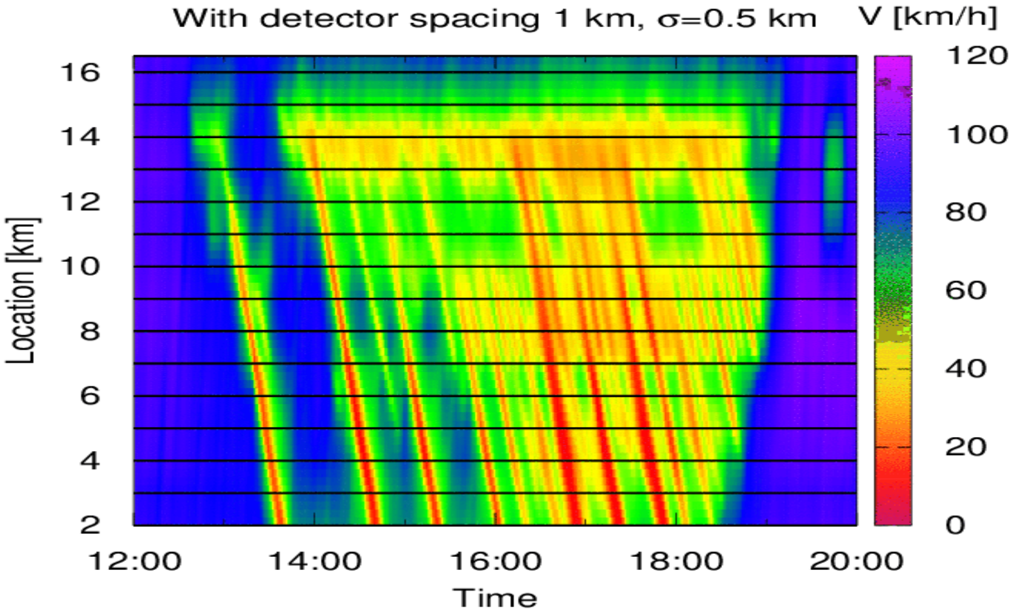
ASM vs. conventional smoothing



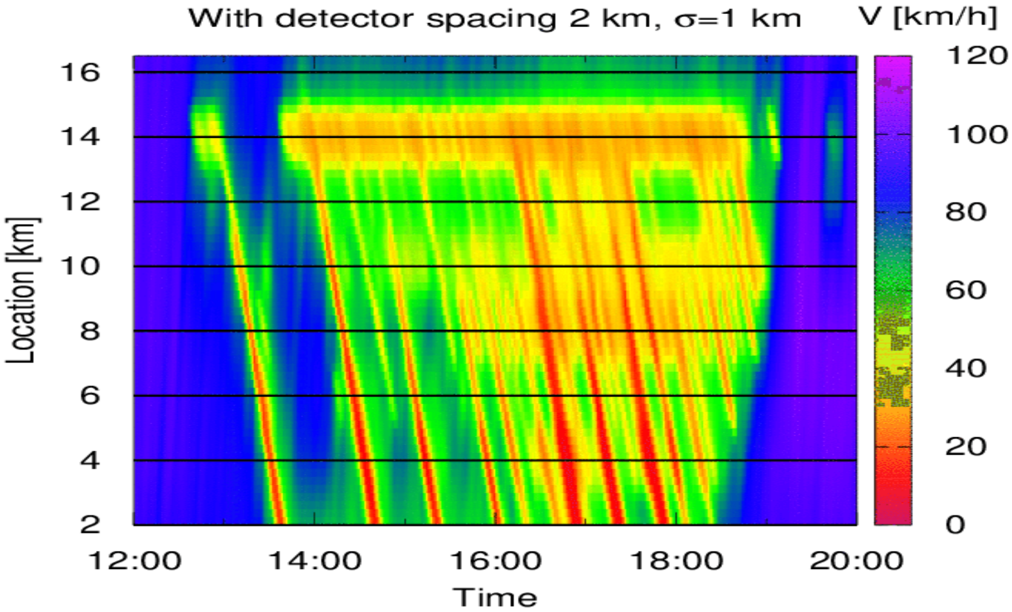
Validation of the Adaptive Smoothing Method: reference



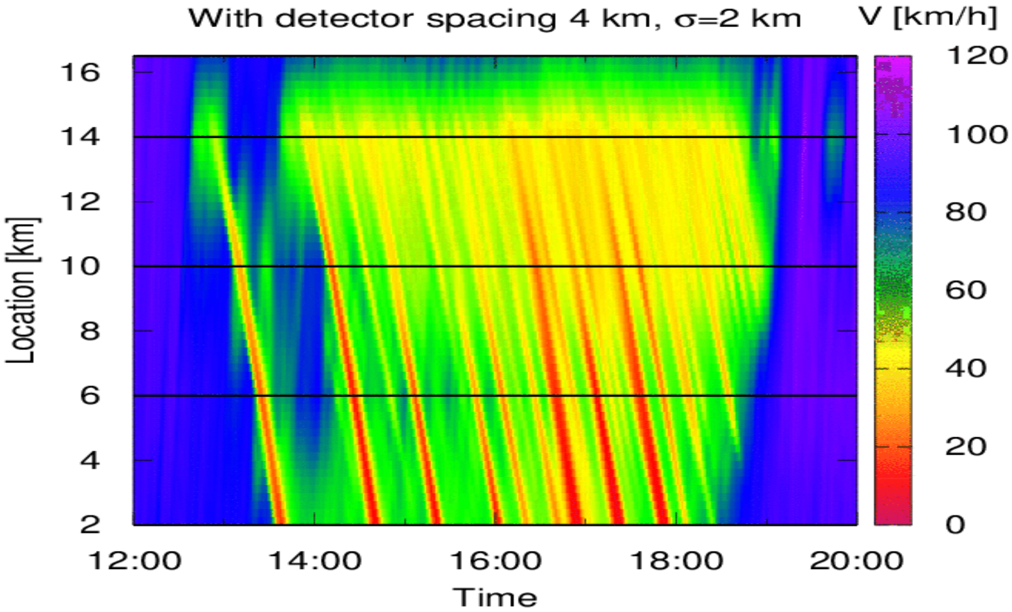
Validation I: detector distance 1 km



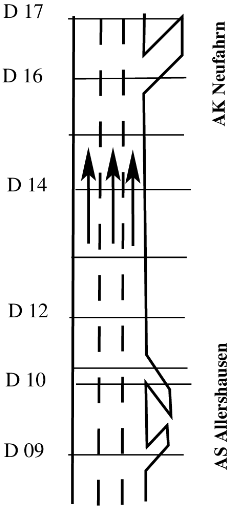
Validation II: detector distance 2 km

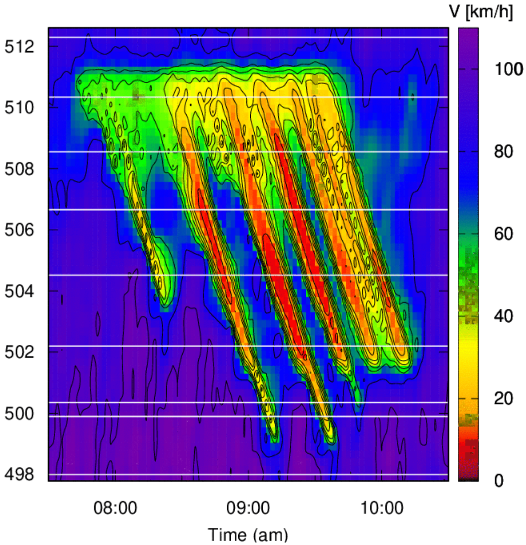


Validation III: detector distance 4 km



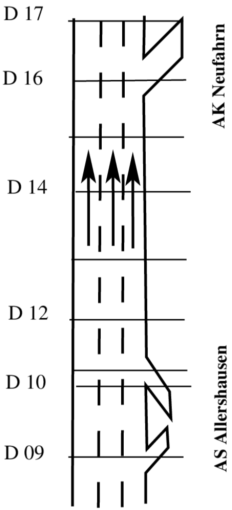
Robustness of the ASM: Sensitivity analysis Reference



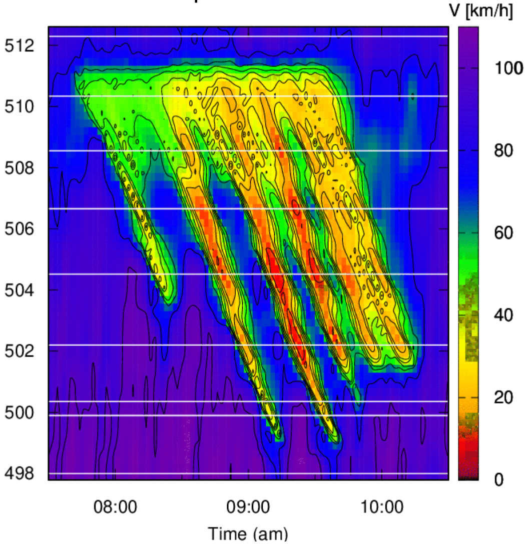


ASM parameters: $\sigma = 600\,\mathrm{m}$, $\tau = 40\,\mathrm{s}$, $c_{\text{free}} = 50\,\mathrm{km/h}$, $w = c_{\text{cong}} = -15\,\mathrm{km/h}$, $vc1 = 50\,\mathrm{km/h}$, $vc2 = 60\,\mathrm{km/h}$

Robustness of the ASM: Sensitivity analysis I



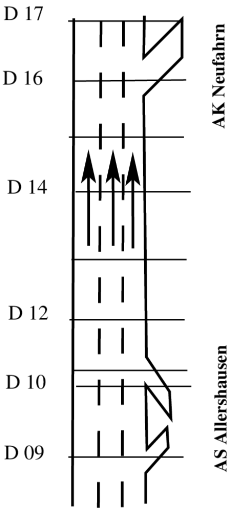
Wave speed w too slow

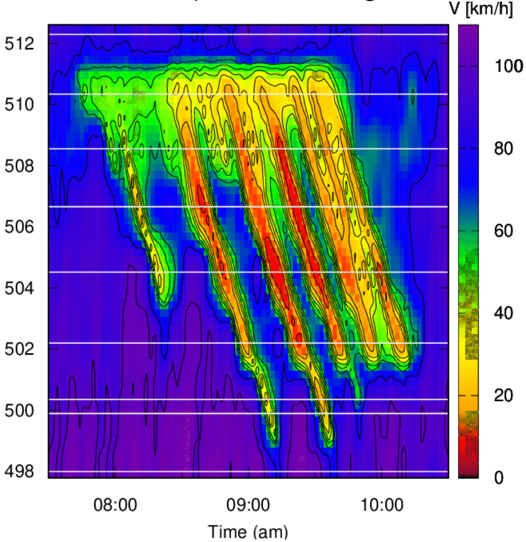


wave speed $w = -10\,\mathrm{km/h}$ instead of $w = -15\,\mathrm{km/h}$

Robustness of the ASM: Sensitivity analysis I

Wave speed w too high

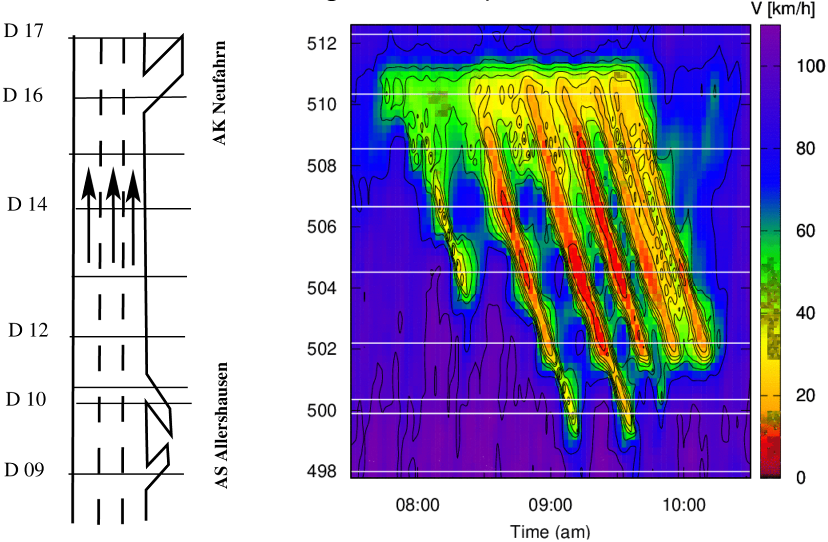




wave speed $w = -20\,\mathrm{km/h}$ instead of $w = -15\,\mathrm{km/h}$

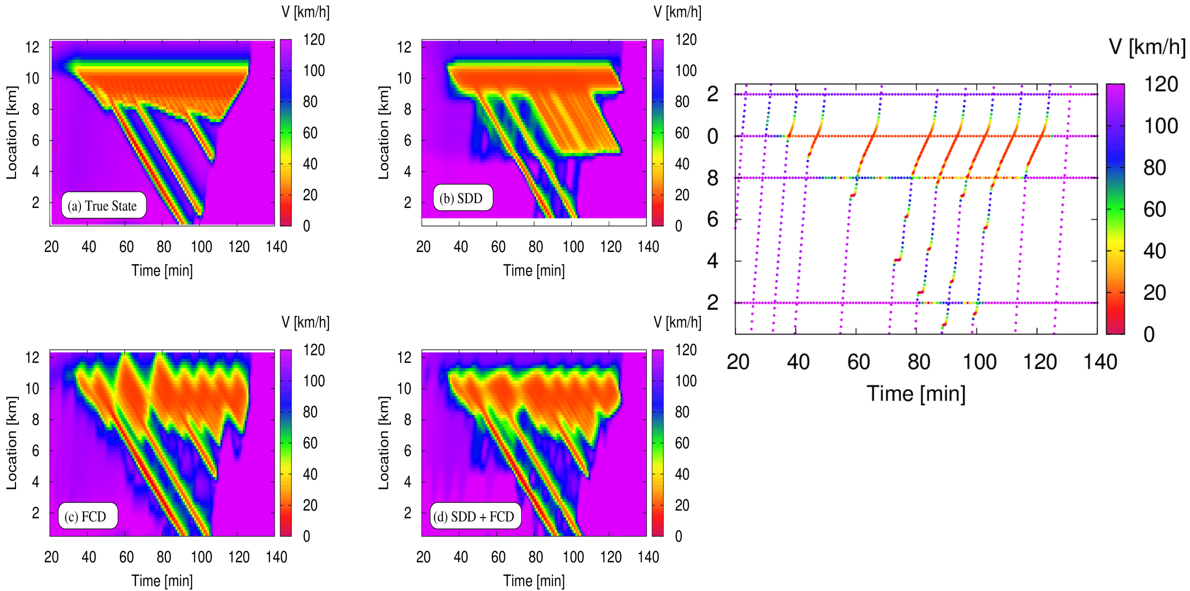
Robustness of the ASM: Sensitivity analysis I

Wrong transition speeds v_{c1} and v_{c2}

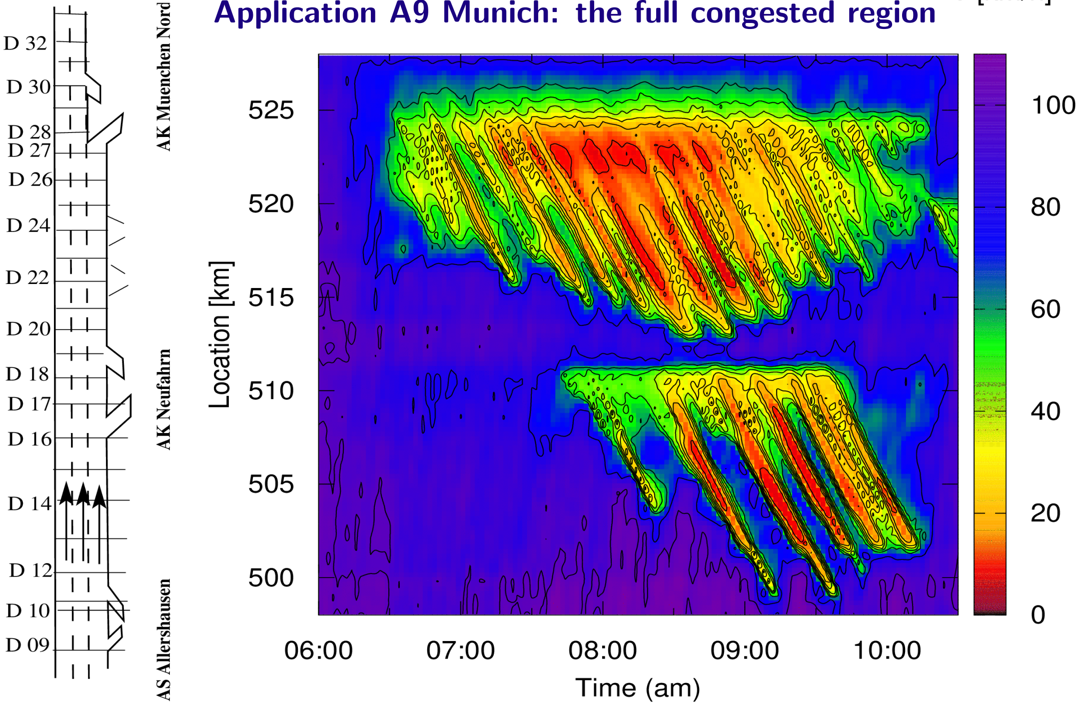


Transit speeds $v_{c1} = 30\,\mathrm{km/h}$ instead of $50\,\mathrm{km/h}$, $v_{c2} = 50\,\mathrm{km/h}$ instead of $60\,\mathrm{km/h}$

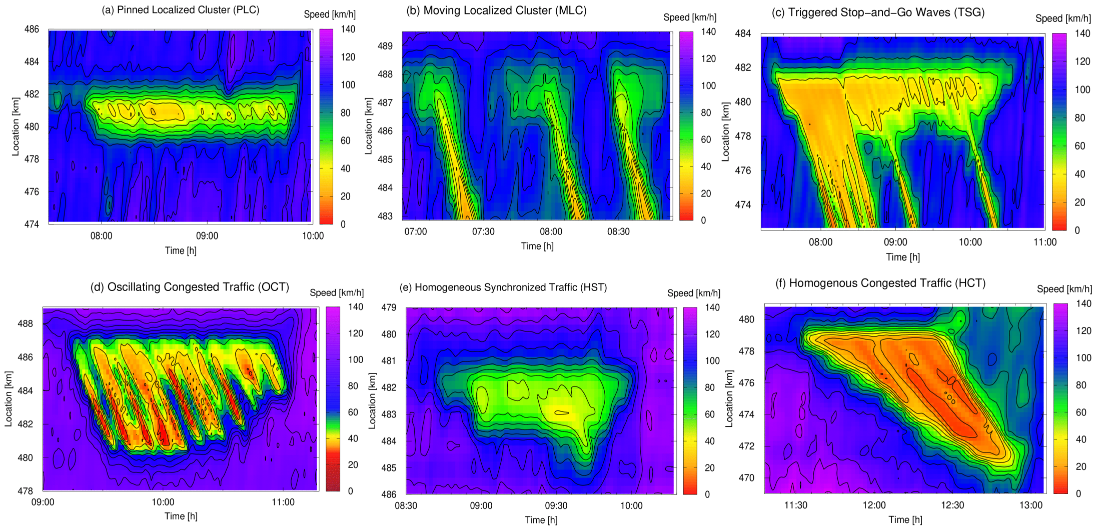
Applying the ASM to SDD, FCD, and both



Application A9 Munich: the full congested region



Application: understanding the dynamics of congestions



$\Rightarrow$ Models!