Maps of Optimal Model Parameters by Efficient Fitting and Extrapolation Techniques

Ralf Thiedmann

rthied@mathematik.uni-ulm.de

University of Ulm, Department of Stochastics

Outline

- **Motivation**

Outline

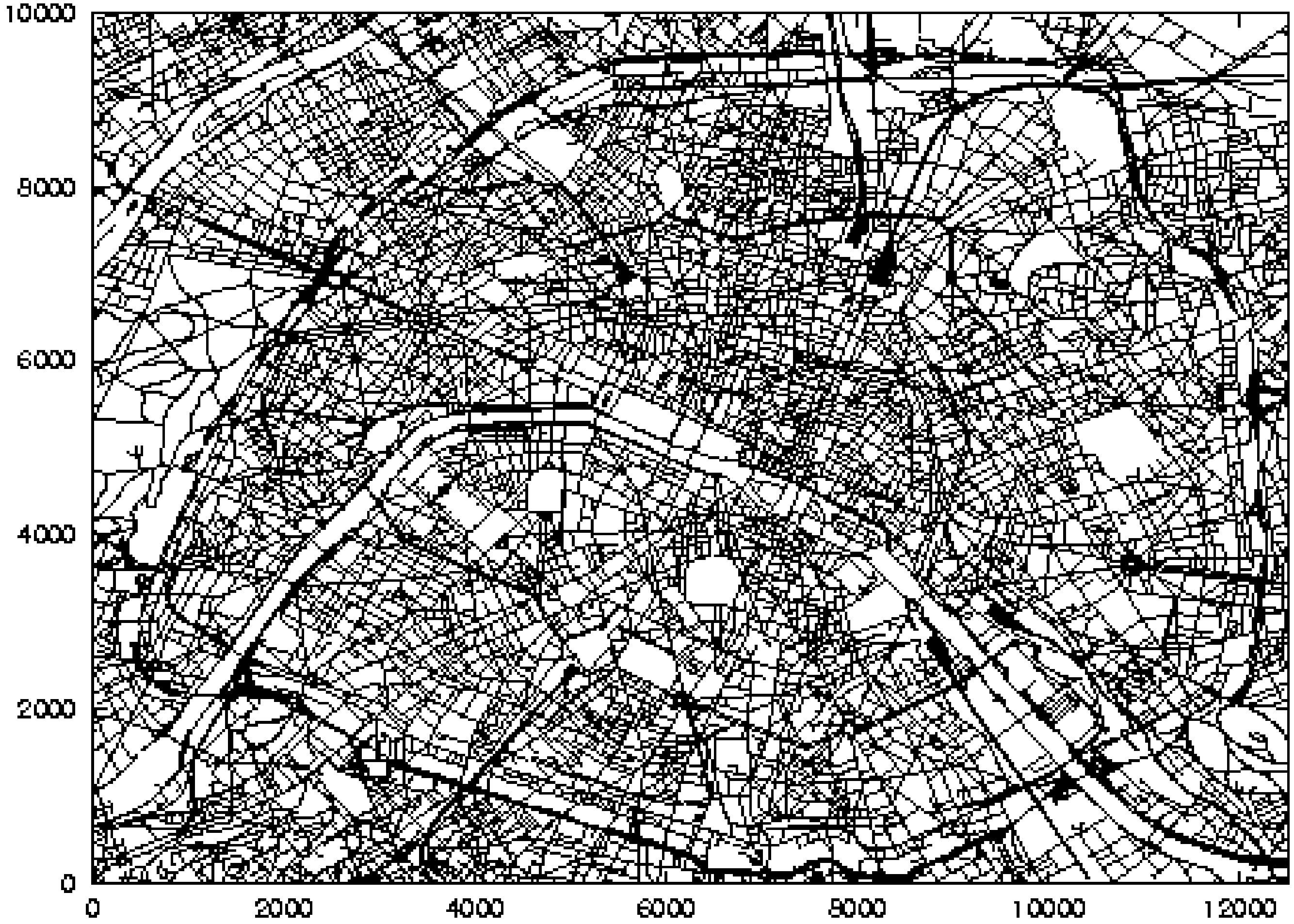
- **Motivation**
- **Modelfitting Procedure**
  - Minimization of Distance Measures
  - Minimization Methods
  - Numerical Examples

Outline

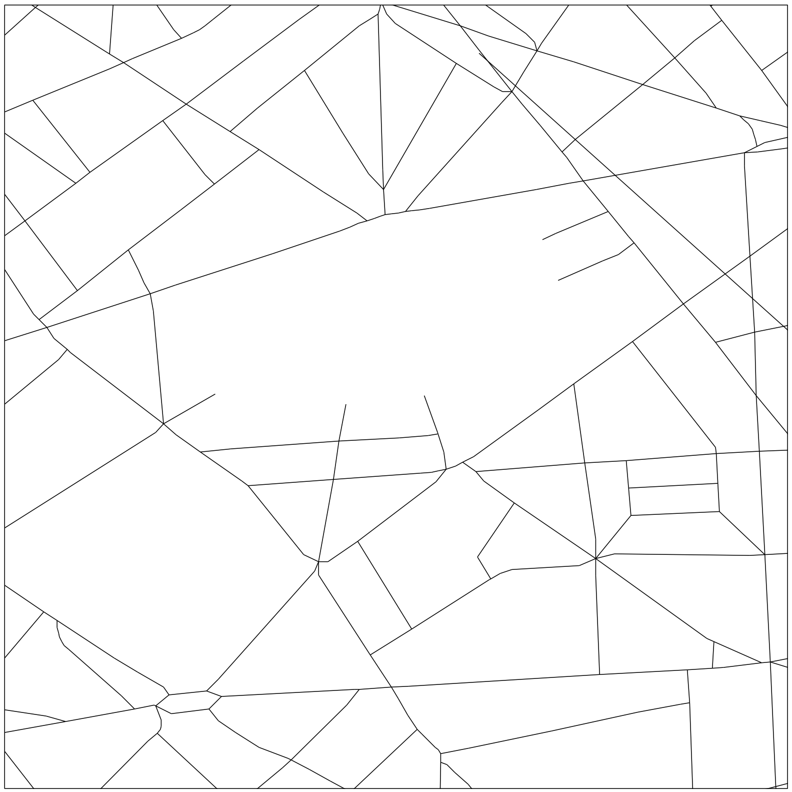
- **Motivation**
- **Modelfitting Procedure**
  - Minimization of Distance Measures
  - Minimization Methods
  - Numerical Examples
- **Extrapolation Techniques**
  - Nonstatistical Extrapolation
  - Intensity Maps

Motivation

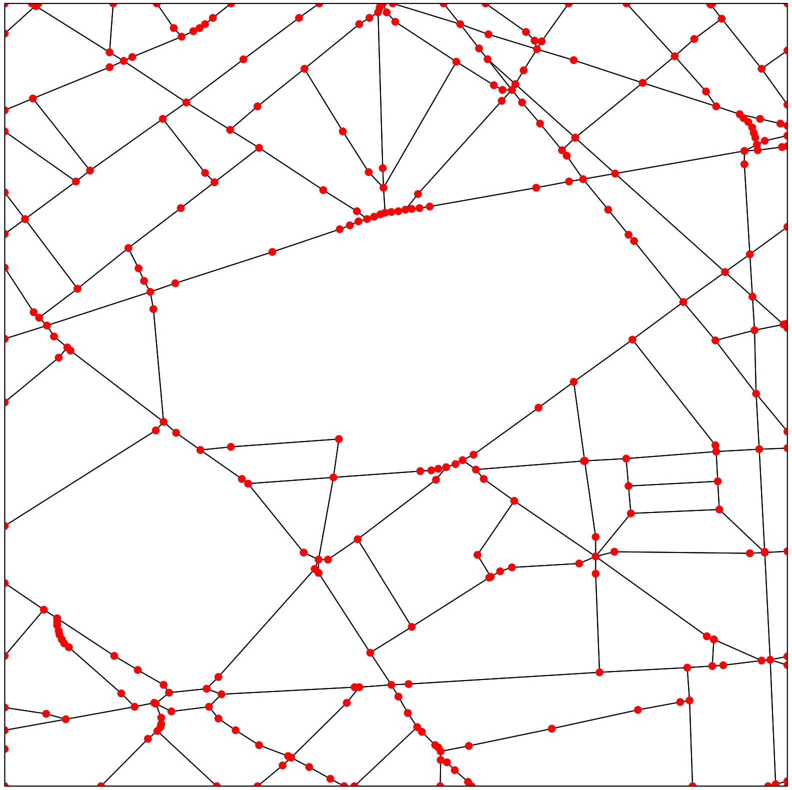
Real Infrastructure Data of Paris



Motivation
Real Infrastructure Data of Paris

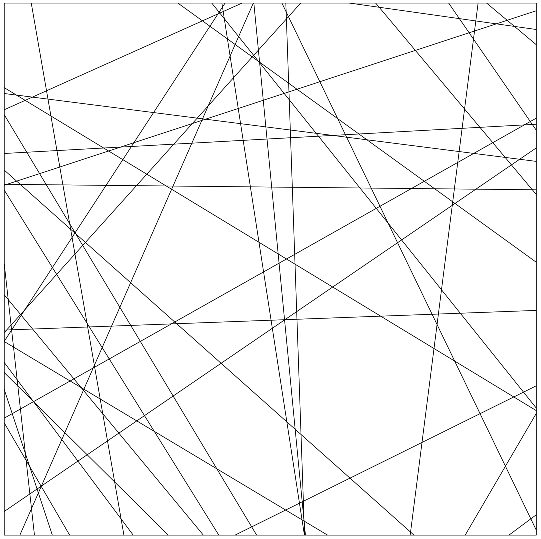


Raw data

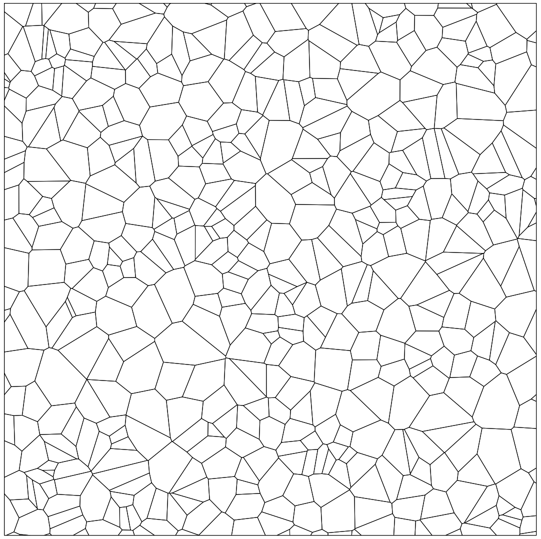


Preprocessed data

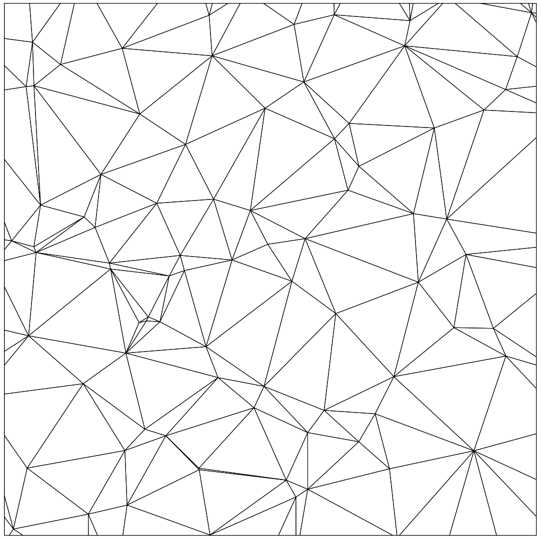
Motivation
Simple Tessellation Models



PLT
(intensity 0.1)



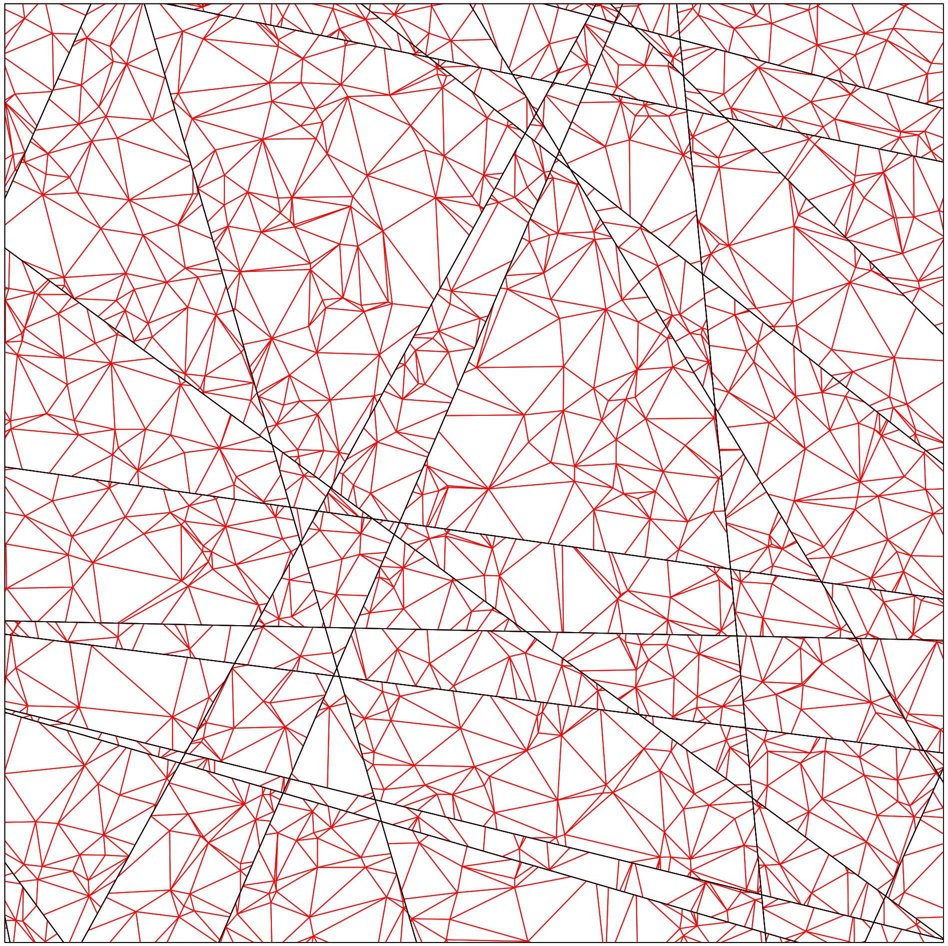
PVT
(intensity 0.005)



PDT
(intensity 0.001)

Motivation

Nested Tessellation Models



PLT/PDT with intensities $\gamma_0 = 0.05$, $\gamma_1 = 0.007$

Minimization of Distance Measures

Simple Tessellations

- Considered tessellations : PLT, PVT, PDT
- Intensity parameter: γ
- Characteristics of simple tessellation per unit area
  - $\lambda_0(\gamma)$ = expected number of nodes
  - $\lambda_1(\gamma)$ = expected number of edge-midpoints
  - $\lambda_2(\gamma)$ = expected number of cell-centroids
  - $\lambda_3(\gamma)$ = expected length of edges

Minimization of Distance Measures

Simple Tessellations

- Considered tessellations : PLT, PVT, PDT
- Intensity parameter: γ
- Characteristics of simple tessellation per unit area
  - $\lambda_0(\gamma)$ = expected number of nodes
  - $\lambda_1(\gamma)$ = expected number of edge-midpoints
  - $\lambda_2(\gamma)$ = expected number of cell-centroids
  - $\lambda_3(\gamma)$ = expected length of edges
- Estimation of these characteristics : $\widehat{\lambda}_0$, $\widehat{\lambda}_1$, $\widehat{\lambda}_2$,$\widehat{\lambda}_3$

Minimization of Distance Measures

Simple Tessellations

	PLT	PVT	PDT
$\lambda_0(\gamma)$	$\frac{1}{\pi}\gamma^2$	2γ	γ
$\lambda_1(\gamma)$	$\frac{2}{\pi}\gamma^2$	3γ	3γ
$\lambda_2(\gamma)$	$\frac{1}{\pi}\gamma^2$	γ	2γ
$\lambda_3(\gamma)$	γ	$2\sqrt{\gamma}$	$\frac{32}{3\pi}\sqrt{\gamma}$

Mean Values for PLT, PVT, and PDT each with Intensity γ

Minimization of Distance Measures

The minimization problem

- Minimization of the relative Euclidean distance

$$F(\gamma) = \sqrt{\sum_{i=0}^{3} \left(\frac{\lambda_i(\gamma) - \widehat{\lambda}_i}{\widehat{\lambda}_i} \right)^2}$$

- Side condition

$$\gamma > 0$$

Minimization of Distance Measures

The minimization problem

- Minimization of the relative Euclidean distance

$$F(\gamma) = \sqrt{\sum_{i=0}^{3} \left(\frac{\lambda_i(\gamma) - \widehat{\lambda}_i}{\widehat{\lambda}_i} \right)^2}$$

- Side condition

$$\gamma > 0$$

- Other distances possible

Minimization of Distance Measures

The minimization problem

- Minimization of the relative Euclidean distance

$$F(\gamma) = \sqrt{\sum_{i=0}^{3} \left(\frac{\lambda_i(\gamma) - \widehat{\lambda}_i}{\widehat{\lambda}_i} \right)^2}$$

- Side condition

$$\gamma > 0$$

- Other distances possible
- Analytical solutions are known for
  - Euclidean distances (absolute and relative)
  - Absolute value distances (absolute and relative)

Minimization of Distance Measures

Nested Tessellations

- Nesting of the simple tessellations PLT, PVT and PDT
- Intensity parameter
  - Intensity of the initial tessellation : γ_0
  - Intensity of the component tessellation : γ_1
  - Bernoulli-thinning parameter : p
- Characteristics of iterated tessellations per unit area
  - $\lambda_0^{It}(\gamma_0, \gamma_1, p)$ = expected number of nodes
  - $\lambda_1^{It}(\gamma_0, \gamma_1, p)$ = expected number of edge-midpoints
  - $\lambda_2^{It}(\gamma_0, \gamma_1, p)$ = expected number of cell-centroids
  - $\lambda_3^{It}(\gamma_0, \gamma_1, p)$ = expected length of edges

Minimization of Distance Measures

Nested Tessellations

Mean value formulae for nested tessellation models

$$
\begin{aligned}
\lambda_0^{It}(\gamma_0, \gamma_1, p) &= \lambda_0(\gamma_0) + p\lambda_0(\gamma_1) + \frac{4p}{\pi}\lambda_3(\gamma_0)\lambda_3(\gamma_1) \\
\lambda_1^{It}(\gamma_0, \gamma_1, p) &= \lambda_1(\gamma_0) + p\lambda_1(\gamma_1) + \frac{6p}{\pi}\lambda_3(\gamma_0)\lambda_3(\gamma_1) \\
\lambda_2^{It}(\gamma_0, \gamma_1, p) &= \lambda_2(\gamma_0) + p\lambda_2(\gamma_1) + \frac{2p}{\pi}\lambda_3(\gamma_0)\lambda_3(\gamma_1) \\
\lambda_3^{It}(\gamma_0, \gamma_1, p) &= \lambda_3(\gamma_0) + p\lambda_3(\gamma_1)
\end{aligned}
$$

Minimization of Distance Measures

The minimization problem

- Minimization of the relative Euclidean distance

$$F(\gamma_0, \gamma_1, p) = \sqrt{\sum_{i=0}^{3} \left(\frac{\lambda_i(\gamma_0, \gamma_1, p) - \widehat{\lambda}_i}{\widehat{\lambda}_i} \right)^2}$$

- Side conditions

$$0 \leq p \leq 1$$

$$\gamma_0, \gamma_1 > 0$$

Minimization of Distance Measures

The minimization problem

- Minimization of the relative Euclidean distance

$$F(\gamma_0, \gamma_1, p) = \sqrt{\sum_{i=0}^{3} \left(\frac{\lambda_i(\gamma_0, \gamma_1, p) - \widehat{\lambda}_i}{\widehat{\lambda}_i} \right)^2}$$

- Side conditions

$$0 \leq p \leq 1$$

$$\gamma_0, \gamma_1 > 0$$

- Other distances possible

Minimization of Distance Measures

Example PLT/PDT

- $$F(\gamma_0,\gamma_1,p)^2 = \left(\frac{\frac{\gamma_0^2}{\pi}+p\gamma_1+\frac{128}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_0}{\widehat{\lambda}_0}\right)^2 + \left(\frac{\frac{2\gamma_0^2}{\pi}+3p\gamma_1+\frac{64}{\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_1}{\widehat{\lambda}_1}\right)^2 + \left(\frac{\frac{\gamma_0^2}{\pi}+2p\gamma_1+\frac{64}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_2}{\widehat{\lambda}_2}\right)^2 + \left(\frac{\gamma_0+\frac{32}{3\pi}p\sqrt{\gamma_1}-\widehat{\lambda}_3}{\widehat{\lambda}_3}\right)^2$$

Minimization of Distance Measures
Example PLT/PDT

- $$F(\gamma_0,\gamma_1,p)^2 = \left(\frac{\frac{\gamma_0^2}{\pi}+p\gamma_1+\frac{128}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_0}{\widehat{\lambda}_0}\right)^2 + \left(\frac{\frac{2\gamma_0^2}{\pi}+3p\gamma_1+\frac{64}{\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_1}{\widehat{\lambda}_1}\right)^2 + \left(\frac{\frac{\gamma_0^2}{\pi}+2p\gamma_1+\frac{64}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_2}{\widehat{\lambda}_2}\right)^2 + \left(\frac{\gamma_0+\frac{32}{3\pi}p\sqrt{\gamma_1}-\widehat{\lambda}_3}{\widehat{\lambda}_3}\right)^2$$
- Not convex in γ_0,γ_1

Minimization of Distance Measures

Example PLT/PDT

- $$F(\gamma_0,\gamma_1,p)^2 = \left(\frac{\frac{\gamma_0^2}{\pi}+p\gamma_1+\frac{128}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_0}{\widehat{\lambda}_0}\right)^2 + \left(\frac{\frac{2\gamma_0^2}{\pi}+3p\gamma_1+\frac{64}{\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_1}{\widehat{\lambda}_1}\right)^2 + \left(\frac{\frac{\gamma_0^2}{\pi}+2p\gamma_1+\frac{64}{3\pi^2}p\gamma_0\sqrt{\gamma_1}-\widehat{\lambda}_2}{\widehat{\lambda}_2}\right)^2 + \left(\frac{\gamma_0+\frac{32}{3\pi}p\sqrt{\gamma_1}-\widehat{\lambda}_3}{\widehat{\lambda}_3}\right)^2$$
- Not convex in γ_0, γ_1
- But convex in p (for all considered tessellations)

Minimization Methods
Sequential Optimization

- $F(\gamma_0, \gamma_1, p)$ is convex in p
- Global minimum of $F(\gamma_0, \gamma_1, p)$ w.r.t. p is root of derivative

$$p(\gamma_0, \gamma_1) = -\left(\frac{\frac{\lambda_0^0 - \widehat{\lambda}_0}{\widehat{\lambda}_0} g_0 + \frac{\lambda_1^0 - \widehat{\lambda}_1}{\widehat{\lambda}_1} g_1 + \frac{\lambda_2^0 - \widehat{\lambda}_2}{\widehat{\lambda}_2} g_2 + \frac{\lambda_3^0 - \widehat{\lambda}_3}{\widehat{\lambda}_3} g_3}{\frac{g_0^2}{\widehat{\lambda}_0} + \frac{g_1^2}{\widehat{\lambda}_1} + \frac{g_2^2}{\widehat{\lambda}_2} + \frac{g_3^2}{\widehat{\lambda}_3}} \right)$$

with

- $g_i = g_i(\gamma_0, \gamma_1) = \left(\frac{\partial \lambda_i^{It}(\gamma_0, \gamma_1, p)}{\partial p} \right) (\gamma_0, \gamma_1) \qquad i = 0, .., 3$
- $\lambda_i^0 = i$-th characteristic of initial tessellation

Minimization Methods
Sequential Optimization

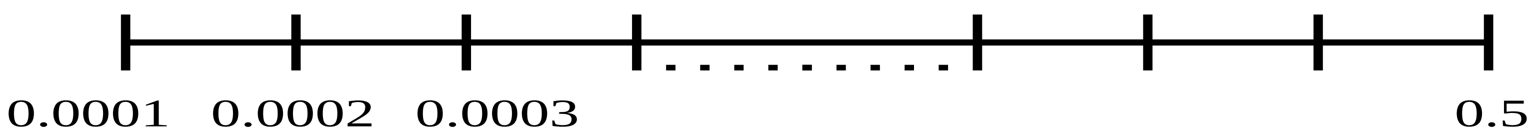
- $F(\gamma_0, \gamma_1, p)$ is convex in p
- Global minimum of $F(\gamma_0, \gamma_1, p)$ w.r.t. p is root of derivative

$$p(\gamma_0, \gamma_1) = -\left(\frac{\frac{\lambda_0^0 - \widehat{\lambda}_0}{\widehat{\lambda}_0} g_0 + \frac{\lambda_1^0 - \widehat{\lambda}_1}{\widehat{\lambda}_1} g_1 + \frac{\lambda_2^0 - \widehat{\lambda}_2}{\widehat{\lambda}_2} g_2 + \frac{\lambda_3^0 - \widehat{\lambda}_3}{\widehat{\lambda}_3} g_3}{\frac{g_0^2}{\widehat{\lambda}_0} + \frac{g_1^2}{\widehat{\lambda}_1} + \frac{g_2^2}{\widehat{\lambda}_2} + \frac{g_3^2}{\widehat{\lambda}_3}} \right)$$

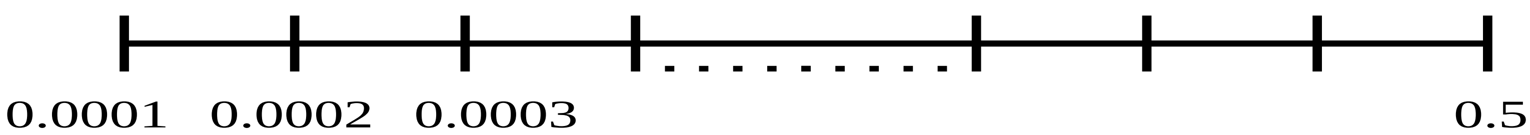
  with
  - $g_i = g_i(\gamma_0, \gamma_1) = \left(\frac{\partial \lambda_i^{It}(\gamma_0, \gamma_1, p)}{\partial p} \right) (\gamma_0, \gamma_1) \qquad i = 0, .., 3$
  - $\lambda_i^0 = i$-th characteristic of initial tessellation
- Insertion into $F(\gamma_0, \gamma_1, p)$
  $\Rightarrow F(\gamma_0, \gamma_1, p(\gamma_0, \gamma_1)) = \widetilde{F}(\gamma_0, \gamma_1)$

Minimization Methods
Traversing method

- $\widetilde{F}(\gamma_0, \gamma_1) \rightarrow min$
- Intensities $\gamma_0 \in (0, \gamma_0^{max}), \gamma_1 \in (0, \gamma_1^{max})$
- Evaluate $\widetilde{F}$
  - for all γ_0 between 0 and γ_0^{max} (i.e.$= 0.5$) with distance, i.e. 0.0001



  - for all γ_1 between 0 and γ_0^{max} (i.e.$= 0.5$) with distance, i.e. 0.0001



- Take minimal value of $\widetilde{F}$

Minimization Methods
Traversing Method

- Advantage : Minimum is always found
- Problem : Critical Runtime
  - i.e 5000^2 computation cases
  - More precise results for smaller steps but much more runtime

Minimization Methods
Nelder Mead (1965)

- Numerical iteration method
- Direct search method without side conditions
- Side conditions are represented by a penalty value
- $F : \mathbb{R}^2 \longrightarrow \mathbb{R}$
- Method based on evaluating F at the vertices of a triangle
- After evaluating F: decision to reflect, expand, contract or shrink the triangle
- Algorithm stops if difference between function values at the vertices is small

Minimization Methods
Nelder Mead - Algorithm

- The vertices $x_1, x_2, x_3 \in \mathbb{R}$ ordered w.l.o.g. $F(x_1) \leq F(x_2) \leq F(x_3)$
- Calculate the middle of the best vertices

$$x^c = \frac{1}{2}(x_1 + x_2)$$

- Calculate the reflection on x^c

$$x^r = x^c + \rho(x^c - x_3) \quad \rho > 0$$

Minimization Methods
Nelder Mead - Algorithm

- If $F(x_1) \leq F(x^r) < F(x_2)$
  - $x_3 = x^r$
- If $F(x^r) < F(x_1)$ try expansion

$$x^e = x^c + \beta(x^r - x^c) \qquad \beta > 1$$

  - If $F(x^e) < F(x^r)$: $x_3 = x^e$
  - Else : $x_3 = x^r$
- If $F(x^r) \geq F(x_2)$ try contraction

Minimization Methods
Nelder Mead - Algorithm

- If $F(x^r) \geq F(x_2)$ try contraction
  - If $F(x^r) \geq F(x_3)$

$$x^k = x^c + \delta(x_3 - x^c) \qquad 0 < \delta < 1$$

    - If $F(x^k) < F(x_3)$: $x_3 = x^k$
    - Else shrinking

$$x_j = \frac{1}{2}(x_j + x_1) \qquad j = 2, 3$$

  - Else

$$x^k = x^c + \delta(x^r - x^c) \qquad 0 < \delta < 1$$

    - If $F(x^k) < F(x^r)$: $x_3 = x^k$
    - Else shrinking

$$x_j = \frac{1}{2}(x_j + x_1) \qquad j = 2, 3$$

Minimization Methods
Nelder Mead - Algorithm

- Stopping condition

$$\sqrt{\frac{1}{3}\sum_{j=1}^{3}|F(x_j) - F(x^c)|^2} < \varepsilon$$

with given $\varepsilon > 0$

Minimization Methods
Nelder Mead (1965)

- Problem : Global minimum may depend on initial values
  - Start with different initial values (randomly chosen)

Minimization Methods

Nelder Mead (1965)

- Problem : Global minimum may depend on initial values
  - Start with different initial values (randomly chosen)
- Advantage :
  - Faster than Traversing-method
  - Most times it delivers more precise results than Traversing-method

Minimization Methods
Projected Gradient Algorithm

- Numerical iteration method
- Usage of the gradient
- Side conditions are included by a projection $Pr_D : \mathbb{R}^n \to D$, where

$$Pr_D(x) = arg \min_{y \in D} ||y - x|| \,.$$

Minimization Methods
Projected Gradient Algorithm

Choose parameters $0 < \beta, \mu < 1$, and $\gamma > 0$

```
Find  x_0 ∈ D
For  k = 0, 1, 2, ...  do:
   Compute  ∇f(x_k)
   IF  x_k  is stationary point THEN
      stop
   ELSE
      x_k(α) = Pr_D(x_k − α ∇ f(x_k))  with  α > 0
      x_{k+1} = x_k(α_k),  with  α_k = β^{m_k} γ  where  m_k  the
      lowest natural number for which
      f(x_{k+1}) ≤ f(x_k) + μ ∇ f(x_k)^T (x_{k+1} − x_k)
      holds.
```

Minimization Methods
Projected Gradient Algorithm

- Advantage :
  - Very fast convergence
  - High precision of the found (local) minimum

Minimization Methods
Projected Gradient Algorithm

- Advantage :
  - Very fast convergence
  - High precision of the found (local) minimum
- Problem : Global minimum is not found for the considered fitting problem even if there are different initial values (randomly chosen)

Minimization Methods

Projected Gradient Algorithm

- Advantage :
  - Very fast convergence
  - High precision of the found (local) minimum
- Problem : Global minimum is not found for the considered fitting problem even if there are different initial values (randomly chosen)
- Application : Improvement of the minimum found by the Nelder-Mead algorithm

Numerical Examples
Comparison of Minimization Methods - Simulated Data

Simulation of a PLT/PLT with intensities $\gamma_0 = 0.1$, $\gamma_1 = 0.06$, $p = 1.0$

Numerical Examples

Comparison of Minimization Methods - Simulated Data

Simulation of a PLT/PLT with intensities $\gamma_0 = 0.1$, $\gamma_1 = 0.06$, $p = 1.0$

Method	Results					
	Best fitting model	Rel. Eucl. distance	γ_0	γ_1	p	Times of NMA time
Traversing-Method	PLT/PLT	0.008977	0.061400	0.100600	1.0	620
Nelder Mead Algorithm (NMA)	PLT/PLT	0.008964	0.061663	0.100305	1.0	1
Projected Gradients (PG)	PLT/PLT	0.017023	0.080599	0.080599	1.0	2
NMA & PG	PLT/PLT	0.008964	0.061663	0.100305	1.0	1

Fixed Bernoulli thinning parameter $p = 1.0$

Numerical Examples

Comparison of Minimization Methods - Simulated Data

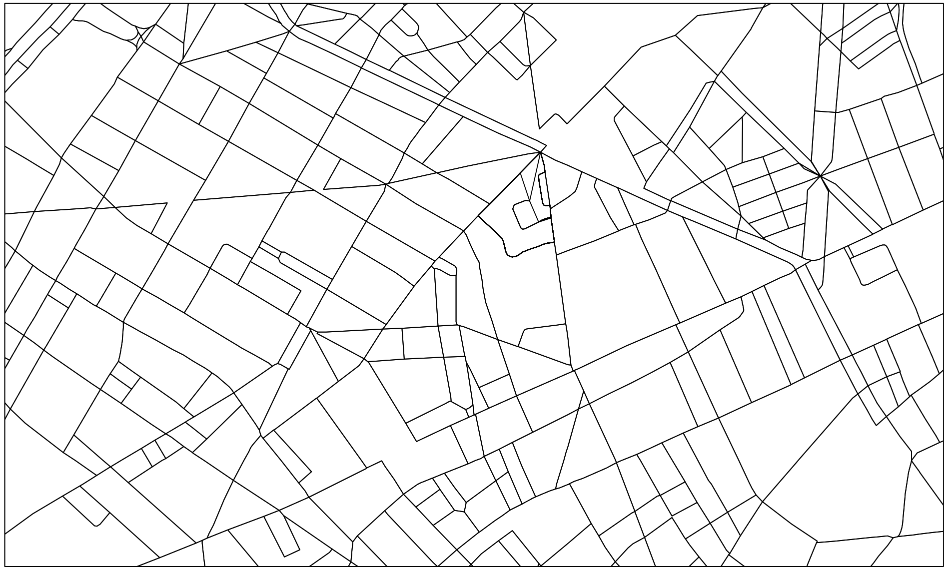
Simulation of a PDT/PVT with intensities $\gamma_0 = 0.001$, $\gamma_1 = 0.005$, $p = 0.9$

Method	Results					
	Best fitting model	Rel. Eucl. distance	γ_0	γ_1	p	Times of NMA time
Traversing-Method	PDT/PVT	0.008605	0.000867	0.007534	0.700000	2400
Nelder Mead Algorithm (NMA)	PDT/PVT	0.008598	0.000868	0.007535	0.700000	1
Projected Gradients (PG)	PVT/PLT	0.046419	0.007372	0.057193	0.940029	1
NMA & PG	PDT/PVT	0.008598	0.000868	0.007535	0.700000	1

Bernoulli thinning parameter p in $[0.7, 1.0]$

Numerical Examples

Comparison of Minimization Methods - Real Data



Part of Paris

Estimated characteristic values:

$\widehat{\lambda}_0 = 0.000106$, $\widehat{\lambda}_1 = 0.000177$, $\widehat{\lambda}_2 = 0.000072$, $\widehat{\lambda}_0 = 0.017659$.

Numerical Examples
Comparison of Minimization Methods - Real Data

Method	Results					
	Best fitting model	Rel. Eucl. distance	γ_0	γ_1	p	Times of NMA time
Traversing-Method	PVT/PDT	0.019028	$6.250 \cdot 10^{-6}$	$1.334 \cdot 10^{-5}$	1.0	230
Nelder Mead Algorithm (NMA)	PVT/PDT	0.018760	$6.307 \cdot 10^{-6}$	$1.329 \cdot 10^{-5}$	1.0	1
Projected Gradients (PG)	PDT/PDT	0.031667	$6.357 \cdot 10^{-6}$	$6.357 \cdot 10^{-6}$	1.0	0.1
NMA & PG	PVT/PDT	0.018760	$6.307 \cdot 10^{-6}$	$1.329 \cdot 10^{-5}$	1.0	1

Fixed Bernoulli thinning parameter $p = 1$

- Simulation of a PLT/PLT with intensities $\gamma_0 = 0.05$, $\gamma_1 = 0.04$, and $p = 1.0$.
- 1 time fitting with traversing method
- 1000 times fitting with Nelder-Mead algorithm

This is repeated 50 times.

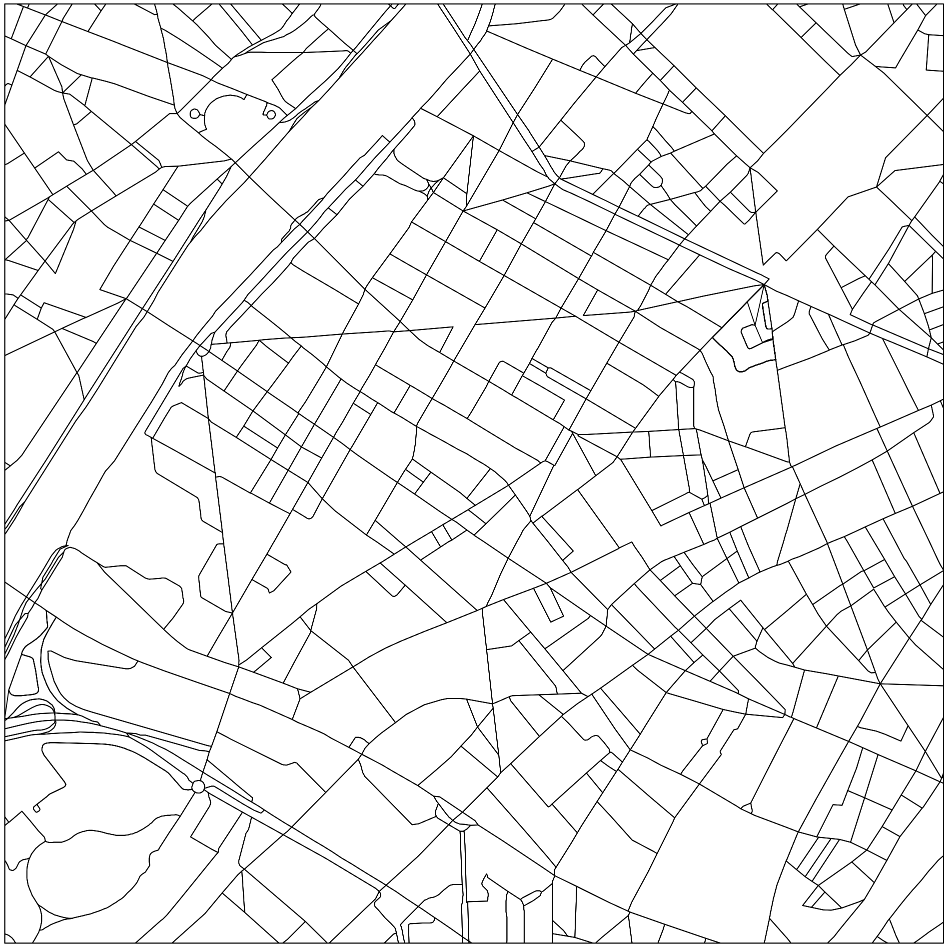
Numerical Examples
Experimental Verification of Nelder-Mead

- Simulation of a PLT/PLT with intensities $\gamma_0 = 0.05$, $\gamma_1 = 0.04$, and $p = 1.0$.
- 1 time fitting with traversing method
- 1000 times fitting with Nelder-Mead algorithm

This is repeated 50 times.

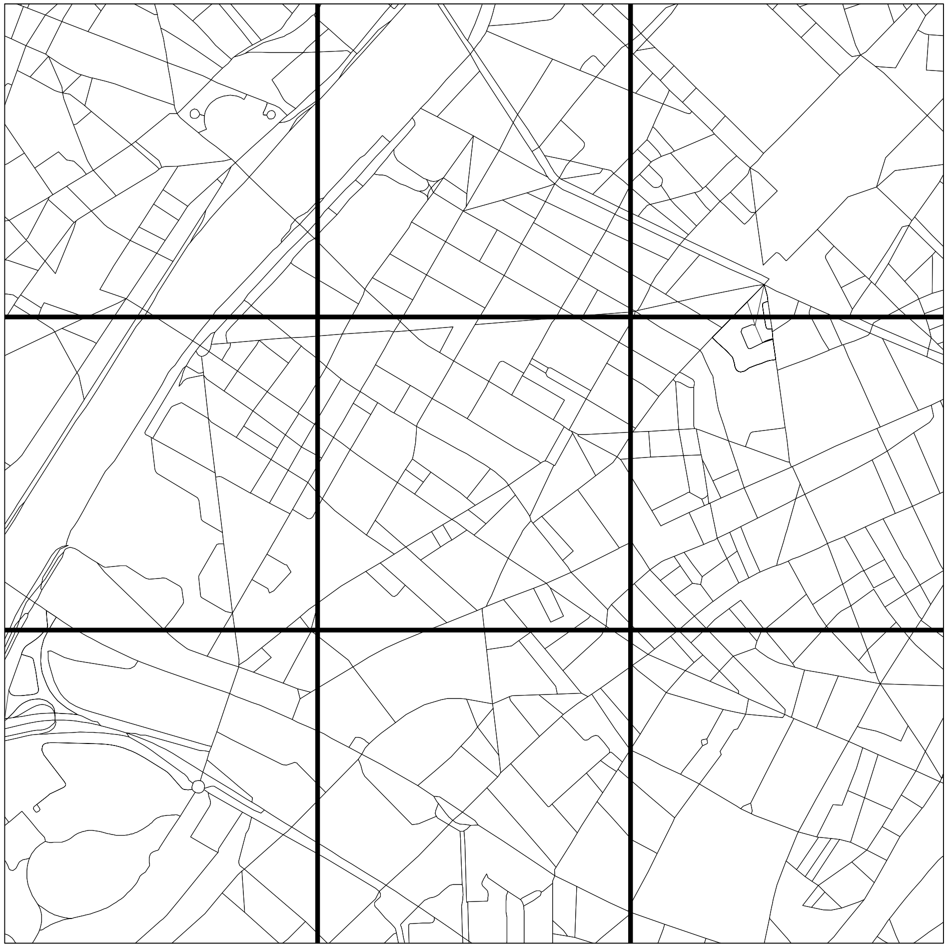
Result:
In all 50.000 cases the minimum found by Nelder-Mead is lower or equal than the minimum found by the Traversing method.

Extrapolation Techniques
Motivation



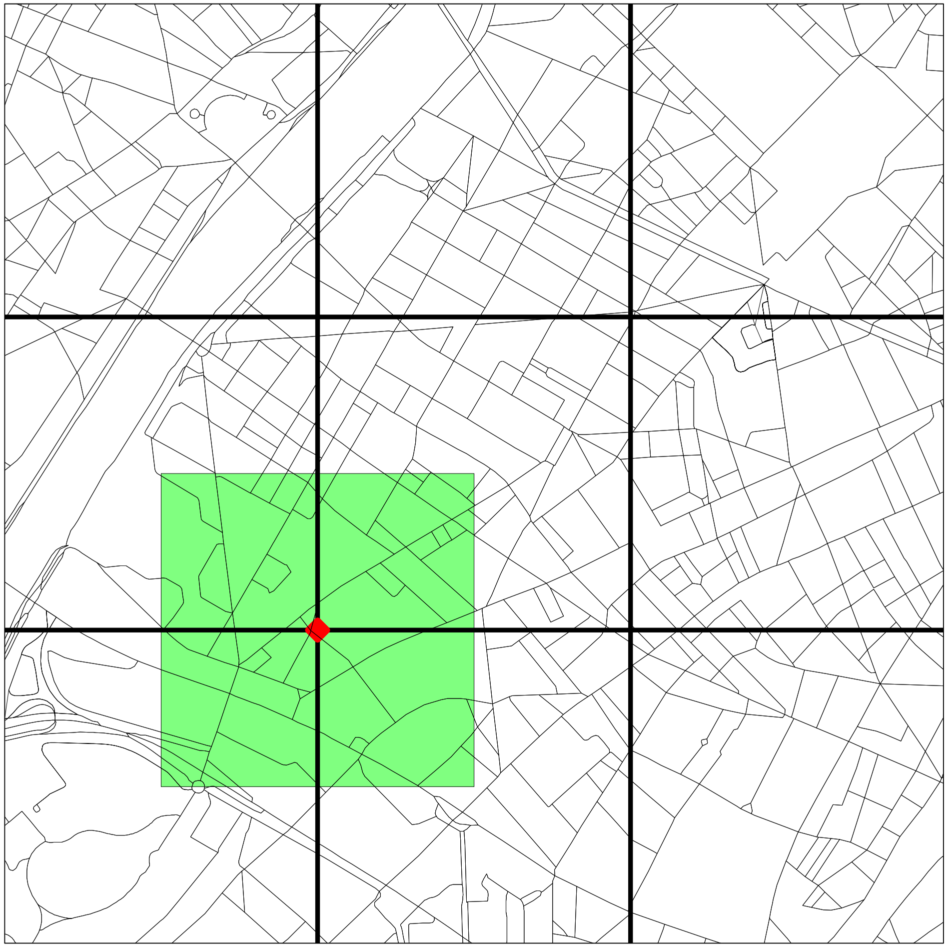
Data of Paris

Extrapolation Techniques
Motivation



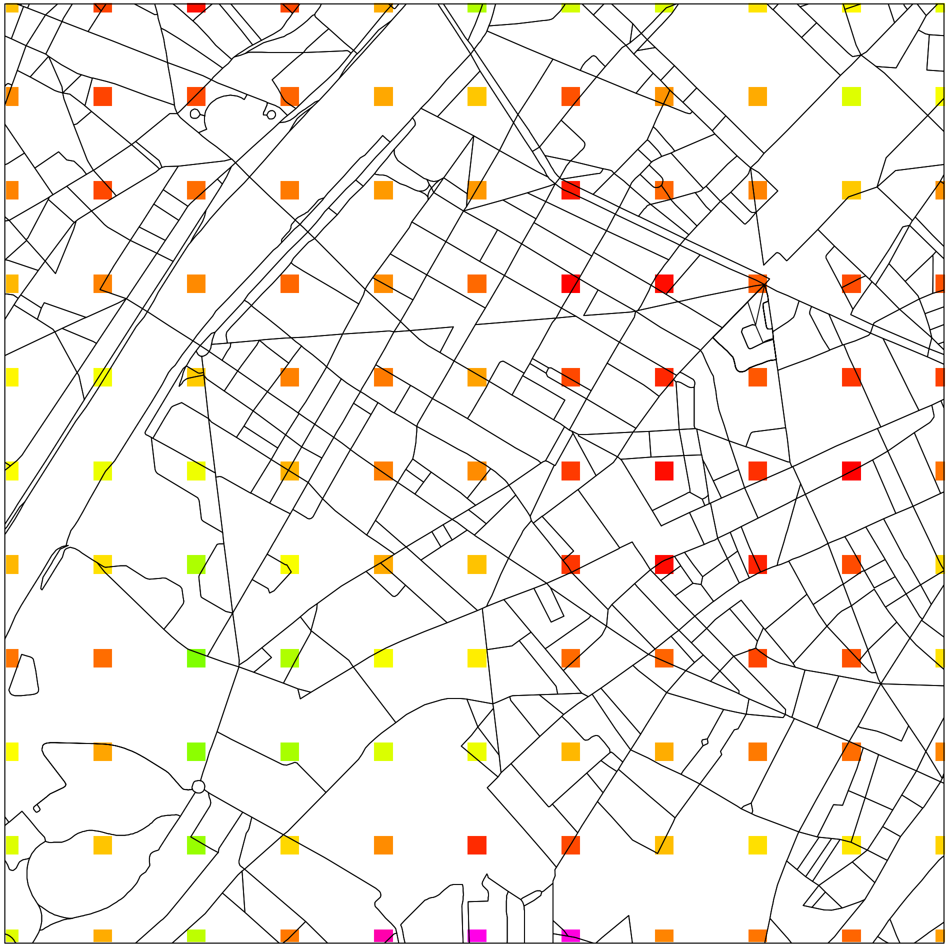
Data of Paris

Extrapolation Techniques
Motivation

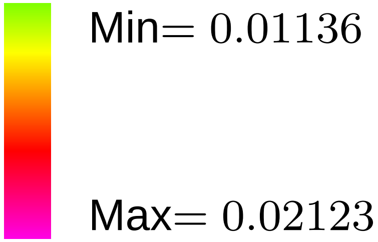


Data of Paris

Nonstatistical Extrapolation
Measurement Points



Measurement Points (PLT)



Nonstatistical Extrapolation

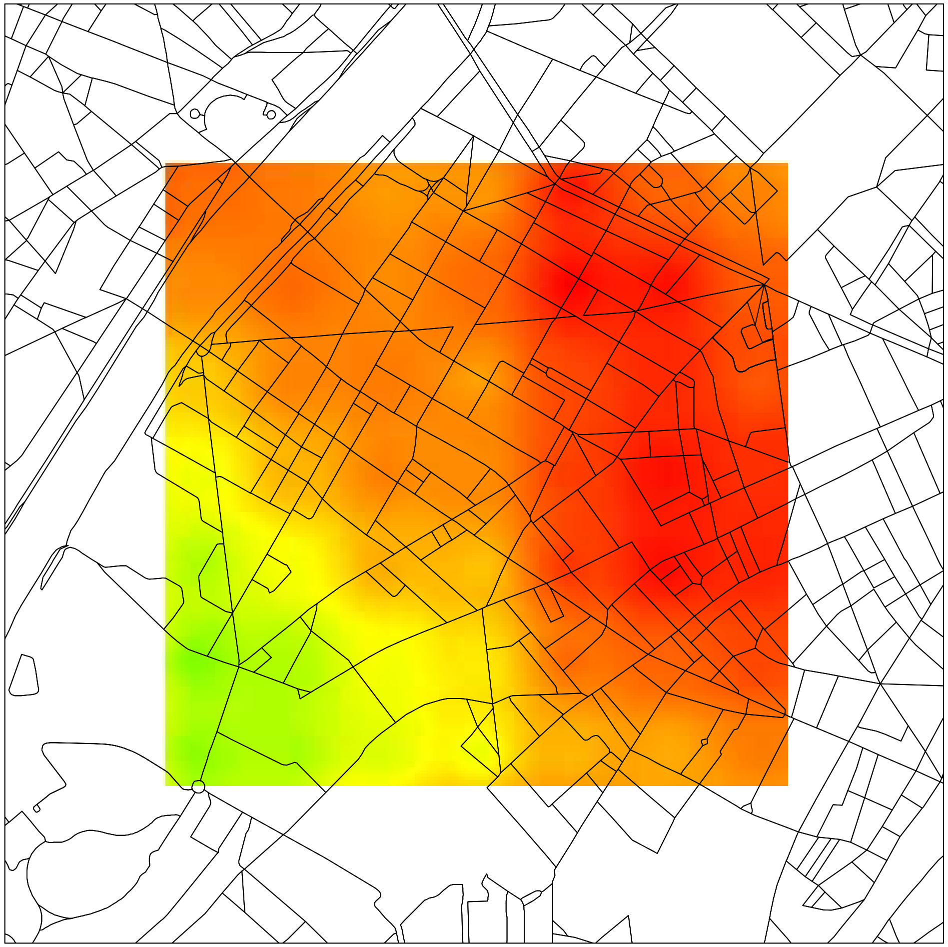
Inverse distance extrapolation

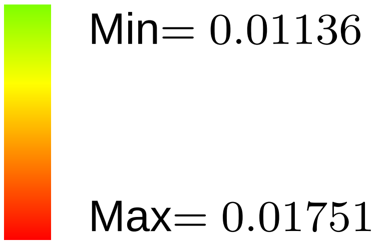
$$\widehat{Z}(x_0) = \sum_{i=1}^{n} w_i(x_0) Z(x_i) \qquad x_i \text{ Sampling Points}$$

where $w_i(x_0) = \frac{1}{h_i^p}\left(\sum_{j=1}^{n} \frac{1}{h_j^p}\right)^{-1}$ with

$h_i = \|x_i - x_0\| \;\; i = 1, \ldots, n, p \in \mathbb{N}$

Nonstatistical Extrapolation
Intensity Map

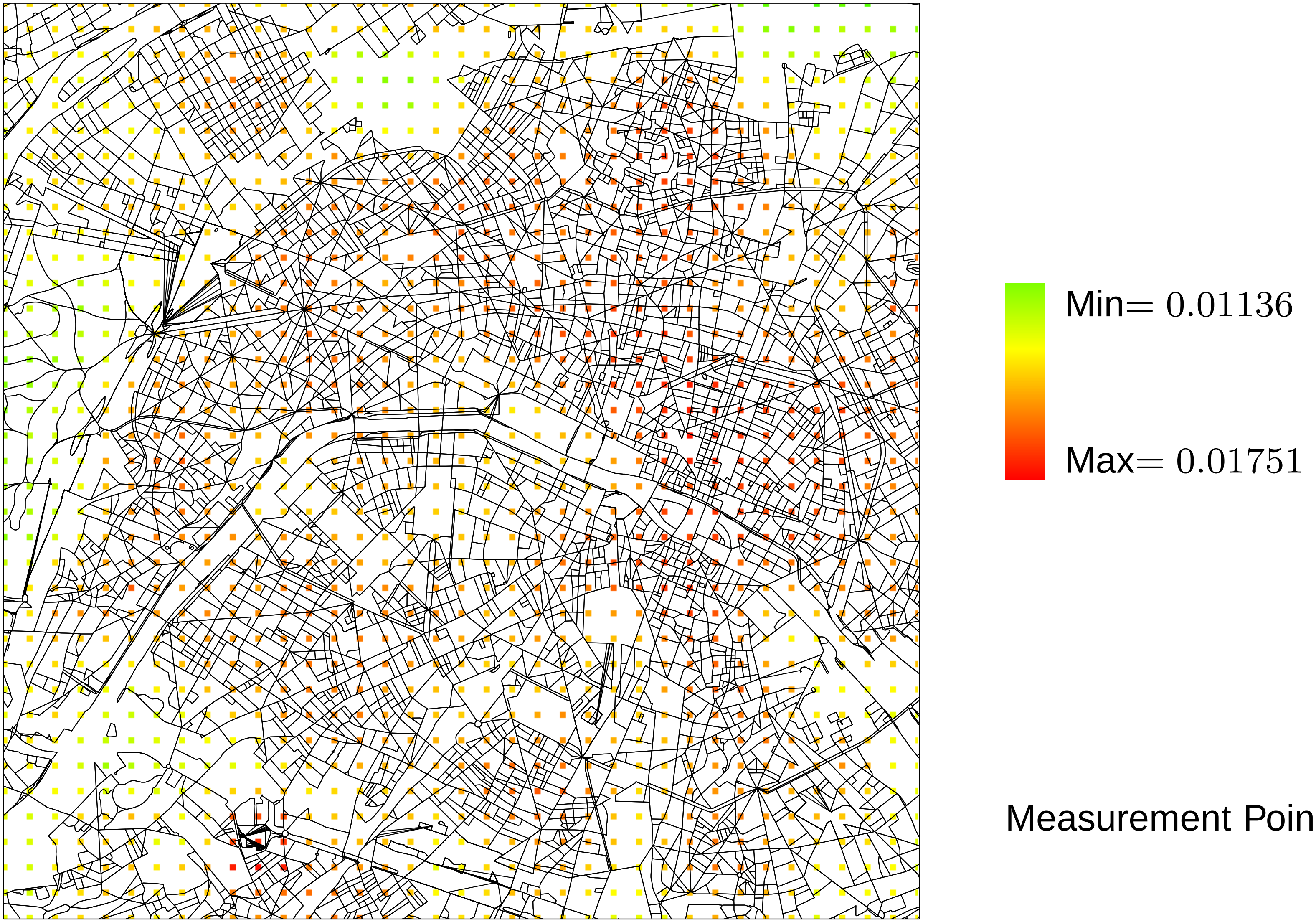




Real data of Paris and Intensity Map (PLT)

Statistical Extrapolation - Kriging

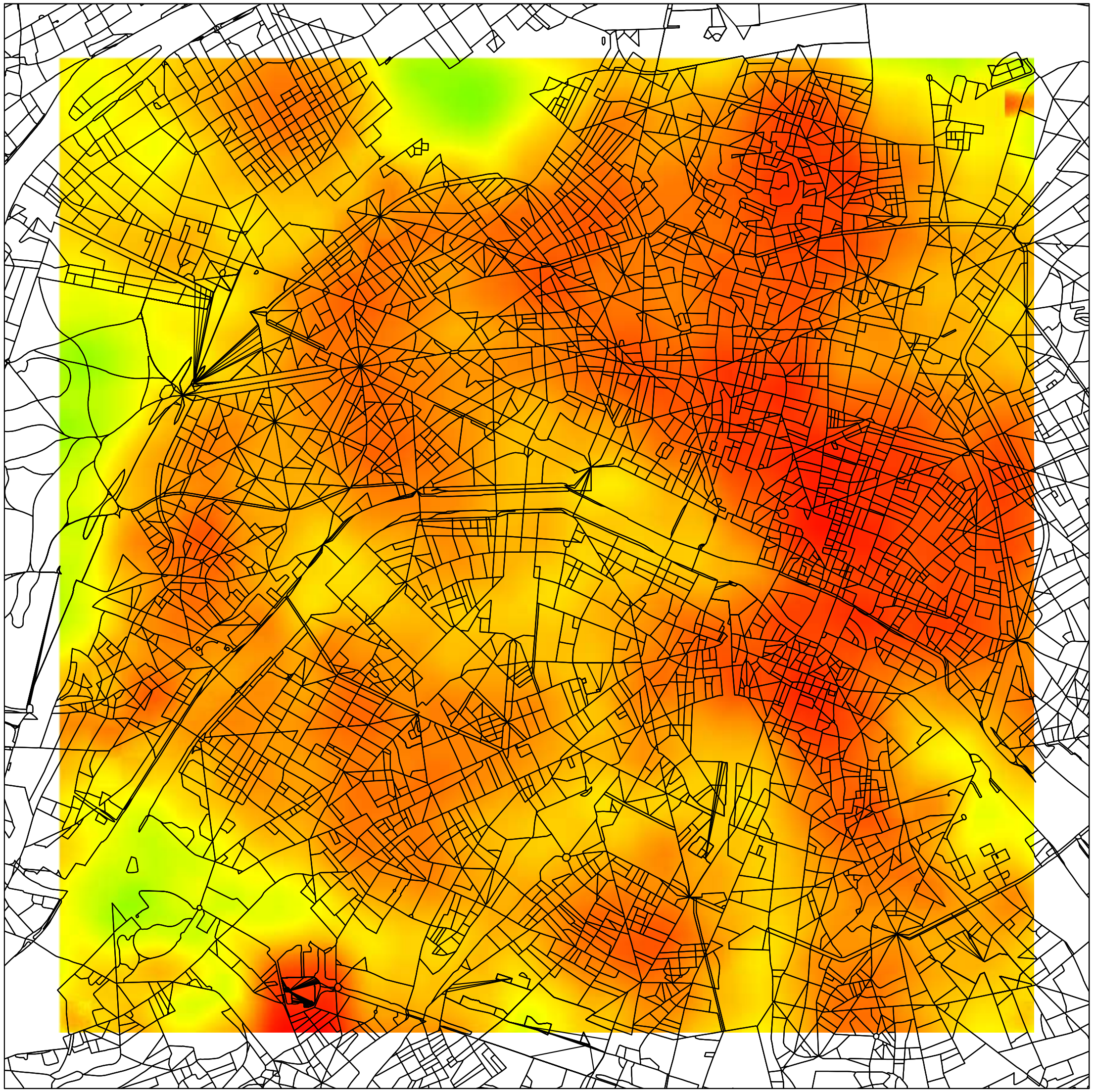
Intensity Map

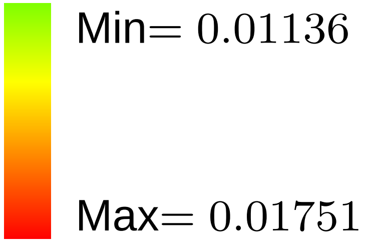


Measurement Points (PLT)

Statistical Extrapolation - Kriging

Intensity Map





Real data of Paris and Intensity Map (PLT)

References

- Gloaguen, C., Fleischer, F., Schmidt, H., Schmidt, V. (2005), Fitting of Stochastic Telecommunication Network Models via Distance Measures and Monte-Carlo Tests. *Telecommunication Systems* (to appear).
- Hörner, S. (2004), *Modelling of Telecommunication Networks*. Diploma Thesis, University of Ulm.
- Stoer, J. (2004), *Optimierung*. Springer, Berlin.
- Thiedmann, R. (2006), *Optimization Methods for Statistical Fitting of Network Models and Application to Intensity Maps*. Diploma Thesis, University of Ulm.
- Wackernagel, H. (1998), *Multivariate Geostatistics*. Springer, Berlin.

Thank you for your attention!

Questions?