aass

the Learning Systems Lab

Örebro Universitet

Statistical Distribution Modelling with the Kernel DM+V Algorithm – Application to Oceanographic Data

Achim J. Lilienthal

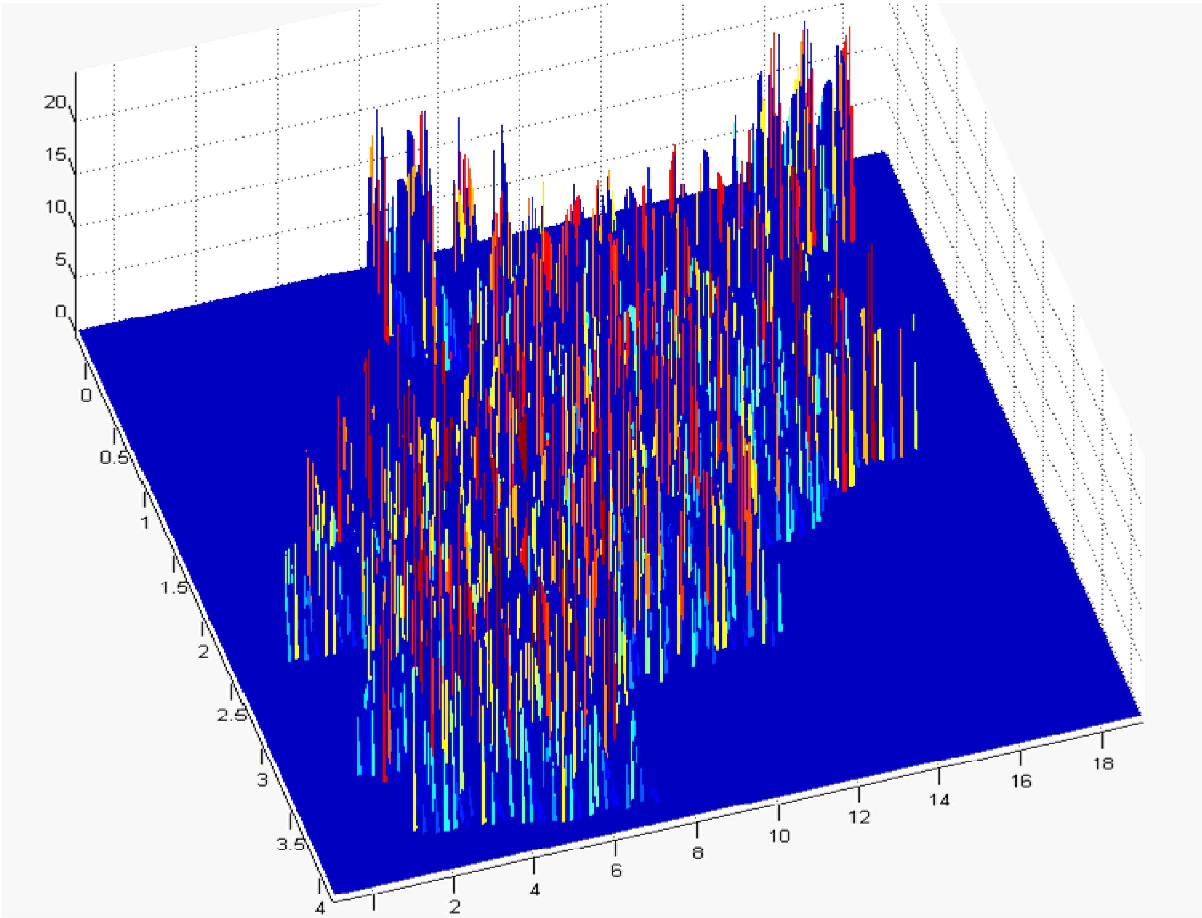
suggested by Andrea Caiti and with help from Alberto Alvarez

Agenda

Introduction

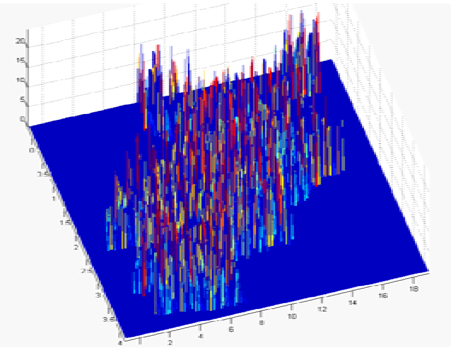
1 Statistical Distribution Modelling

- Input: Spatially Distributed Data

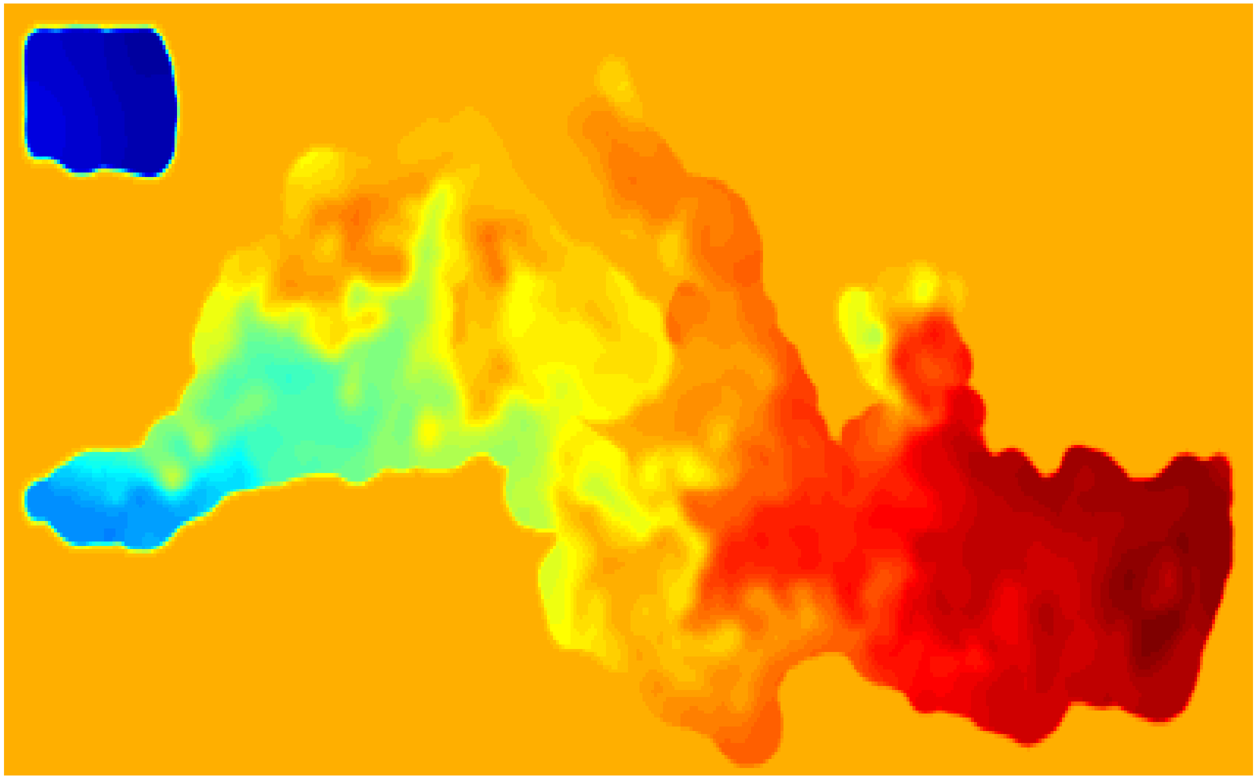


1 Statistical Distribution Modelling

- Input: Spatially Distributed Data

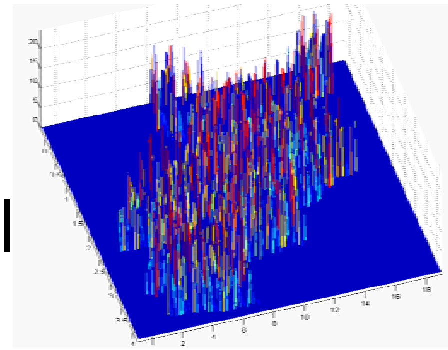


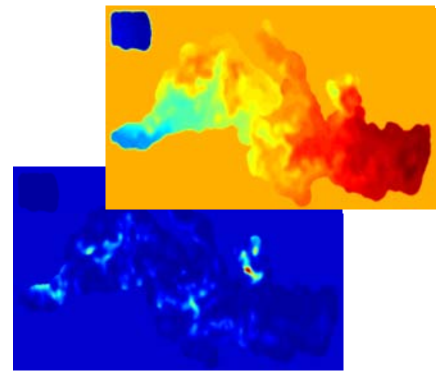
- → Output: Data-Driven, Statistical Model



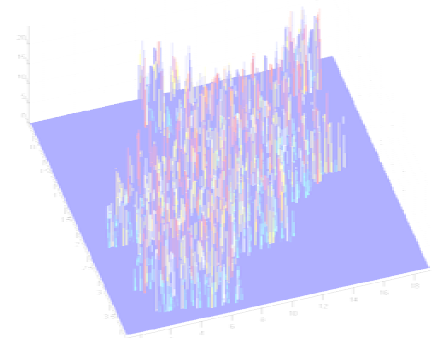
1 Statistical Distribution Modelling

- Input: Spatially Distributed Data
- → Output: Data-Driven, Statistical Model
  - provides a comprehensive view on the data
    - truthful representation of the distribution
  - allows to plan new measurement locations
    - where information density is low
    - where predictive concentration is high
    - where predictive variance is high

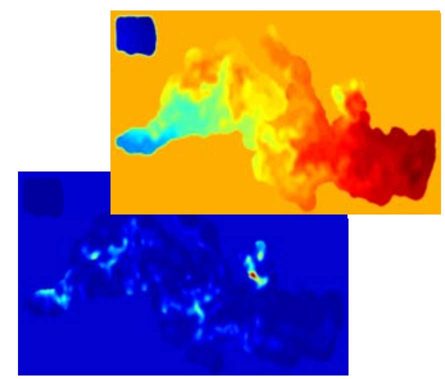




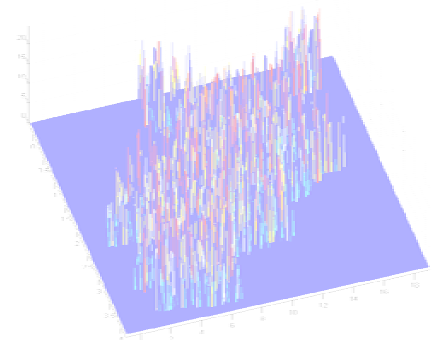
1 Statistical Distribution Modelling



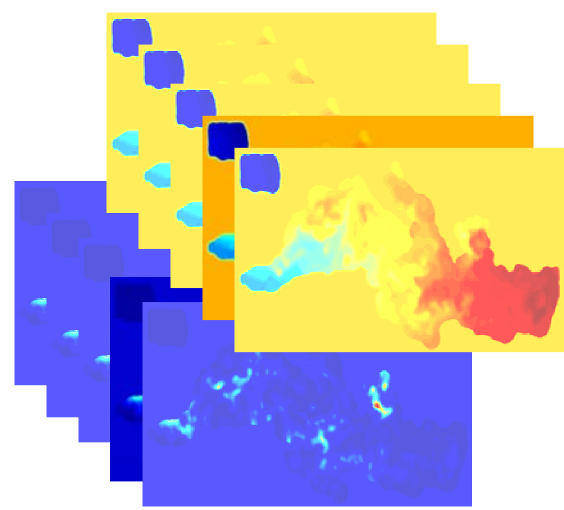
- Input: Spatially Distributed Data
- → Output: Data-Driven, Statistical Model
- Statistical Model (Kernel DM+V)
  - treat measurements as random variables
  - estimate generating random process by cross-validation optimization on the data
  - predicts mean and variance



1 Statistical Distribution Modelling



- Input: Spatially Distributed Data
- → Output: Data-Driven, Statistical Model
- Statistical Model (Kernel DM+V)
- Good Model? Truthful Representation?
  - allows to infer unseen measurements "explains observations and accurately predict new ones"
  - allows to infer hidden parameters



1 Statistical Distribution Modelling

- Problem Definition: Stat. Gas Distribution Modelling
  - learn predictive model

measurements

$$p\left(r_* \mid \vec{x}_*, \vec{x}_{1:n}, r_{1:n}\right)$$

prediction

query location

measurement locations

1 Statistical Distribution Modelling

- Problem Definition: Stat. Gas Distribution Modelling
  - learn predictive model

$$p\left(r_* \mid \vec{x}_*, \vec{x}_{1:n}, r_{1:n}\right)$$

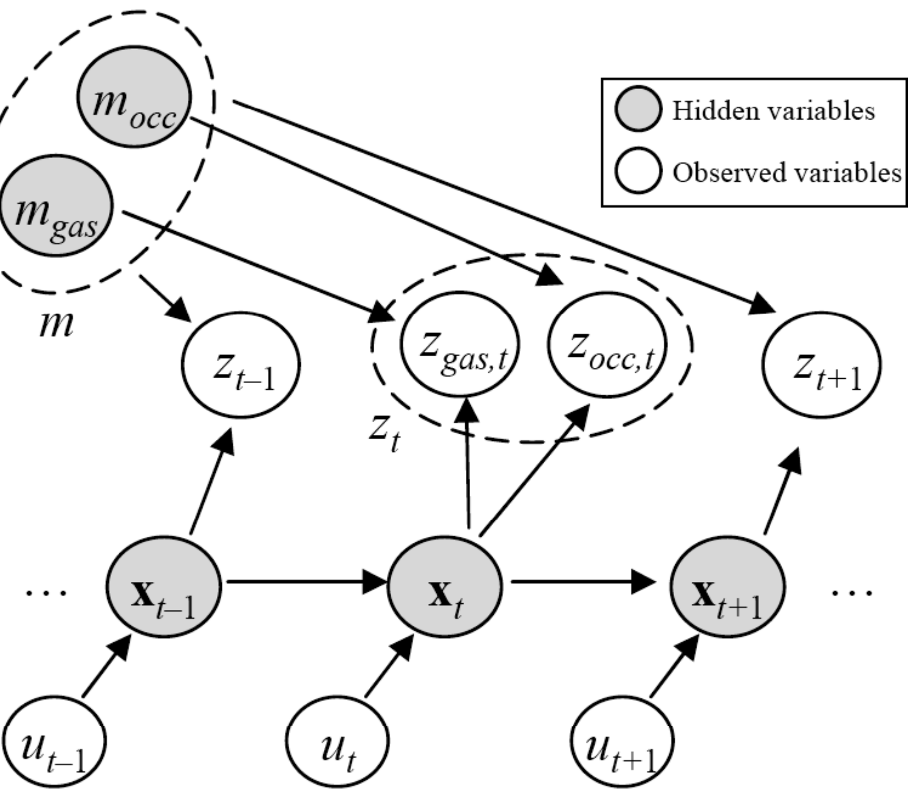
  - or: estimate posterior over gas distribution models

$$p\left(\mathbf{m} \mid \vec{x}_{1:n}, r_{1:n}\right)$$

1 Statistical Distribution Modelling

- Problem Definition: Stat. Gas Distribution Modelling
  - estimate posterior probability over gas distribution models

Bayesian network of the GDM/SLAM problem



Kernel DM+V for Spatial Distribution Modelling

2 Kernel DM+V

- Kernel Distribution Mapping+V (Kernel DM+V)
  - predictive uncertainty (predictive mean + variance)

➡ *A Statistical Approach to Gas Distribution Modelling with Mobile Robots – The Kernel DM+V Algorithm,* Achim J. Lilienthal, Matteo Reggente, Marco Trincavelli, Jose Luis Blanco and Javier Gonzalez. IROS, 2009, pp. 570-576.

2 Kernel DM+V

- Example: Gas Data Collected with Mobile Robot

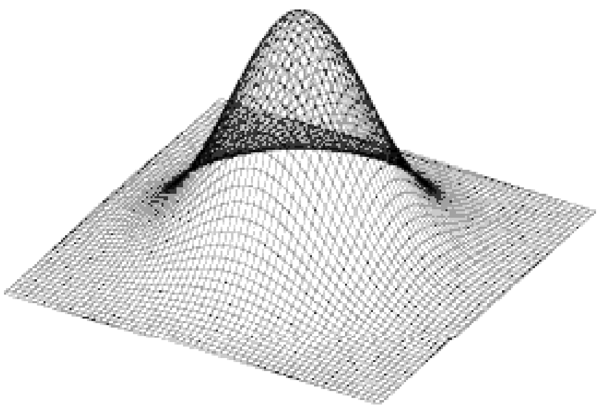
Kernel DM+V

- integrated weights, integrated readings

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i} \qquad R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i$$

$$\omega_{k,i} = Gauss\left(\left|\vec{x}_i - \vec{x}_k(c)\right|, \sigma\right)$$

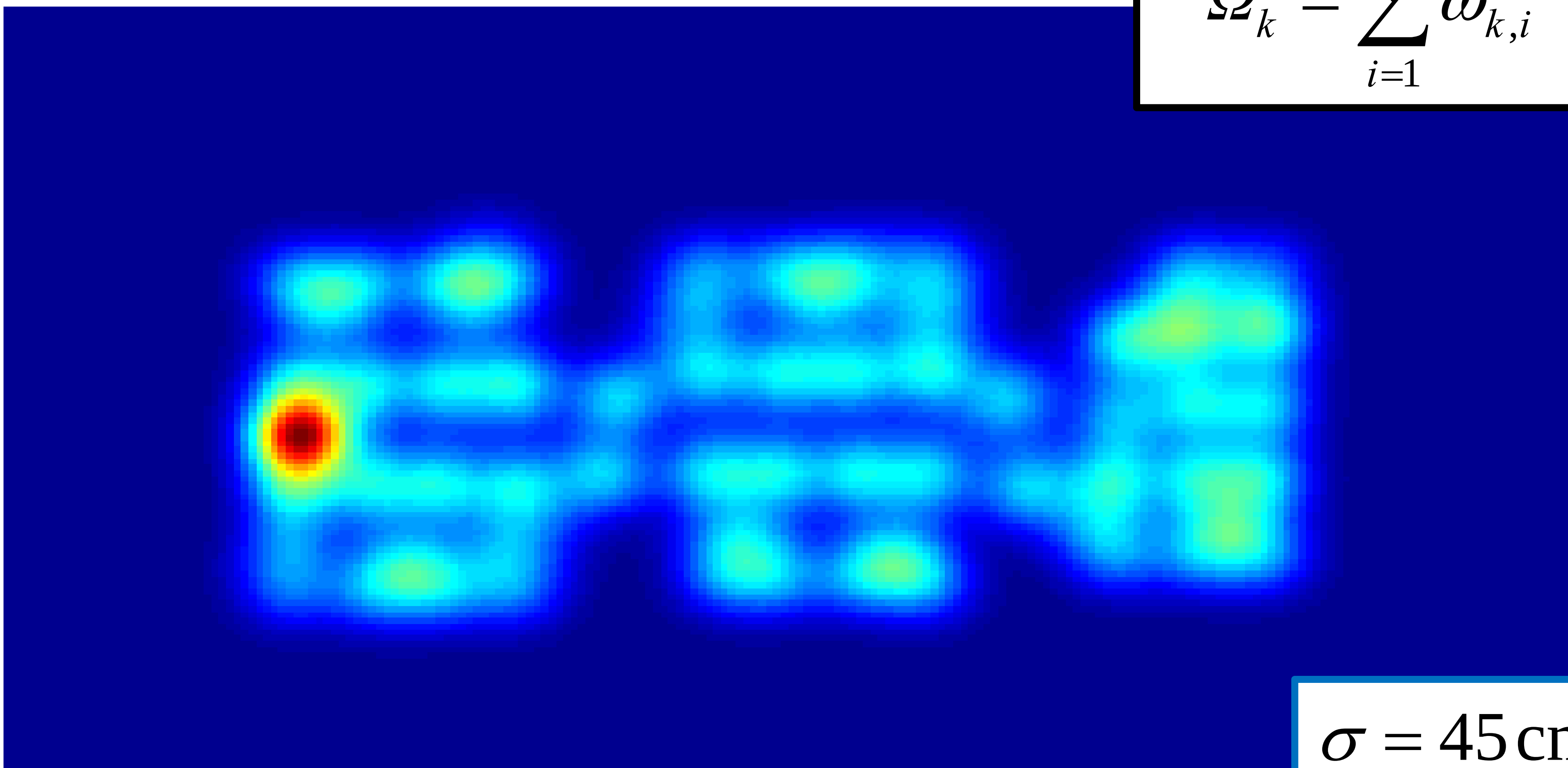
  - models the spatial information content of r_i



2 Kernel DM+V

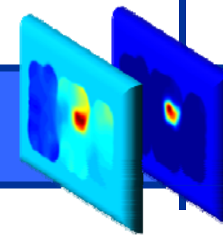
- Kernel DM+V
  - integrated weights

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i}$$



$$\sigma = 45\,\mathrm{cm}$$

2 Kernel DM+V



- Kernel DM+V
  - integrated weights, integrated readings

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i} \qquad R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i$$

  - confidence map

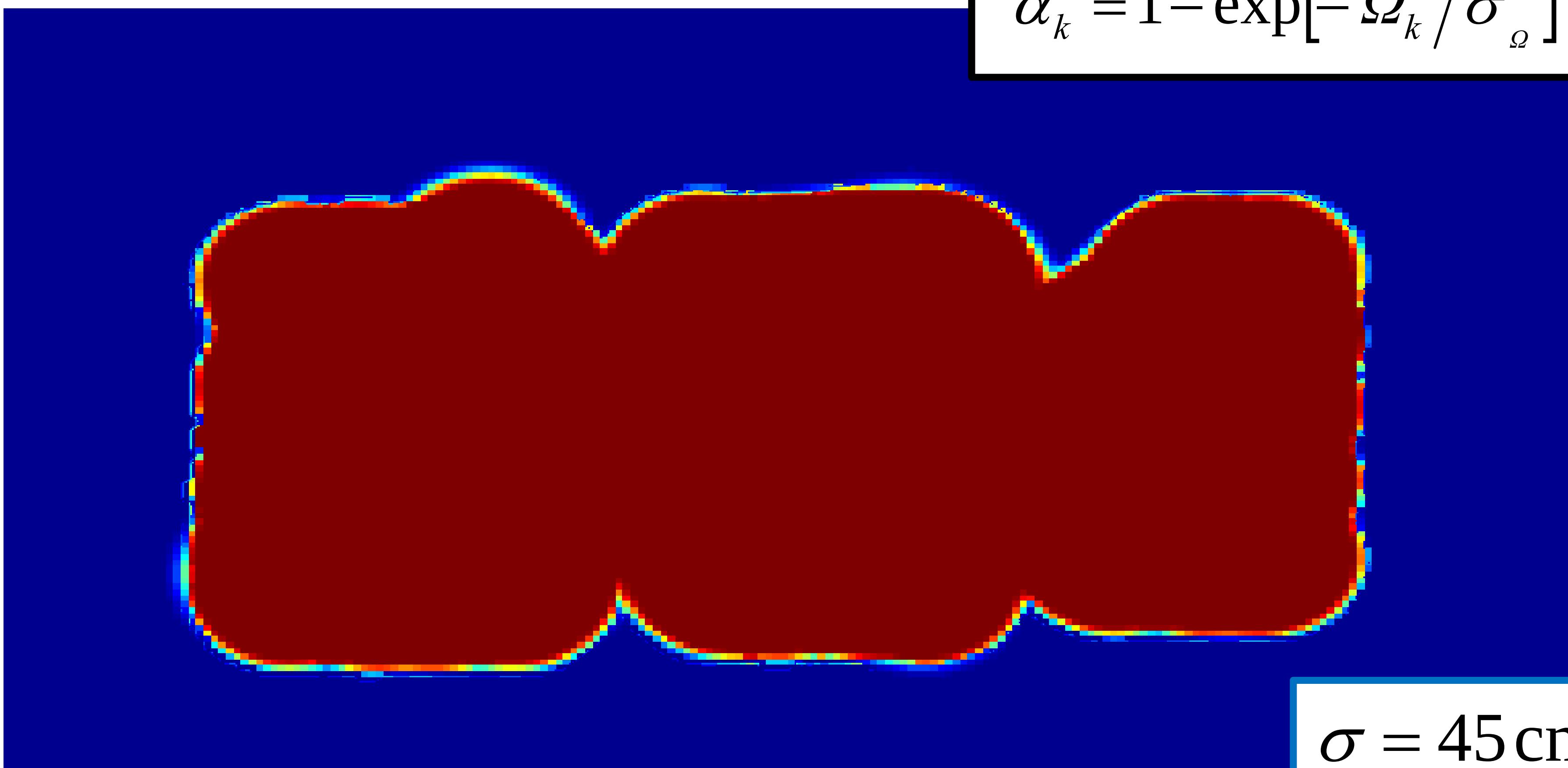
$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2 \right]$$

    - for high values of Ω_k: $\alpha_k \to 1$
    - for high values of $\Omega_k = 0$: $\alpha_k = 0$
    - "high" and "low" relative to σ_Ω

2 Kernel DM+V

- Kernel DM+V
  - confidence map

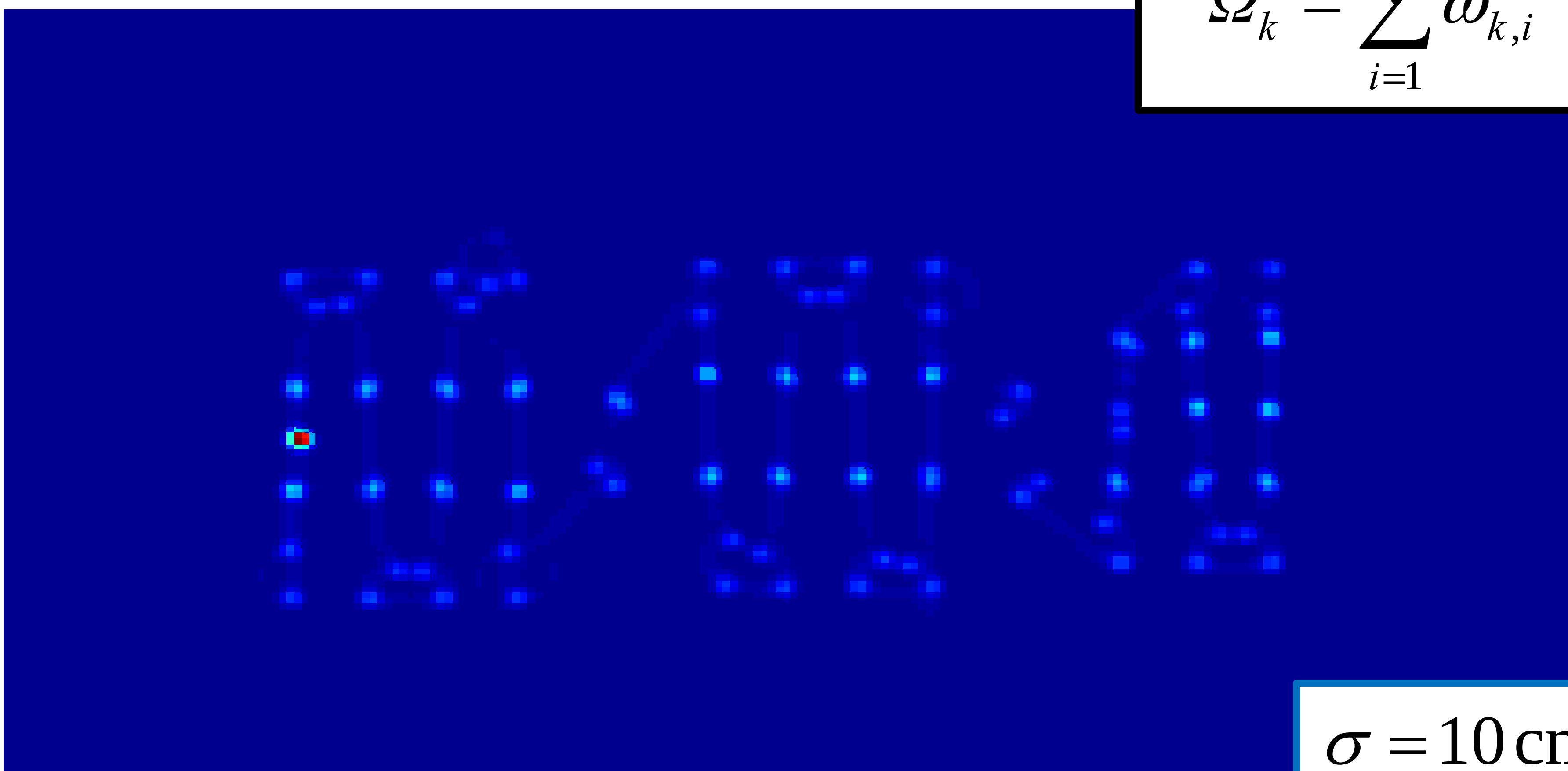
$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$



$$\sigma = 45\,\mathrm{cm}$$

- Kernel DM+V
  - integrated weights

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i}$$

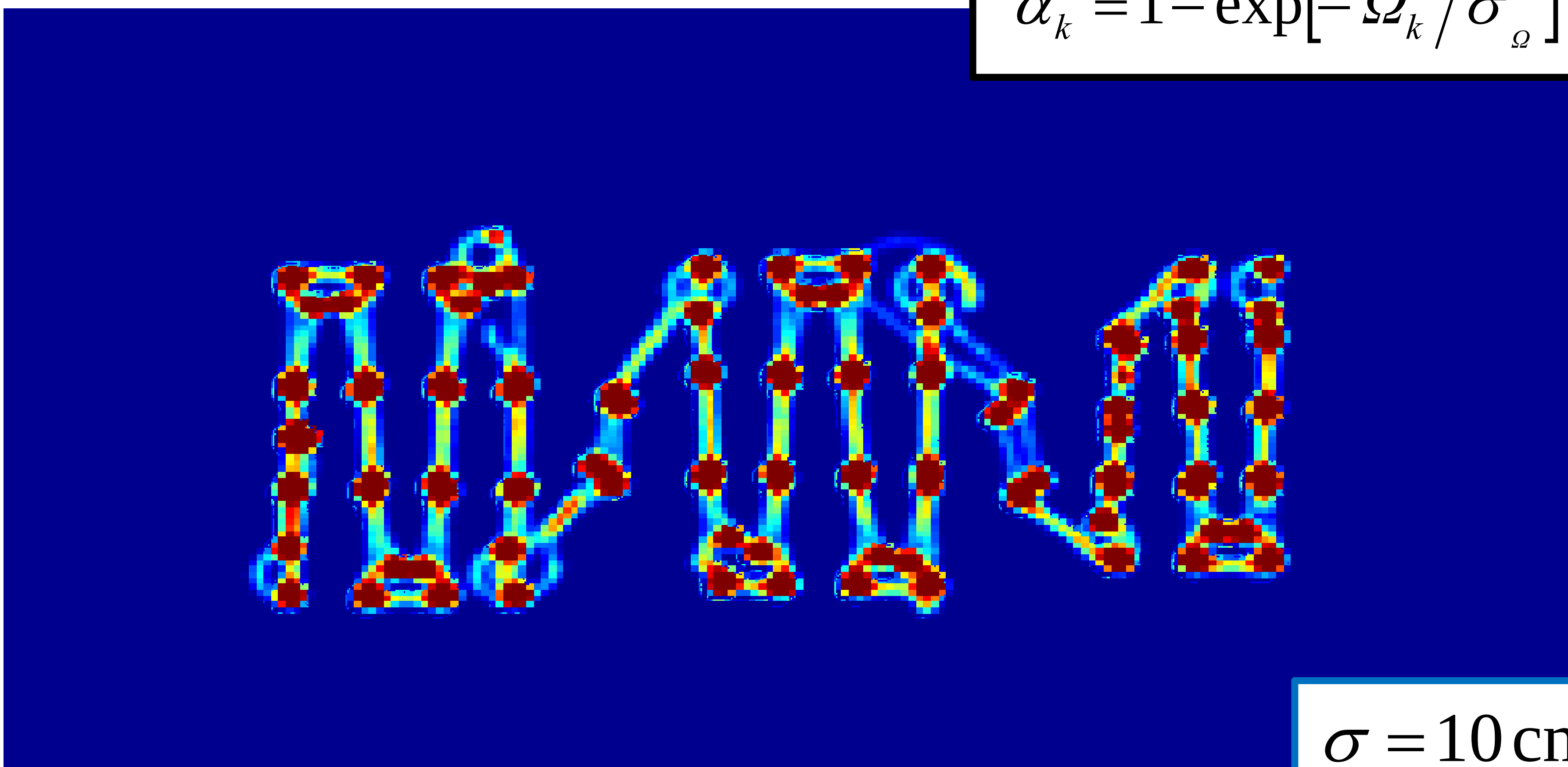


$$\sigma = 10\,\mathrm{cm}$$

2 Kernel DM+V

- Kernel DM+V
  - confidence map

$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$



$$\sigma = 10\,\mathrm{cm}$$

- Kernel DM+V
  - integrated weights, integrated readings

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i} \qquad R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i$$

  - confidence map

$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$

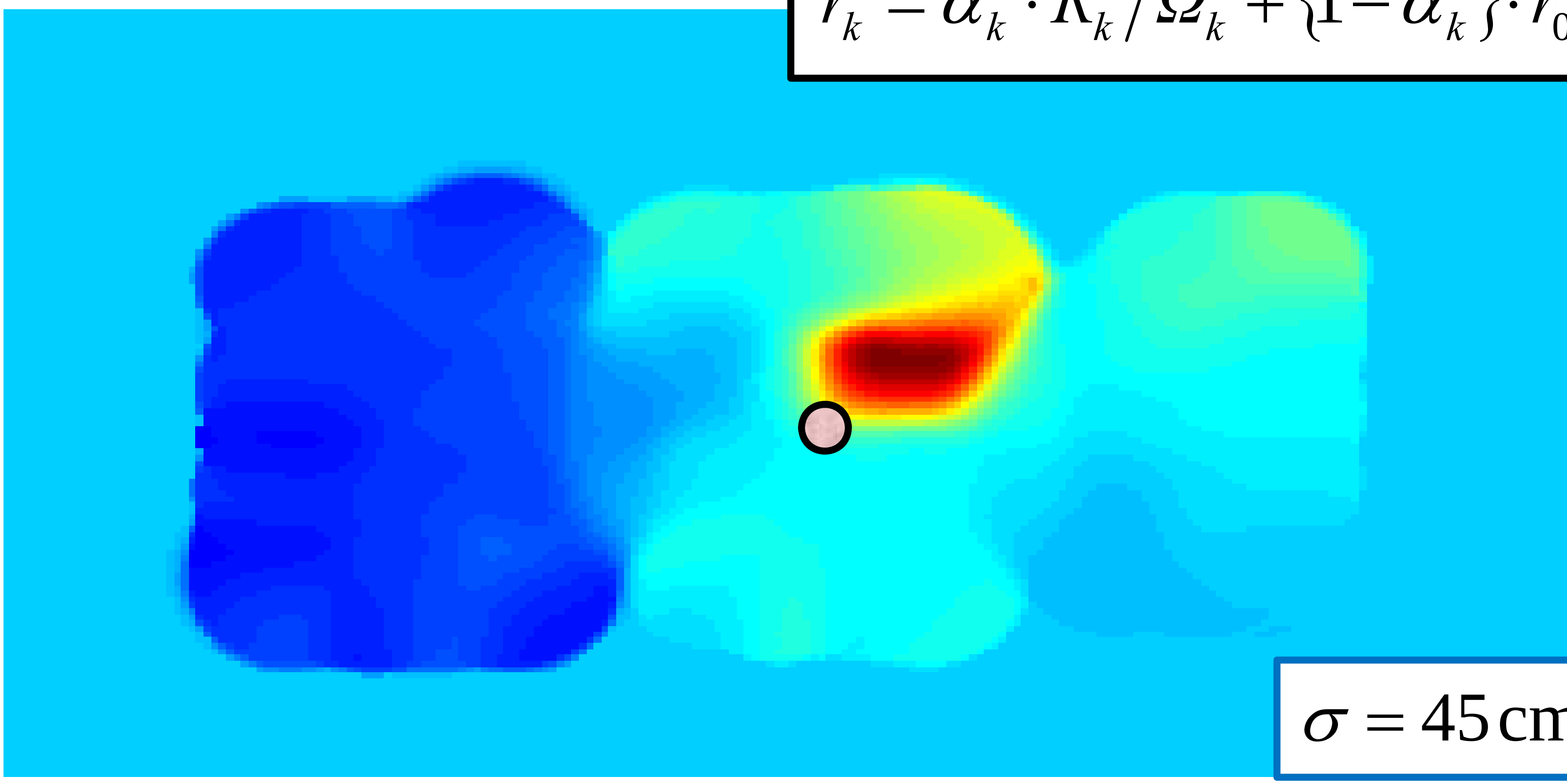
  - predictive mean

$$r_k = \alpha_k \cdot R_k / \Omega_k + \{1 - \alpha_k\} \cdot r_0 \qquad r_0 = 1 \big/ |D| \sum_{i=1}^{|D|} r_i$$

Kernel DM+V – Example

- predictive mean

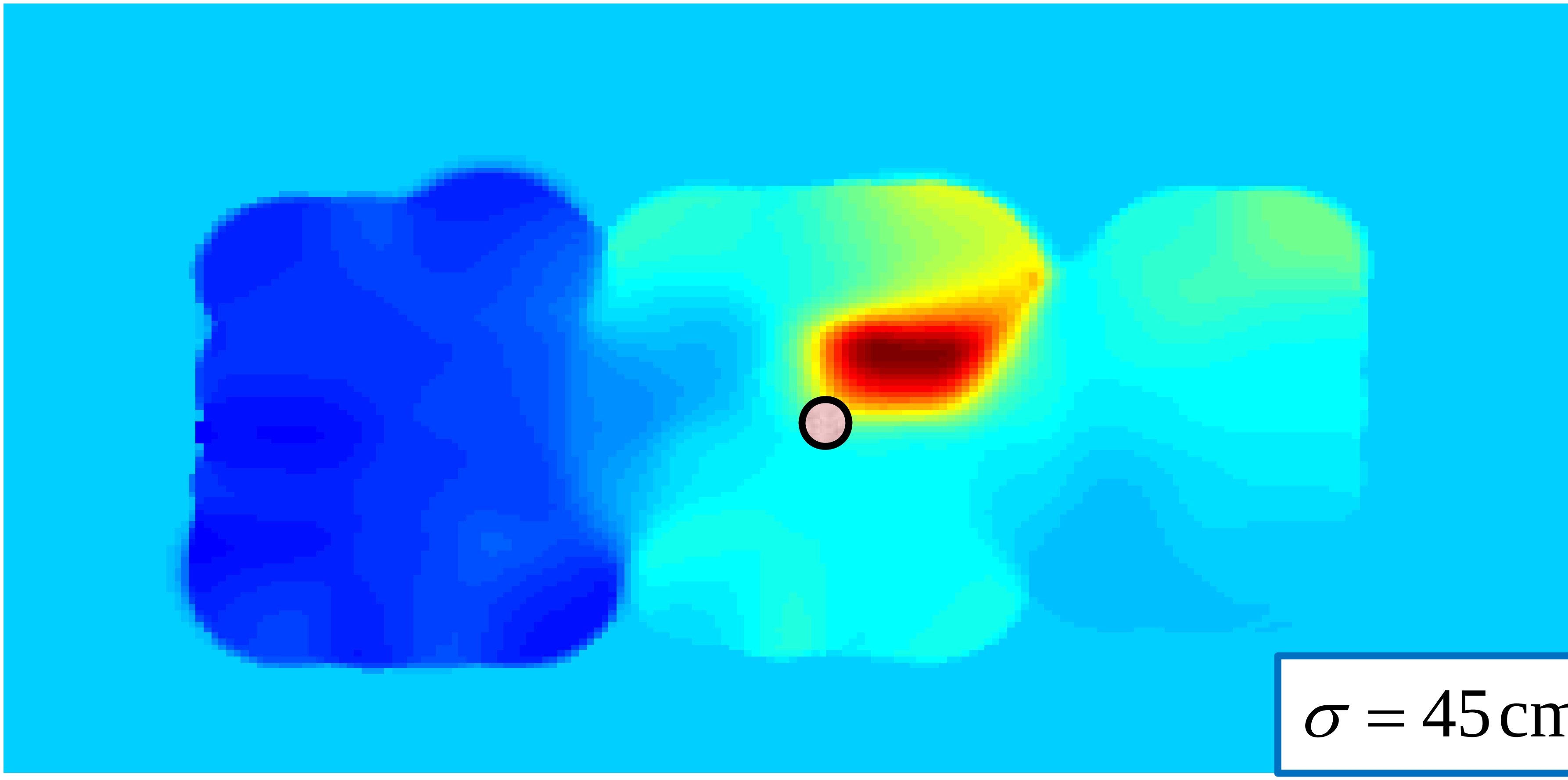
$$r_k = \alpha_k \cdot R_k / \Omega_k + \{1 - \alpha_k\} \cdot r_0$$



$$\sigma = 45\,\text{cm}$$

2 Kernel DM+V – Remarks

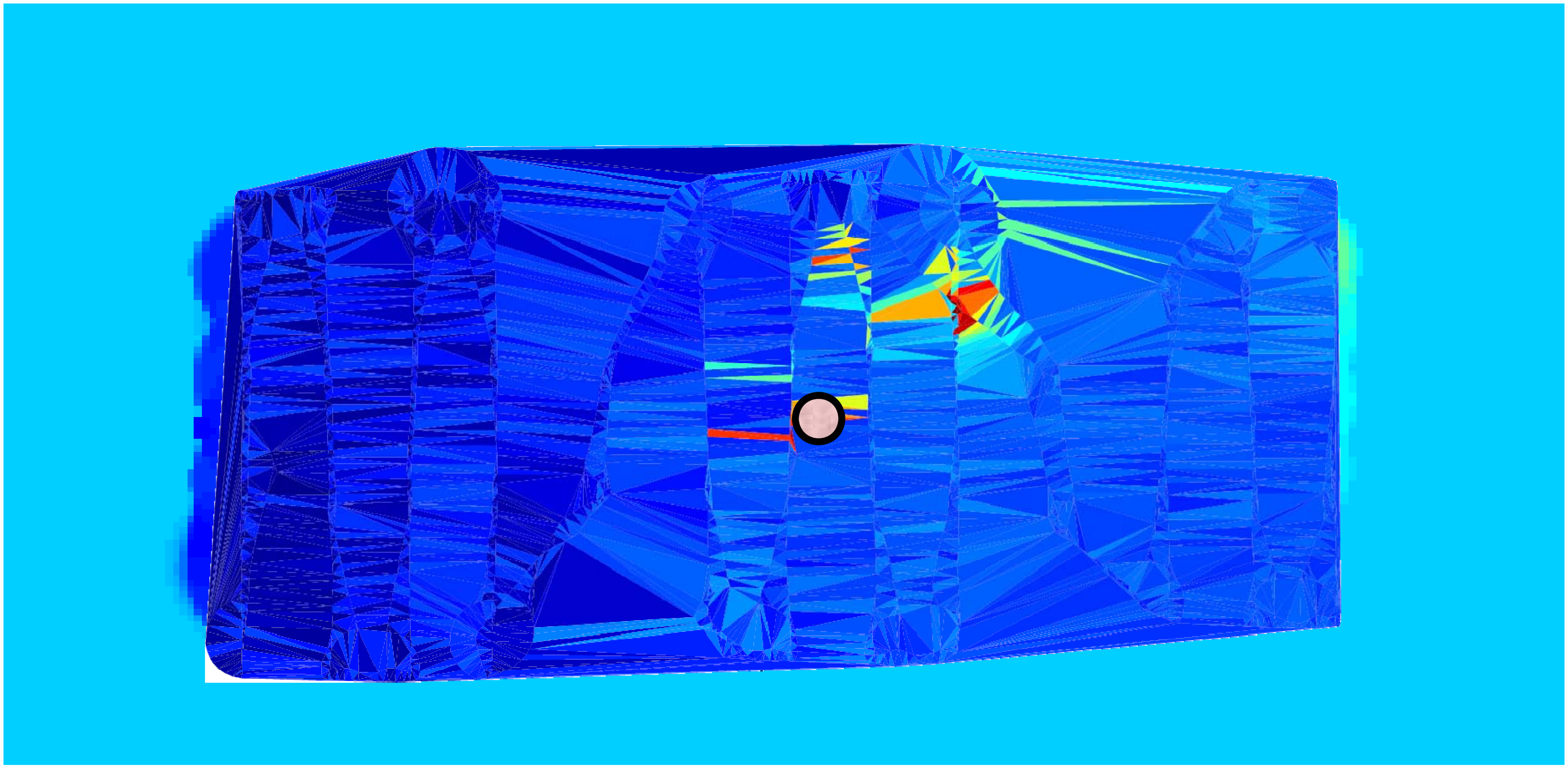
- Kernel DM+V
  - comparison with interpolation map



$$\sigma = 45\,\mathrm{cm}$$

2 Kernel DM+V – Remarks

- Comparison with Map from Trilinear Interpolation
  - MATLAB function `trisurf`



Kernel DM+V

- integrated weights, integrated readings

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i} \qquad R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i$$

- confidence map

$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$

- predictive variance estimated separately
- variance contributions for each measurement

$$\tau_i = \left(r_i - r_{k(i)}\right)^2$$

$k(i)$ = cell closest to x_i

2 Kernel DM+V

- Kernel DM+V
  - integrated weights, integrated readings

$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i} \qquad R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i$$

  - confidence map

$$\alpha_k = 1 - \exp\left[-\Omega_k^2 \big/ \sigma_\Omega^2 \right]$$

  - predictive variance estimated separately

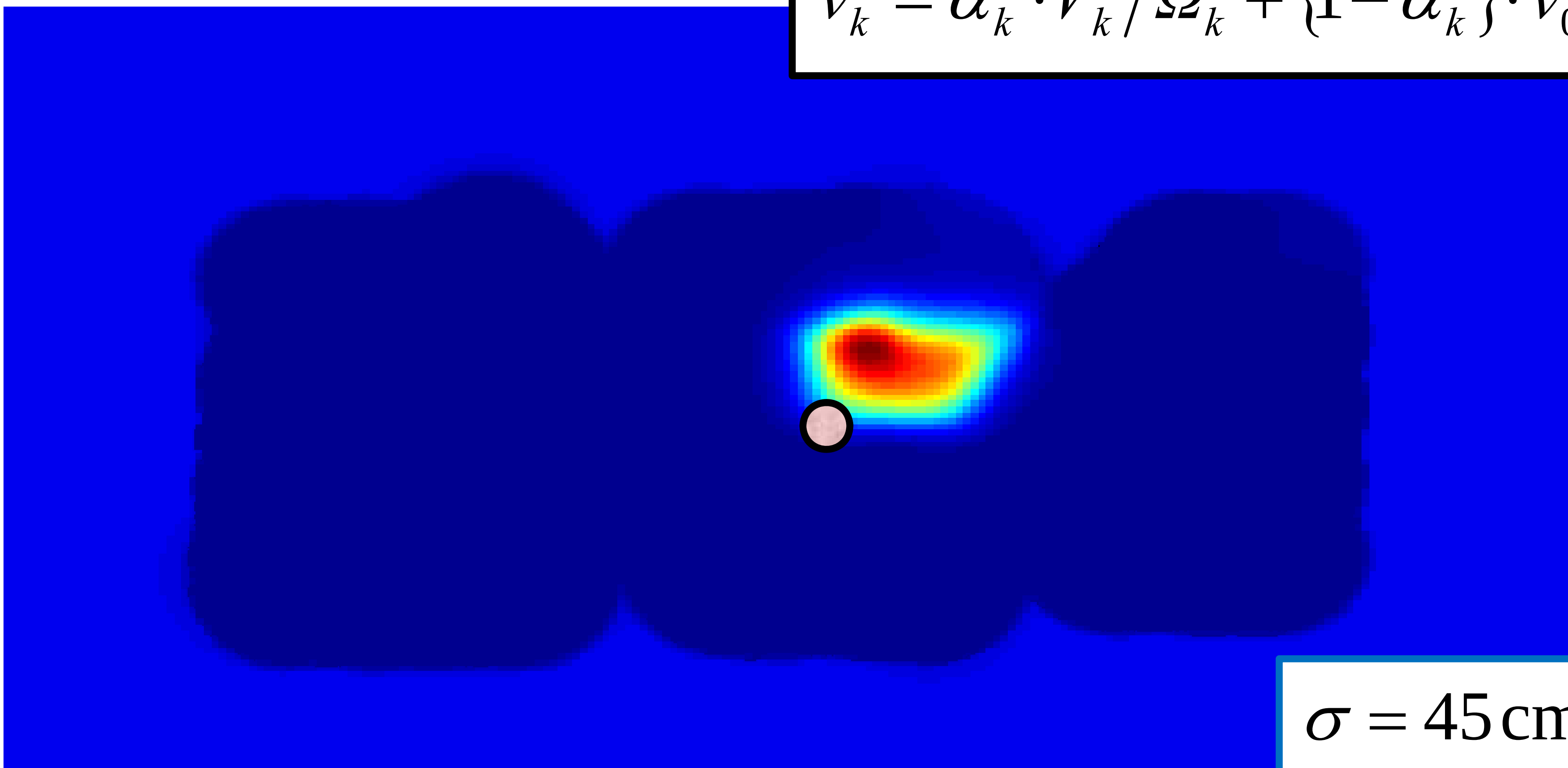
$$v_k = \alpha_k \cdot V_k \big/ \Omega_k + \{1 - \alpha_k\} \cdot v_0$$

$$V_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot \tau_i \qquad v_o = 1 \big/ |D| \sum_{i=1}^{|D|} \tau_i$$

Kernel DM+V – Example

- predictive variance

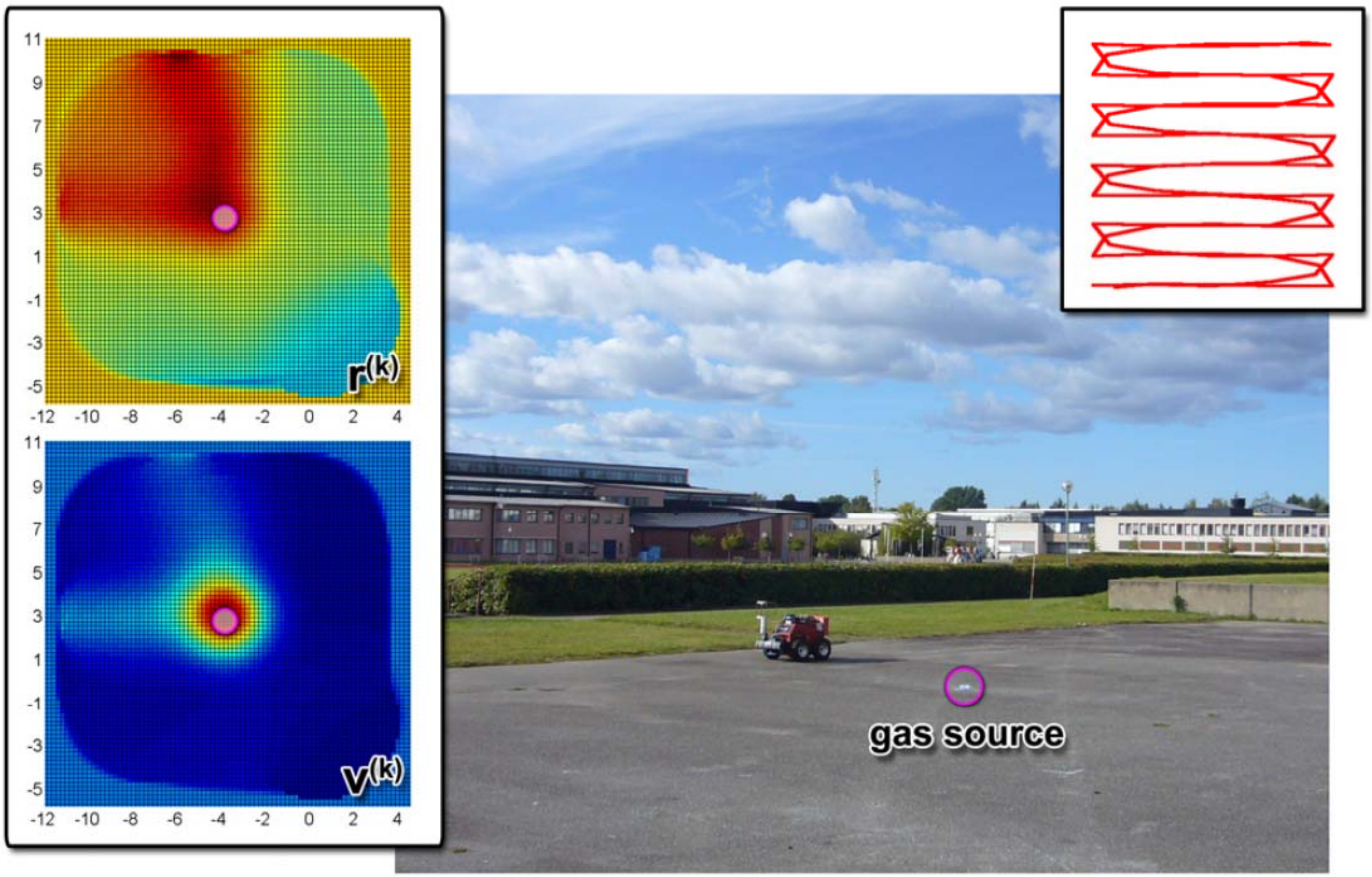
$$v_k = \alpha_k \cdot V_k / \Omega_k + \{1 - \alpha_k\} \cdot v_0$$



$$\sigma = 45\,\mathrm{cm}$$

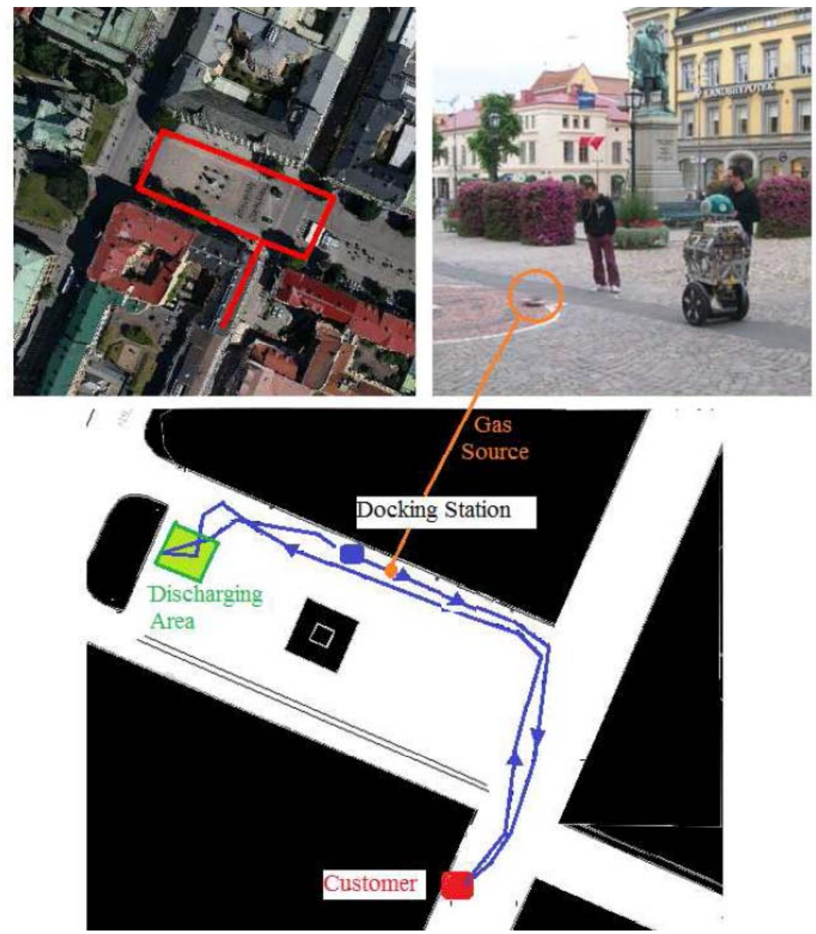
2 Kernel DM+V – GDM Experience

- Gas Distribution Modelling
  - outdoor experiments with mobile robot



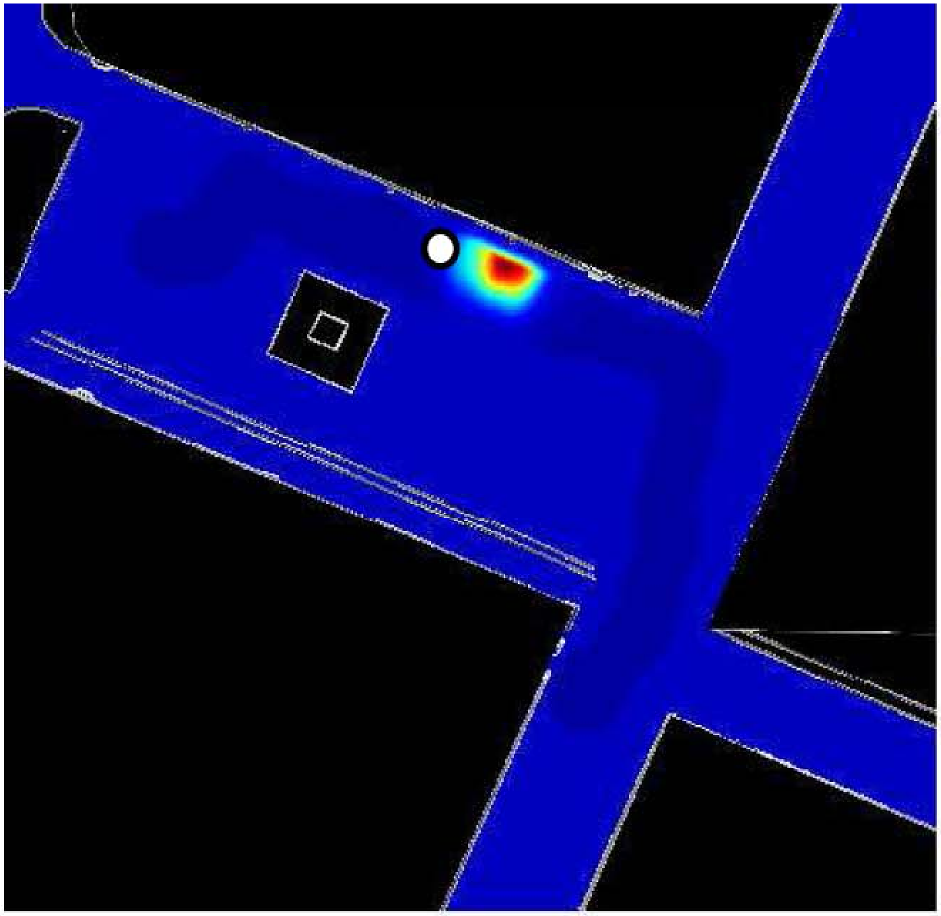
2 Kernel DM+V – GDM Experience

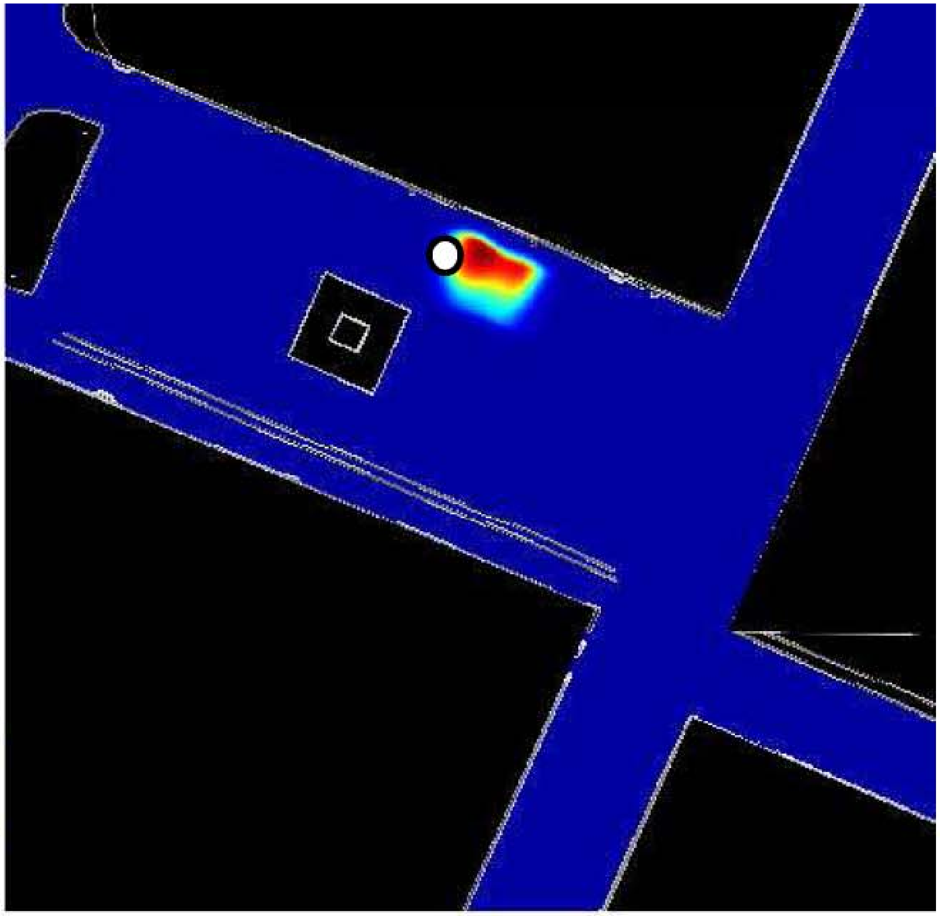
- Gas Distribution Modelling
  - outdoor experiments with mobile robot (DustBot)
    - DustTrak 8520 sensor (PM10)



2 Kernel DM+V – GDM Experience

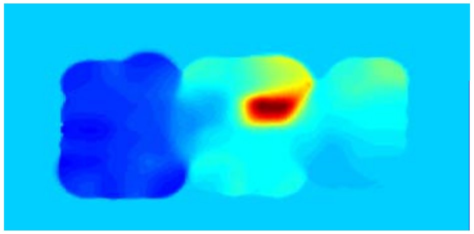
- Gas Distribution Modelling
  - outdoor experiments with mobile robot (DustBot)



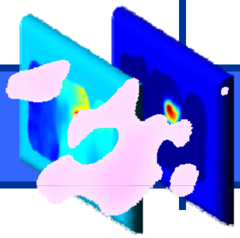


2 Kernel DM+V

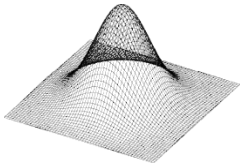
- Experiments in the "Microscope Room"



2 Kernel DM+V

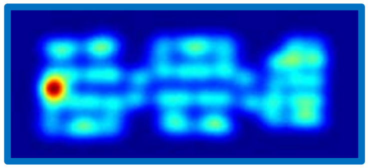


Kernel DM+V – Summary

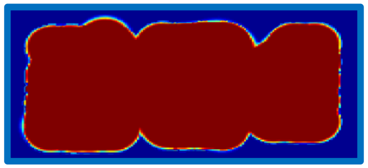


$$\omega_{k,i} = Gauss\left(\left|\vec{x}_i - \vec{x}_k\right|, \sigma\right)$$

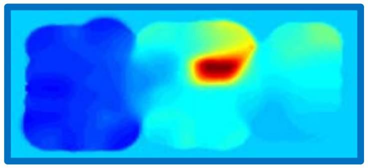
$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i}$$



$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$

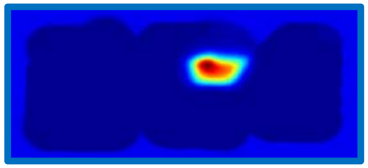


$$r_k = \alpha_k \cdot R_k / \Omega_k + \{1 - \alpha_k\} \cdot r_0$$



$$R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i \qquad r_0 = \frac{1}{|D|} \sum_{i=1}^{|D|} r_i$$

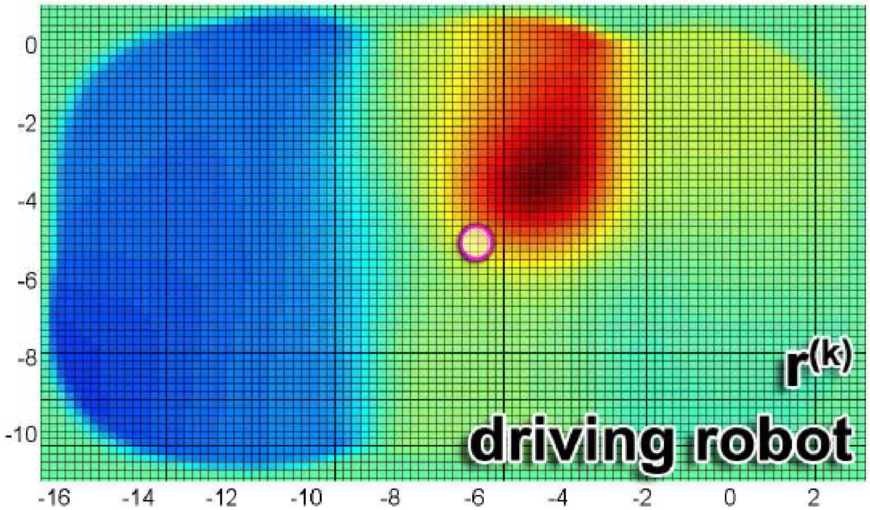
$$v_k = \alpha_k \cdot V_k / \Omega_k + \{1 - \alpha_k\} \cdot v_0$$

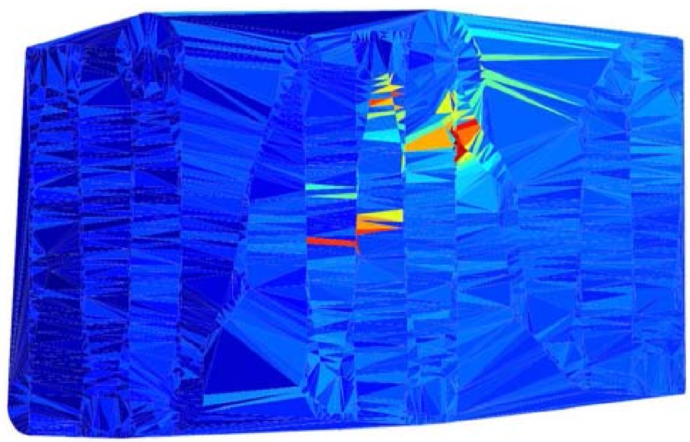


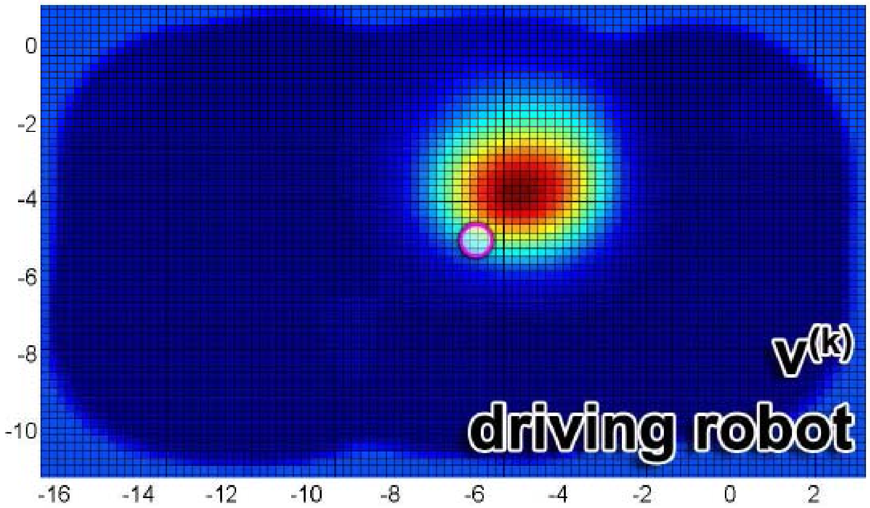
$$\tau_i = \left(r_i - r_{k(i)}\right)^2 \quad V_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot \tau_i \qquad v_o = \frac{1}{|D|} \sum_{i=1}^{|D|} \tau_i$$

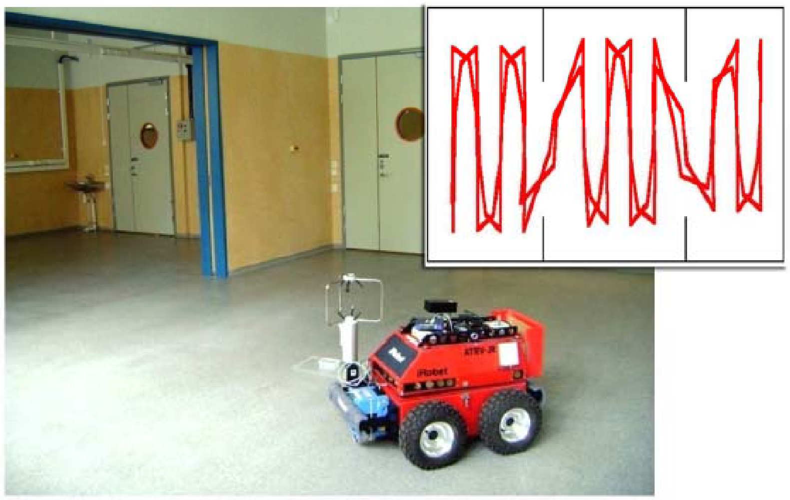
2 Kernel DM+V – Remarks

- Model Consistency

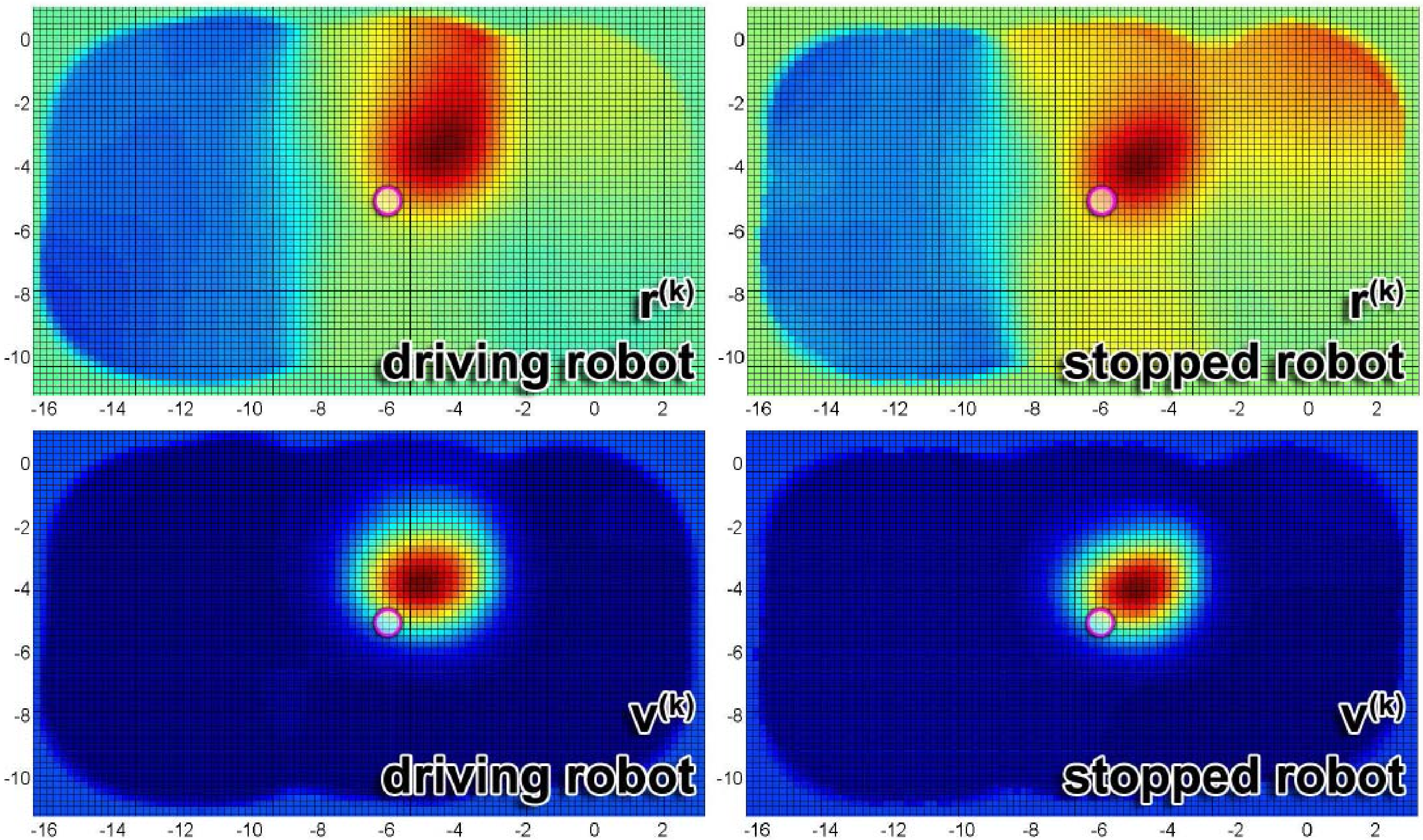








- Model Consistency



2 Kernel DM+V – Remarks

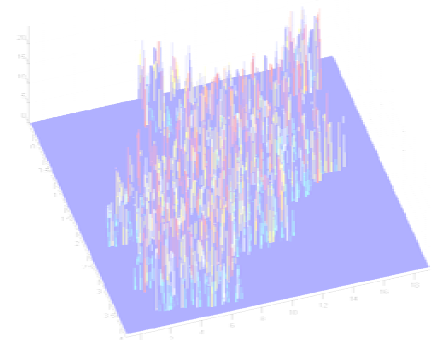
- Parameter Selection
  - 3 hyper-parameters: c, σ, σ_Ω
  - cross-validation to find optimal hyper-parameters
- Complexity (given σ)
  - O[training points $\times$ $(\sigma/c)^2 + (L/c)^2$]
- Availability
  - Matlab implementation (available on request)
  - C++ implementation: Ubuntu/MRPT (→ Jose Luis Blanco)

2 Kernel DM+V – Remarks

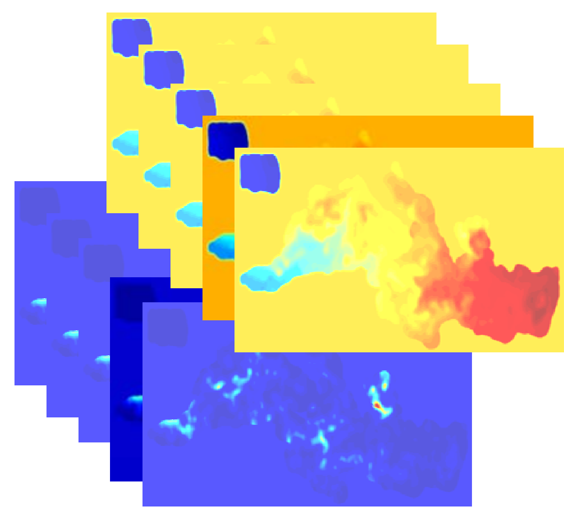
- Parameter Selection
  - 3 hyper-parameters $c, \sigma, \sigma_{\Omega}$
  - cross-validation to find optimal hyper-parameters

1 Statistical Distribution Modelling

- Input: Spatially Distributed Data
- → Output: Data-Driven, Statistical Model



- Statistical Model (Kernel DM+V)
- Good Model? Truthful Representation?
  - allows to infer unseen measurements "explains observations and accurately predict new ones"
  - allows to infer hidden parameters

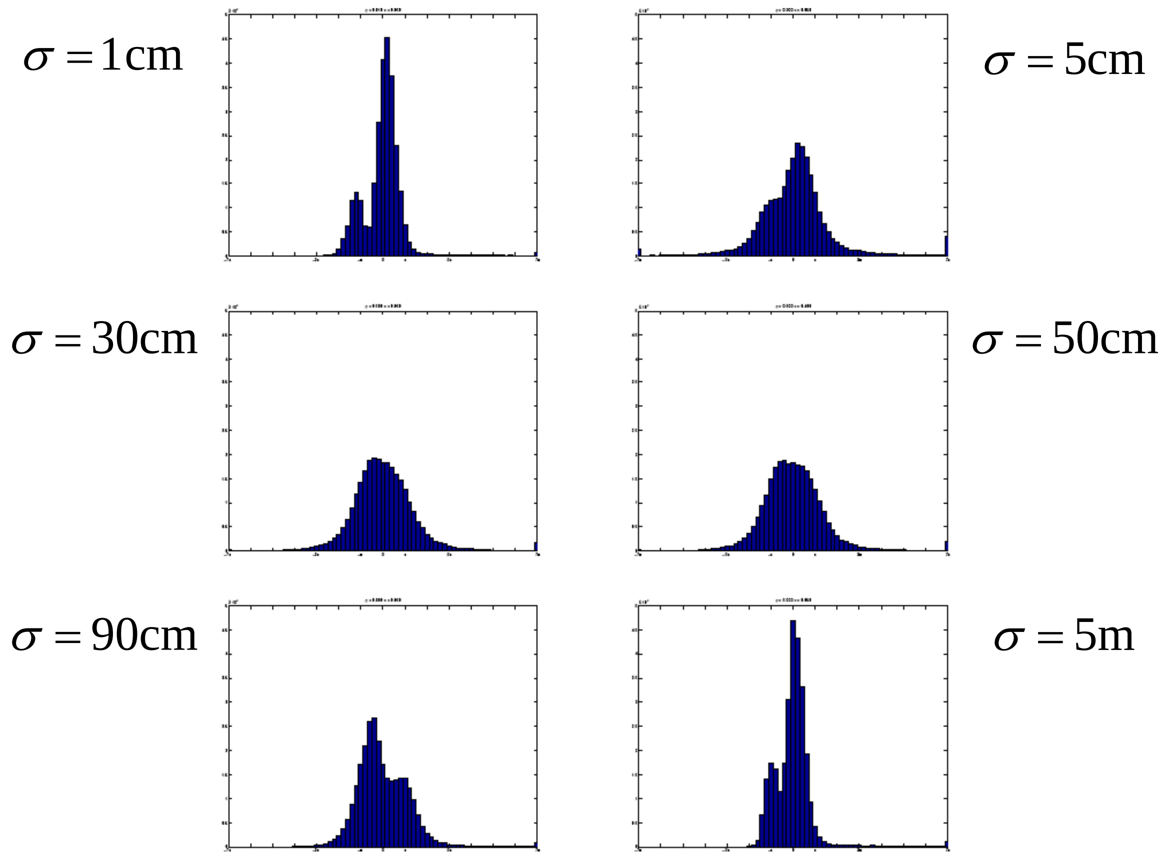


2 Kernel DM+V – Remarks

- Parameter Selection
  - 3 hyper-parameters $c, \sigma, \sigma_{\Omega}$
  - cross-validation to find optimal hyper-parameters
- For Gas Distribution Modelling
  - σ_{Ω} can often be related to σ
  - relatively weak dependence on *c*

2 Kernel DM+V – Preliminary Results

- Alternative Parameter Selection Method
  - error distribution



Salinity Data

3 Salinity Data

- Data
  - output of the Mediterranean Forecasting System (MFS), Bologna, from Nov 20, 2006
    - simulation (3D ocean model)
  - salinity field at 50m depth over the whole med sea
    - 2D slice, values between 35.384 ppt and 39.503 ppt
  - 61933 values in a sparse 253 x 677 matrix (36.2%)
    - x = [0, 4036.2] km ($\Delta \approx$ 6.0km), y = [2.1, 1739.1] km ($\Delta \approx$ 6.9km)

thanks to the Mediterranean Ocean Forecasting System (run by INGV[Istituto Nazionale di Geofisica e Vulcanologia], Bologna)

... and to Alberto Alvarez, NURC

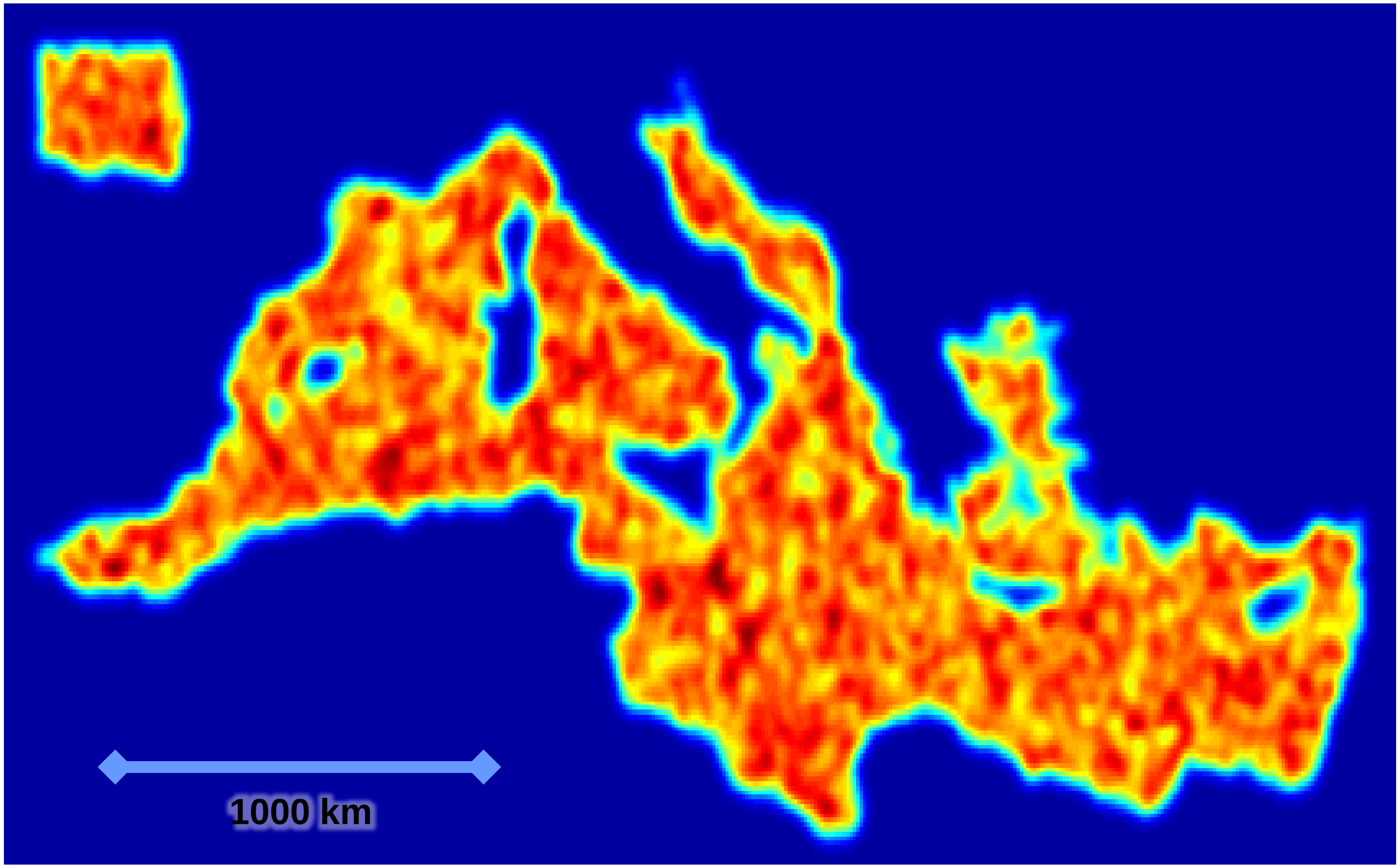
Application of Kernel DM+V to Salinity Data

4 Kernel DM+V on Salinity Data

- Area

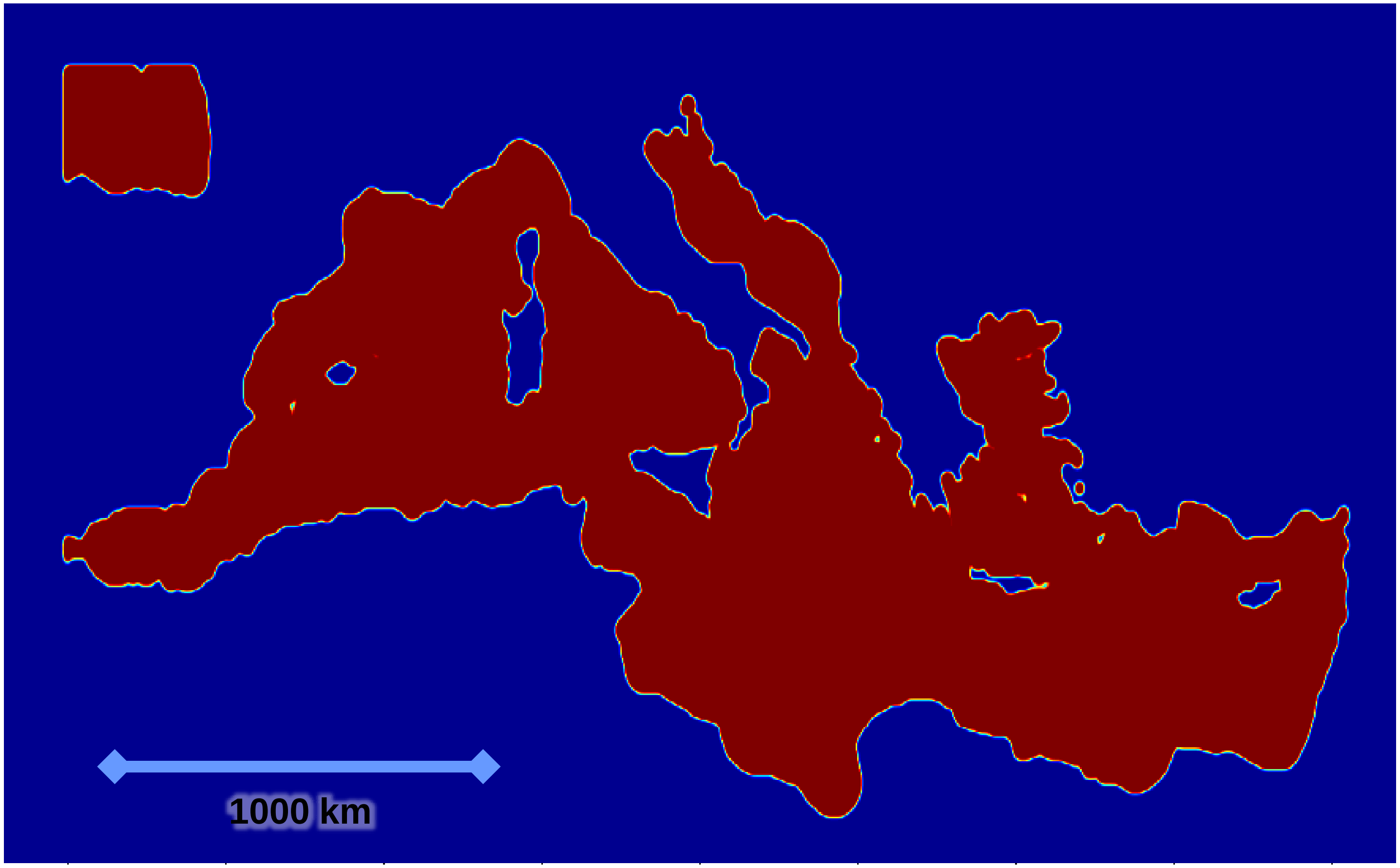
4 Kernel DM+V on Salinity Data

- Information Density
  - cell size: 10 km, kernel width: 20 km



4 Kernel DM+V on Salinity Data

- Confidence Map
  - cell size: 2.5 km, kernel width: 8.5 km

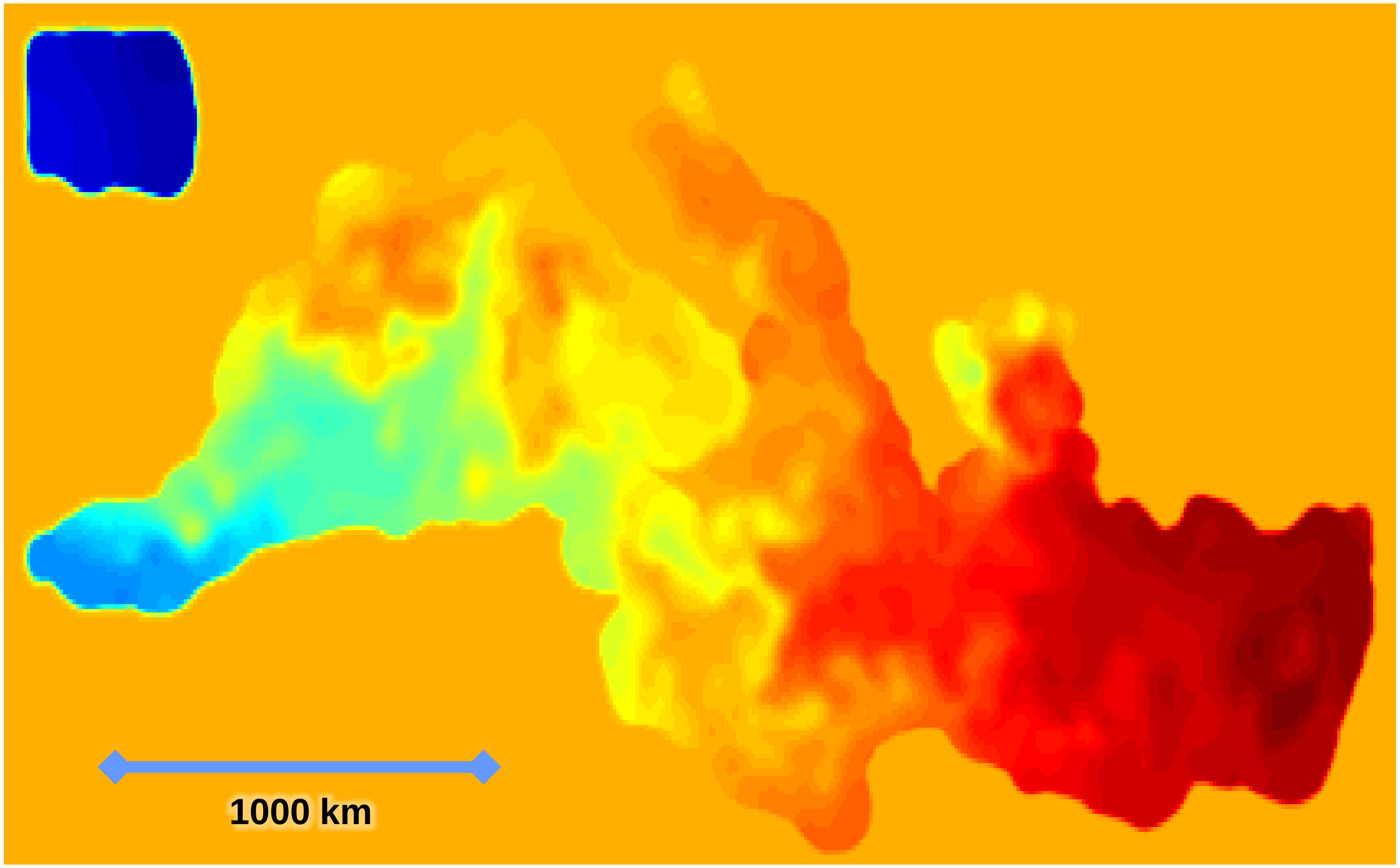


4 Kernel DM+V on Salinity Data

- Area

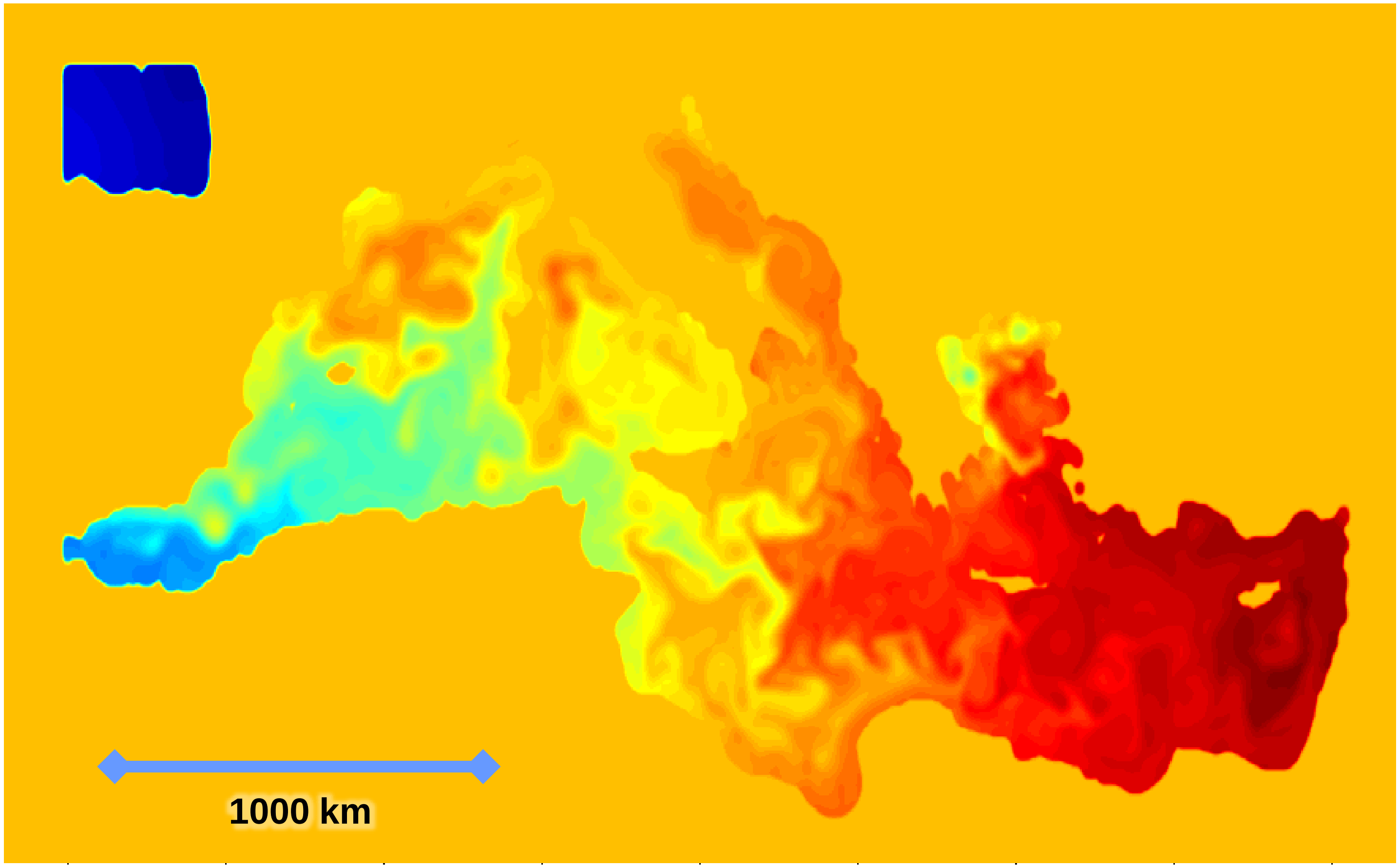
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Mean
  - cell size: 10 km, kernel width: 20 km



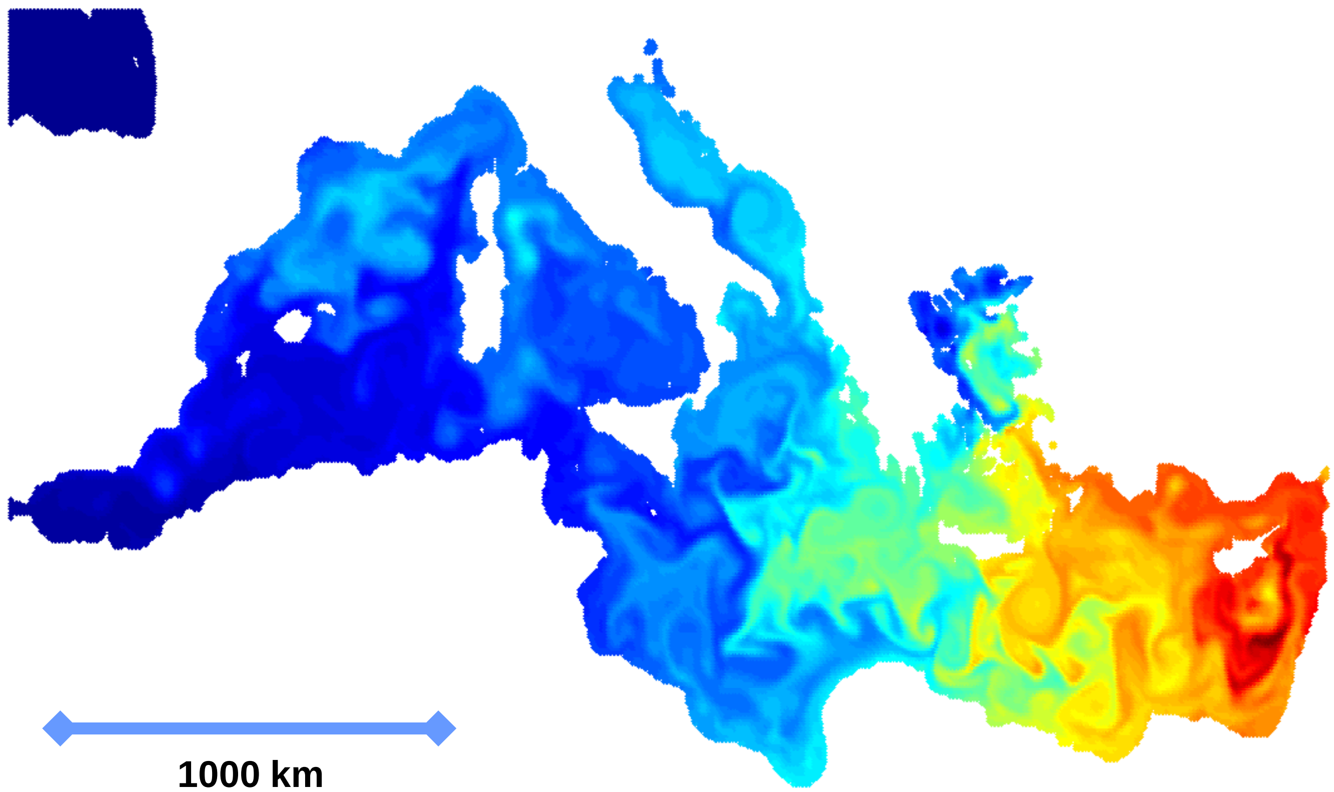
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Mean
  - cell size: 2.5 km, kernel width: 8.5 km



4 Kernel DM+V on Salinity Data

- Salinity Distribution, Mean
  - original data, scatter plot

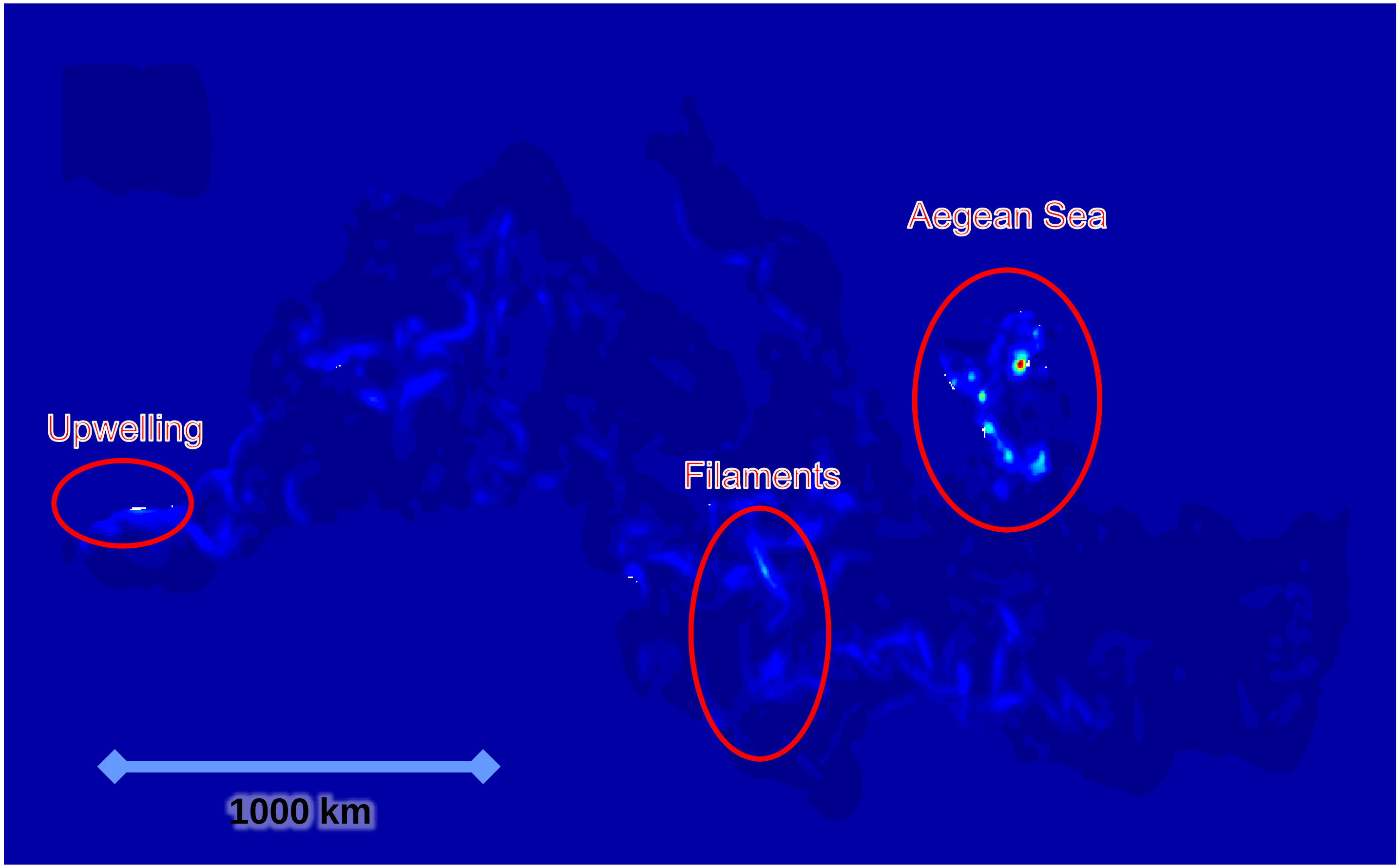


4 Kernel DM+V on Salinity Data

- Area

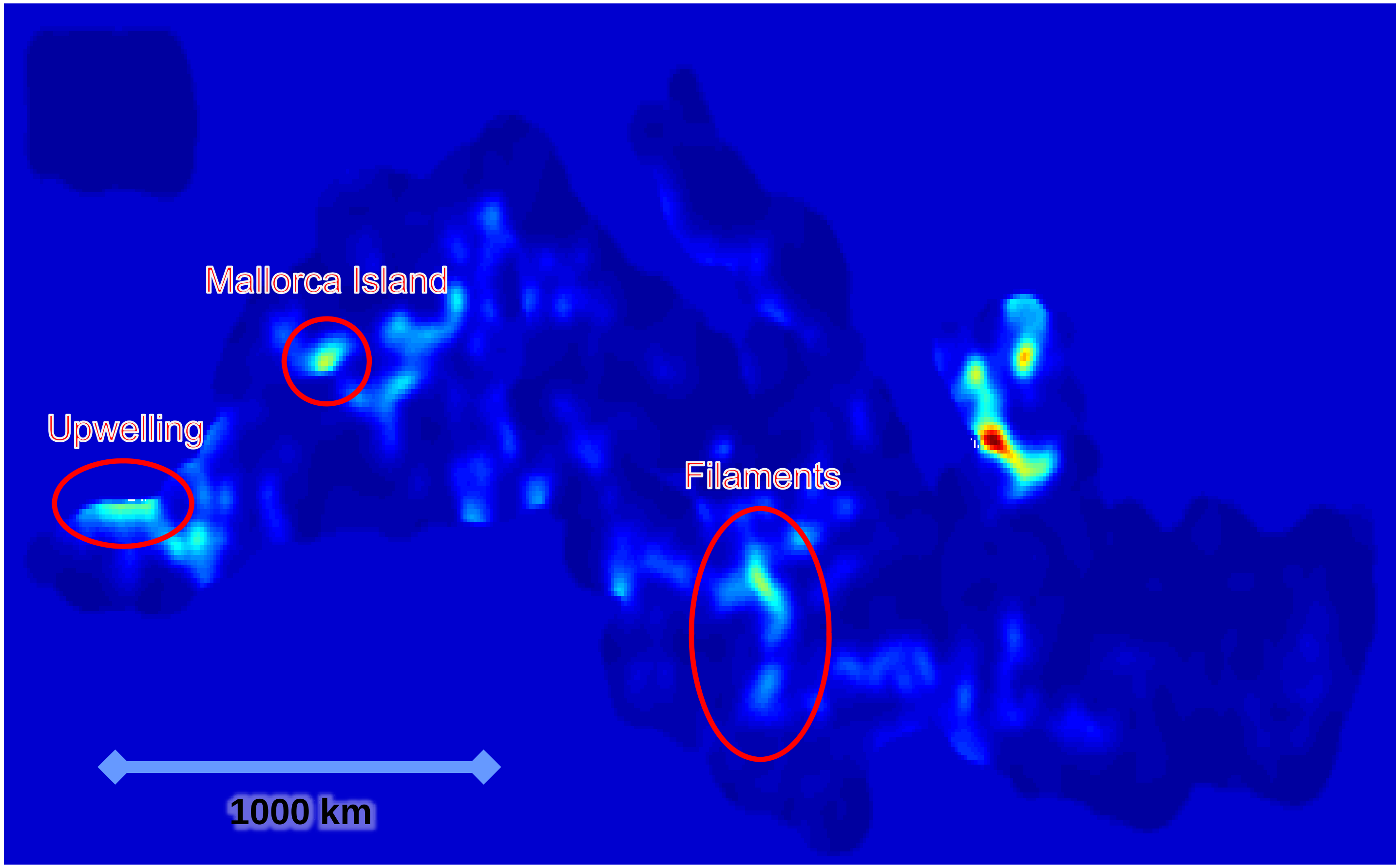
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance
  - cell size: 2.5 km, kernel width: 8.5 km



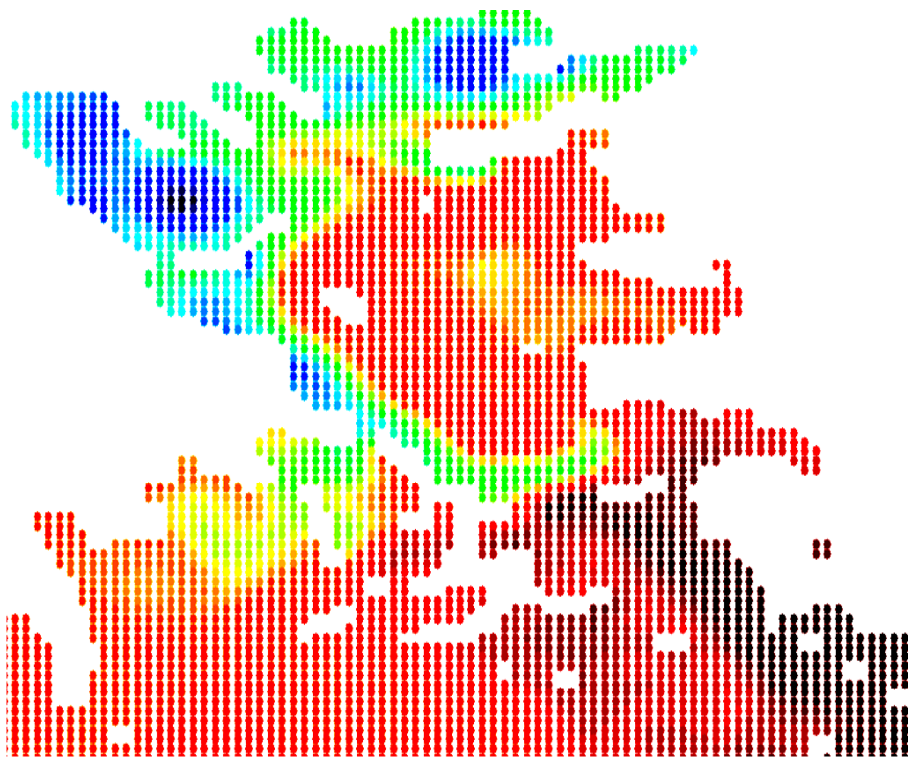
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance
  - cell size: 10 km, kernel width: 20 km



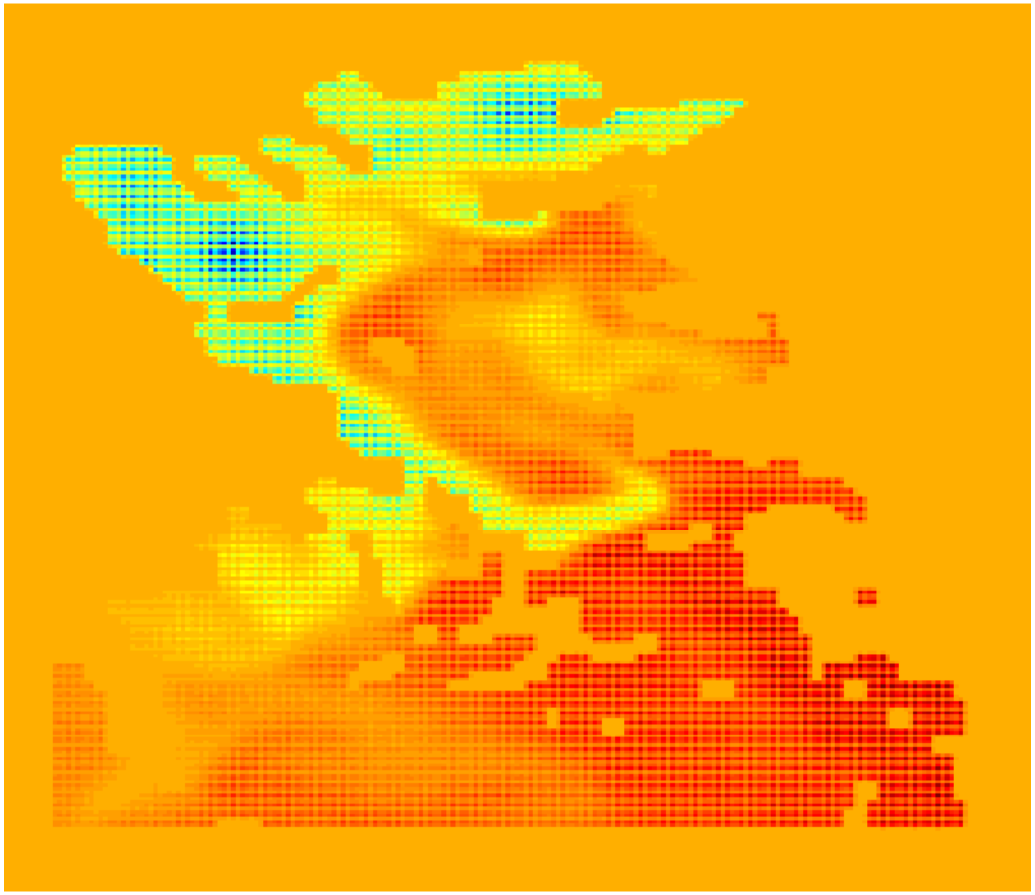
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - scatter plot



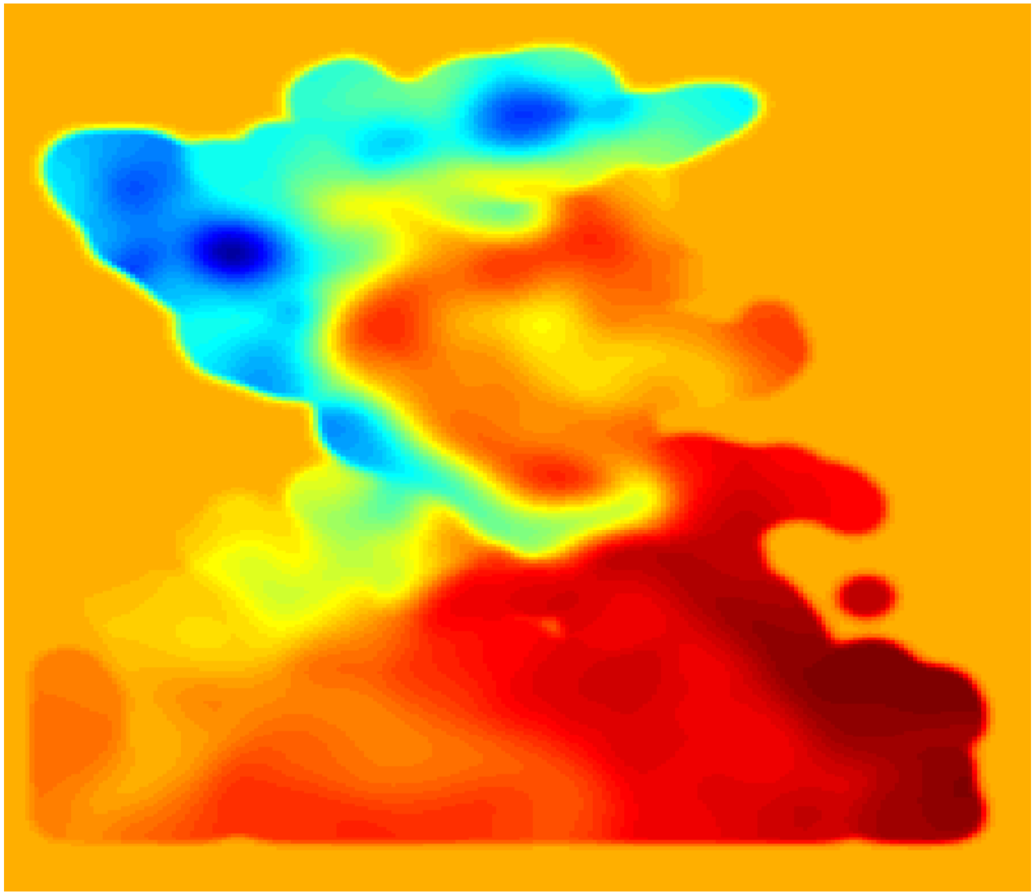
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - cell size: 2.5 km, kernel width: 8.487 km



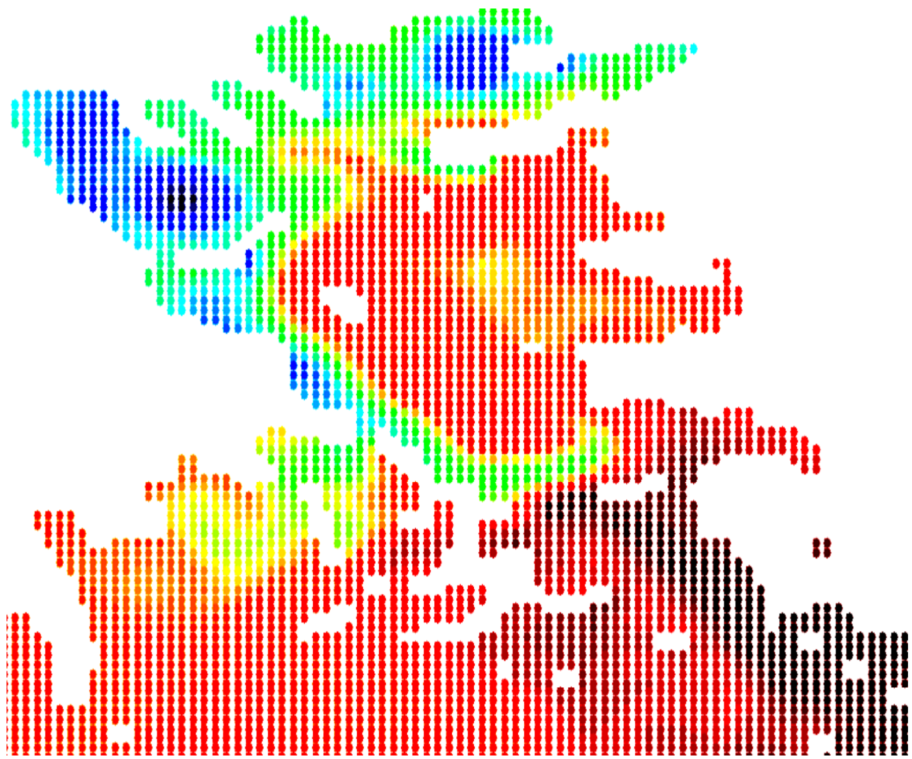
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - cell size: 2.5 km, kernel width: 8.487 km



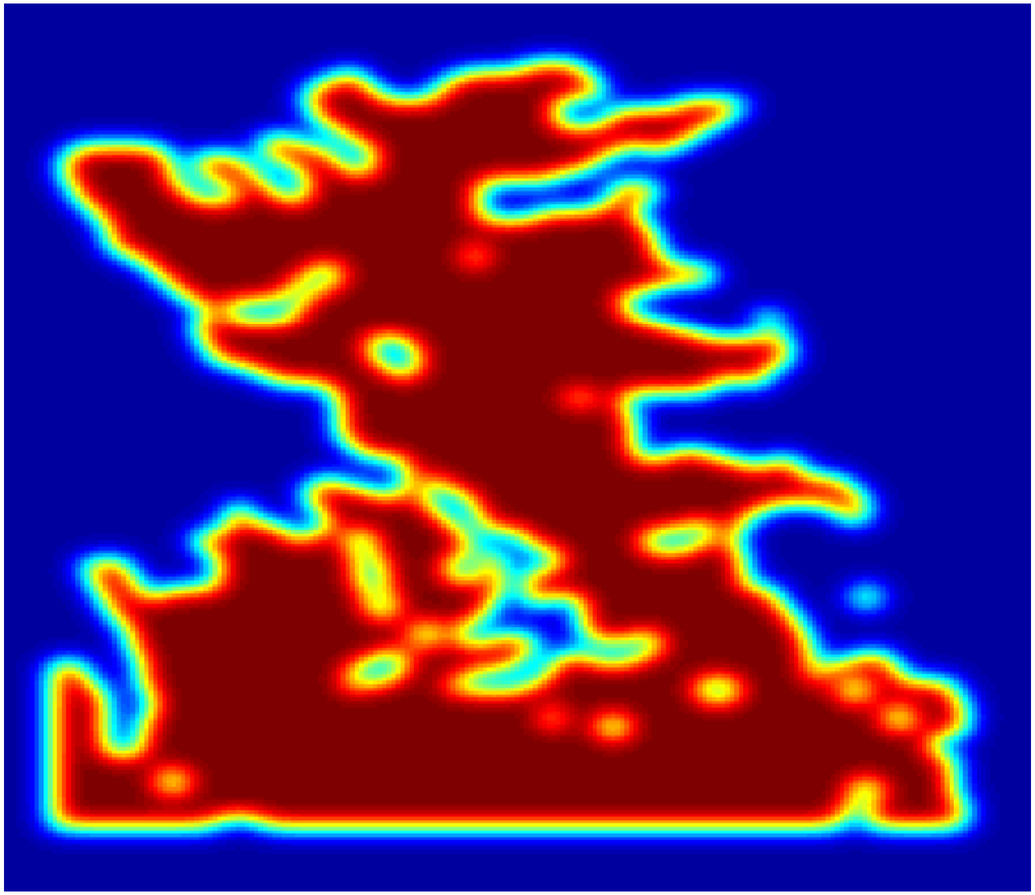
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - scatter plot



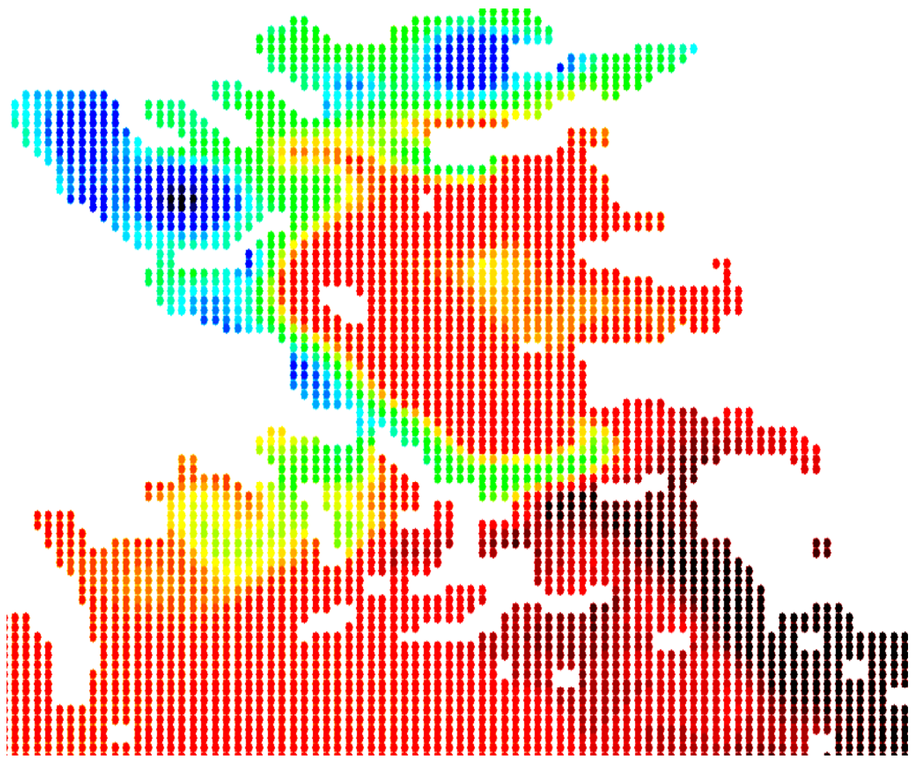
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - cell size: 2.5 km, kernel width: 8.487 km



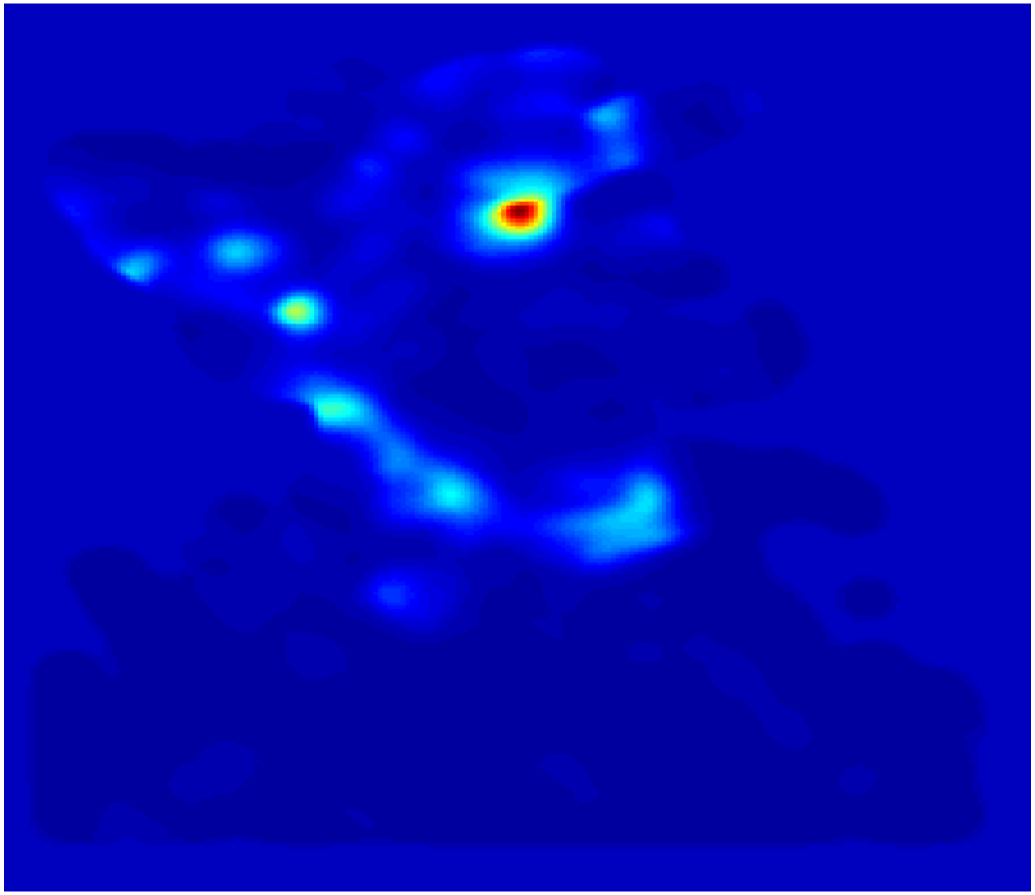
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - scatter plot



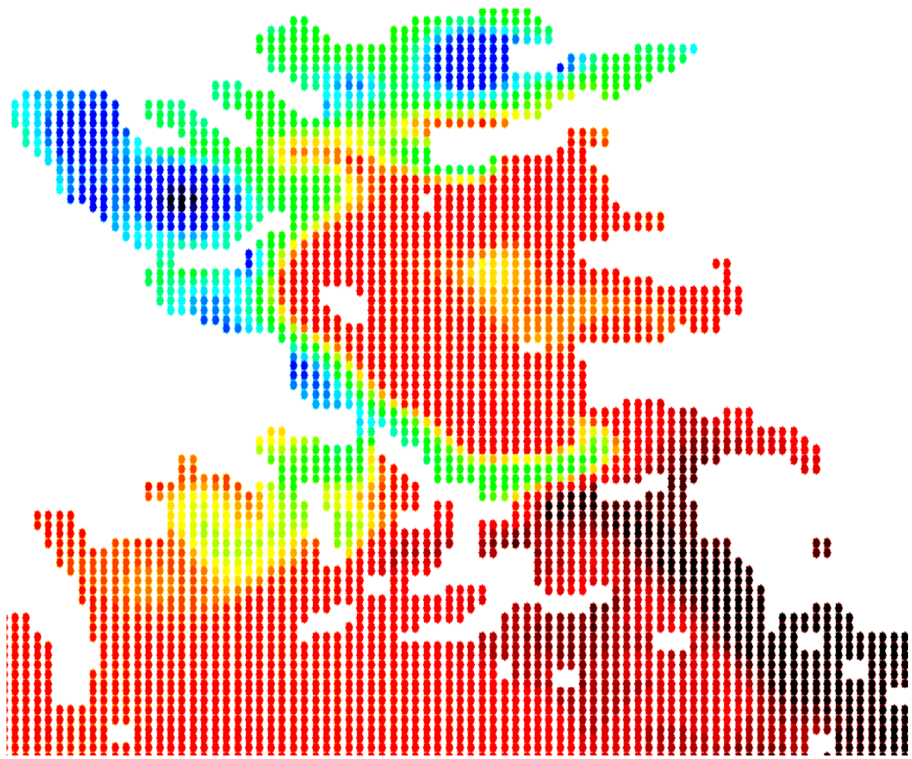
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - cell size: 2.5 km, kernel width: 8.487 km



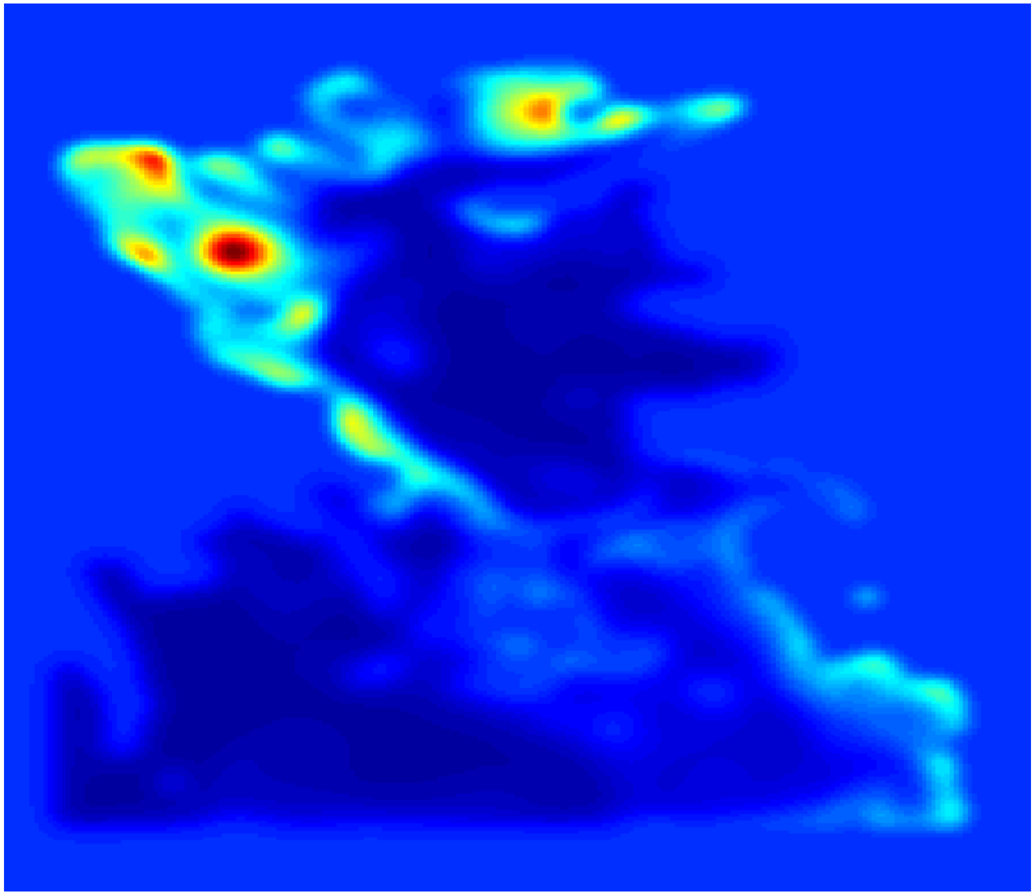
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - scatter plot



4 Kernel DM+V on Salinity Data

- Salinity Distribution, Variance, Aegean Sea
  - cell size: 2.5 km, kernel width: 8.487 km, σ_{Ω}=10xGauss$_{\sigma}$(0,0)



4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 30km,
          - σ_0 = 10km (within [1km 50km]) → σ = 20.962 km
          - → NLPD = -1.41727 (test), -1.4228 (all)
        - c = 20km
          - σ_0 = 10km (within [1km 50km]) → σ = 15.513 km
          - → NLPD = -1.7486 (test), -1.6449 (all)
        - c = 10km
          - σ_0 = 10km (within [1km 50km]) → σ = 10.764 km
          - → NLPD = -2.26314 (test), -2.3235 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 10km
          - σ_0 = 10km (within [1km 50km]) → σ = 10.764 km
          - → NLPD = -2.26314 (test), -2.3235 (all)
        - c = 5km
          - σ_0 = 2km (within [1km 50km]) → σ = 8.786 km
          - → NLPD = -2.7168 (test), -2.8239 (all)
        - c = 2.5km
          - σ_0 = 5km (within [1km 50km]) → σ = 8.487 km
          - → NLPD = -2.91587 (test), -3.0443 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 2.5km
          - σ_0 = 5km (within [1km 50km]) → σ = 8.487 km
          - → NLPD = -2.91587 (test), -3.0443 (all)
        - c = 1.0km
          - σ_0 = 5km (within [1km 50km]) → σ = 7.950 km
          - → NLPD = -3.08507 (test), -3.2308 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 30km,
          - σ_0 = 10km (within [1km 50km]) → σ = 20.962 km
          - → NLPD = -1.41727 (test), -1.4228 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 30km,
          - σ_0 = 10km (within [1km 50km]) → σ = 20.707 km
          - → NLPD = -1.42026 (test), -1.4237 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 20km
          - σ_0 = 10km (within [1km 50km]) → σ = 15.513 km
          - → NLPD = -1.7486 (test), -1.6449 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 20km
          - σ_0 = 2km (within [1km 50km]) → σ = 15.450 km
          - → NLPD = -1.7268 (test), -1.7495 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 10km
          - σ_0 = 10km (within [1km 50km]) → σ = 10.764 km
          - → NLPD = -2.26314 (test), -2.3235 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 10km
          - σ_0 = 2km (within [1km 50km]) → σ = 10.049 km
          - → NLPD = -2.29108 (test), -2.3571 (all)

4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 5km
          - σ_0 = 2km (within [1km 50km]) → σ = 8.786 km
          - → NLPD = -2.7168 (test), -2.8239 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 5km
          - σ_0 = 2km (within [1km 50km]) → σ = 8.731 km
          - → NLPD = -2.73475 (test), -2.8301 (all)

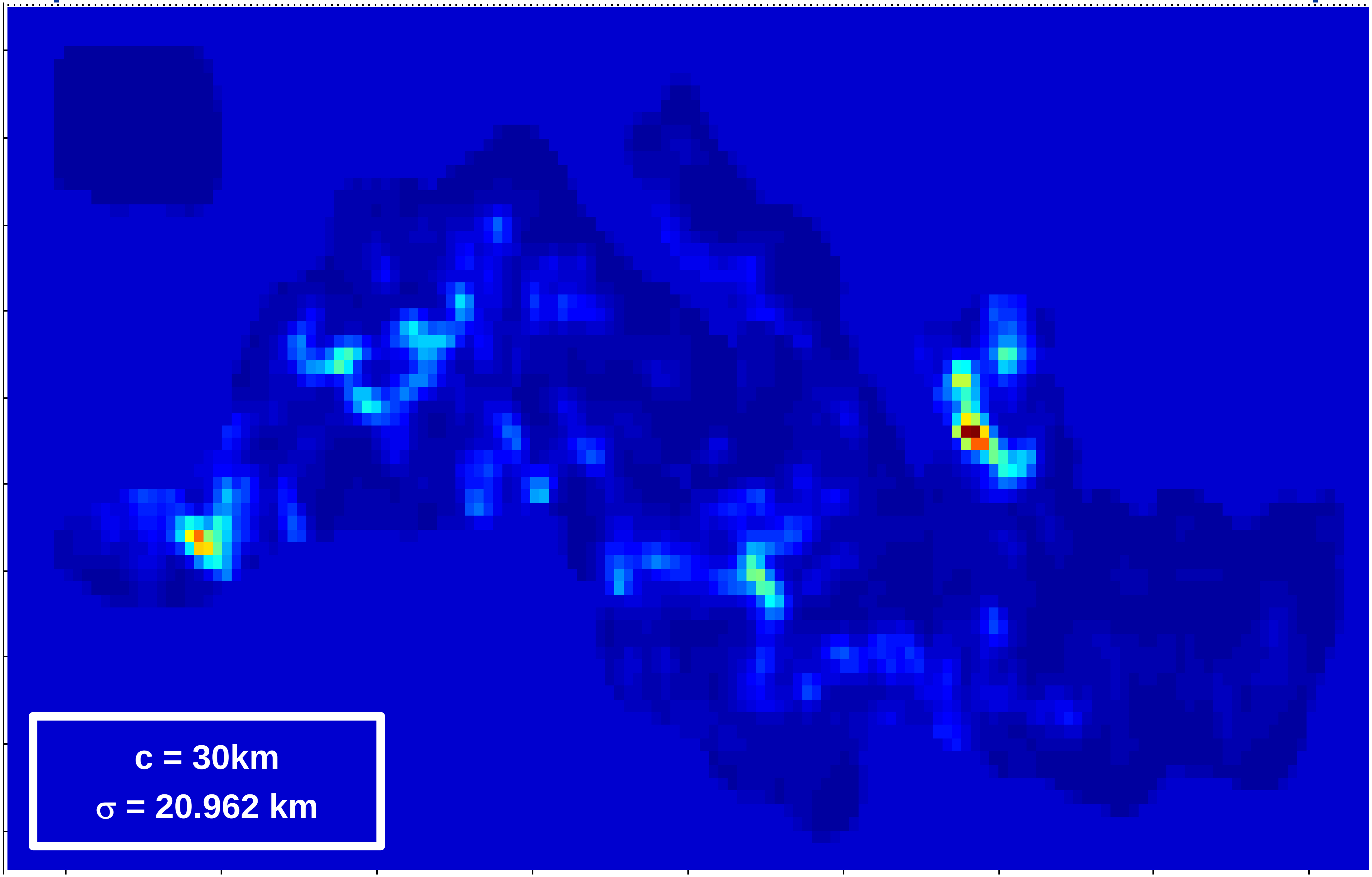
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 2.5km
          - σ_0 = 5km (within [1km 50km]) → σ = 8.487 km
          - → NLPD = -2.91587 (test), -3.0443 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 2.5km
          - σ_0 = 5km (within [1km 50km]) → σ = 8.327 km
          - → NLPD = -2.9424 (test), -3.0707 (all)

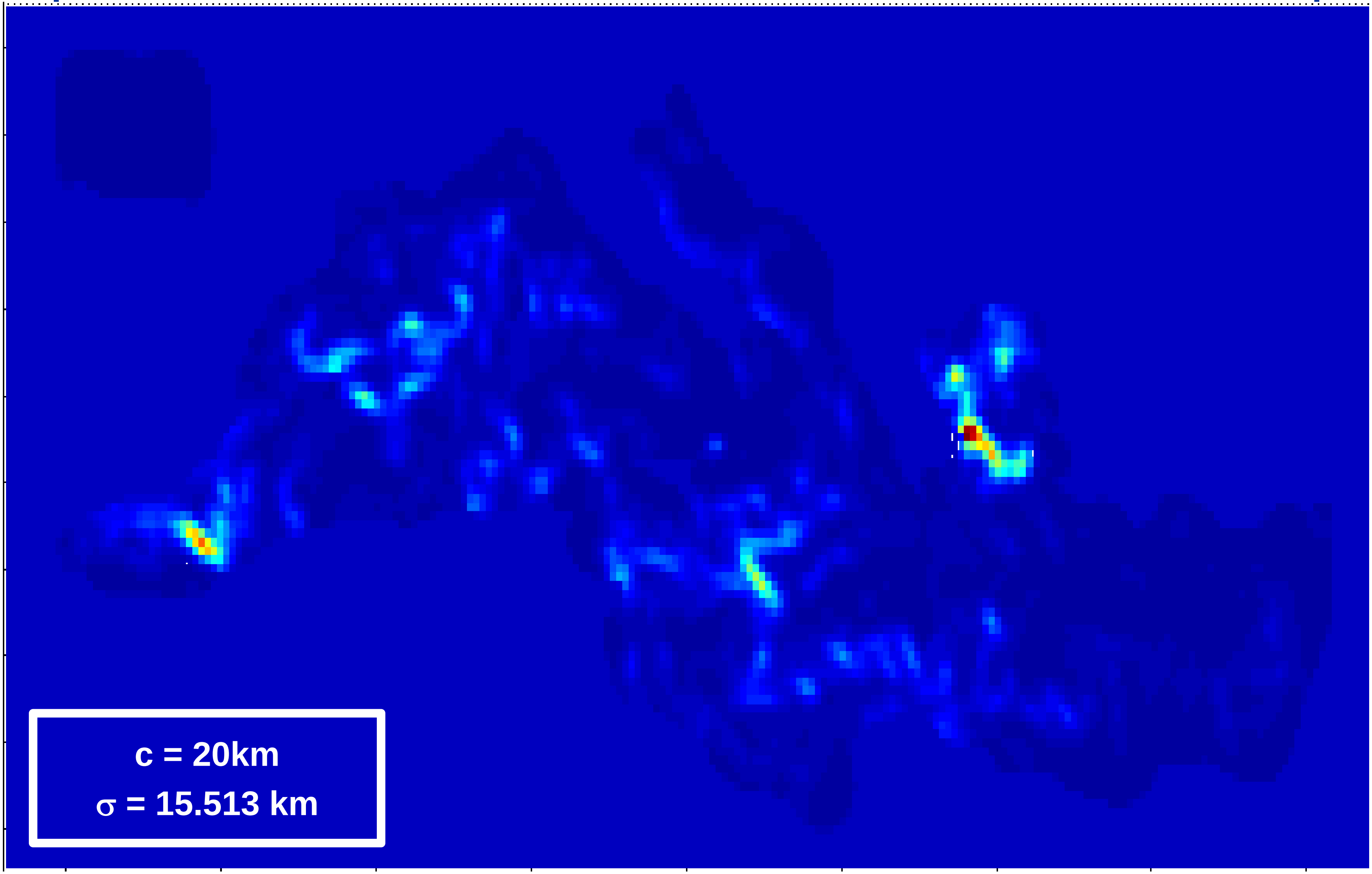
4 Kernel DM+V on Salinity Data

- Salinity Distribution, Learning
  - #data points: 61933
    - training (Levenberg-Marquardt parameters: [0.01 0.01 100 1000])
      - training: [1:2:61933], test: [2:2:61933]
        - c = 1.0km
          - σ_0 = 5km (within [1km 50km]) → σ = 7.950 km
          - → NLPD = -3.08507 (test), -3.2308 (all)
      - training: [2:2:61933], test: [1:2:61933]
        - c = 1.0km
          - σ_0 = 5km (within [1km 50km]) → σ = 7.806 km
          - → NLPD = -3.11175 (test), -3.2607 (all)

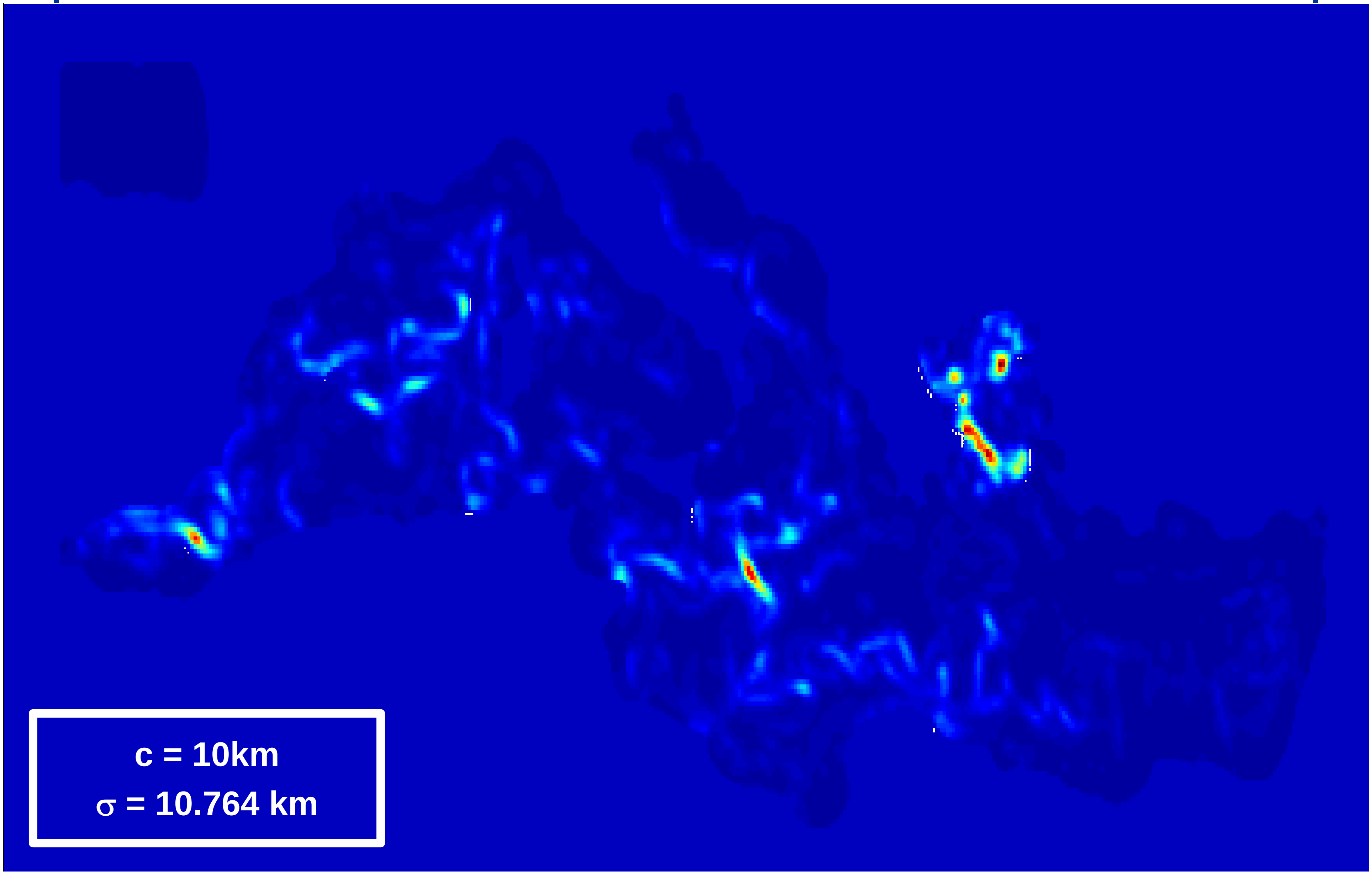
4 Kernel DM+V on Salinity Data



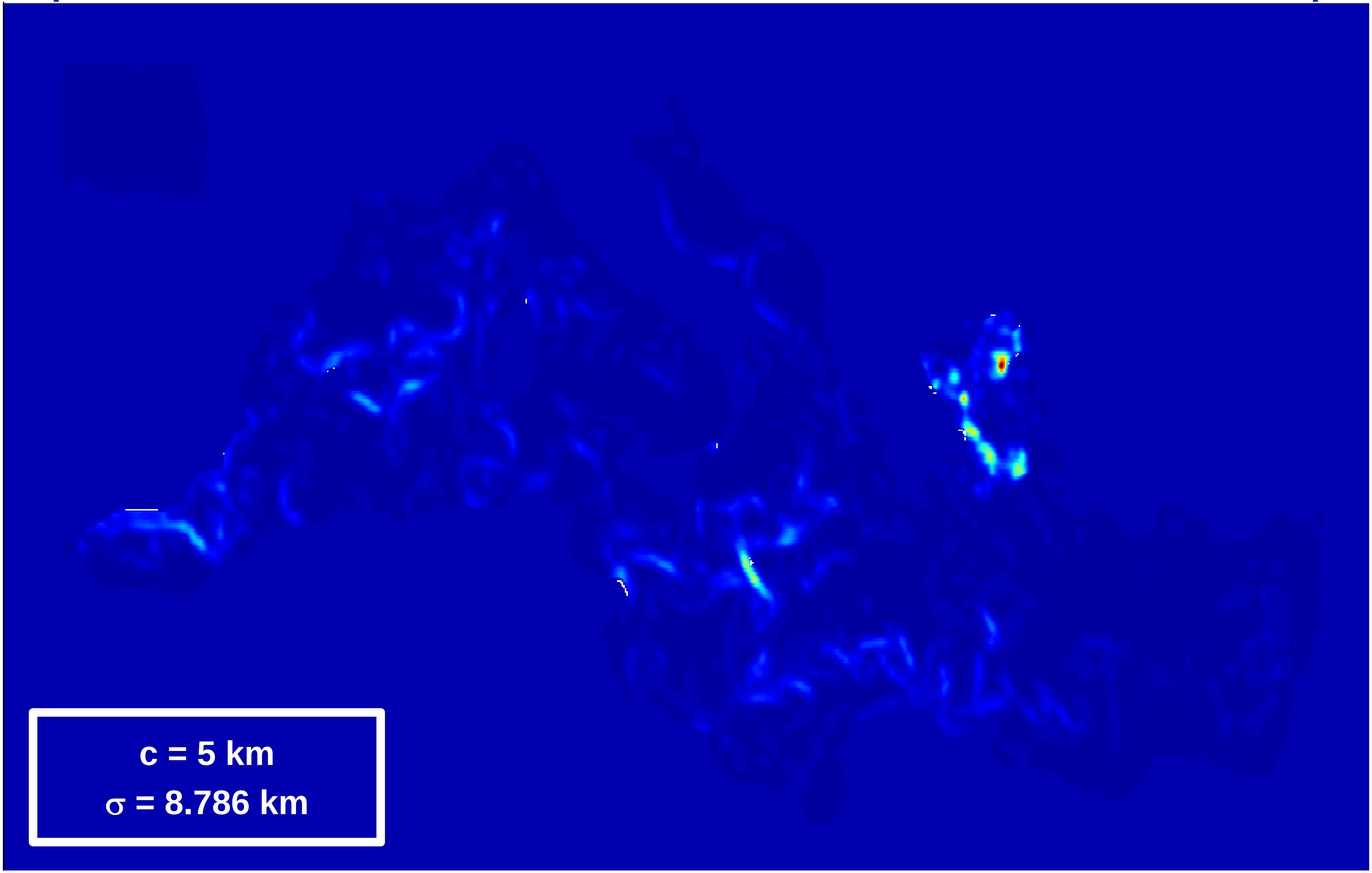
4 Kernel DM+V on Salinity Data



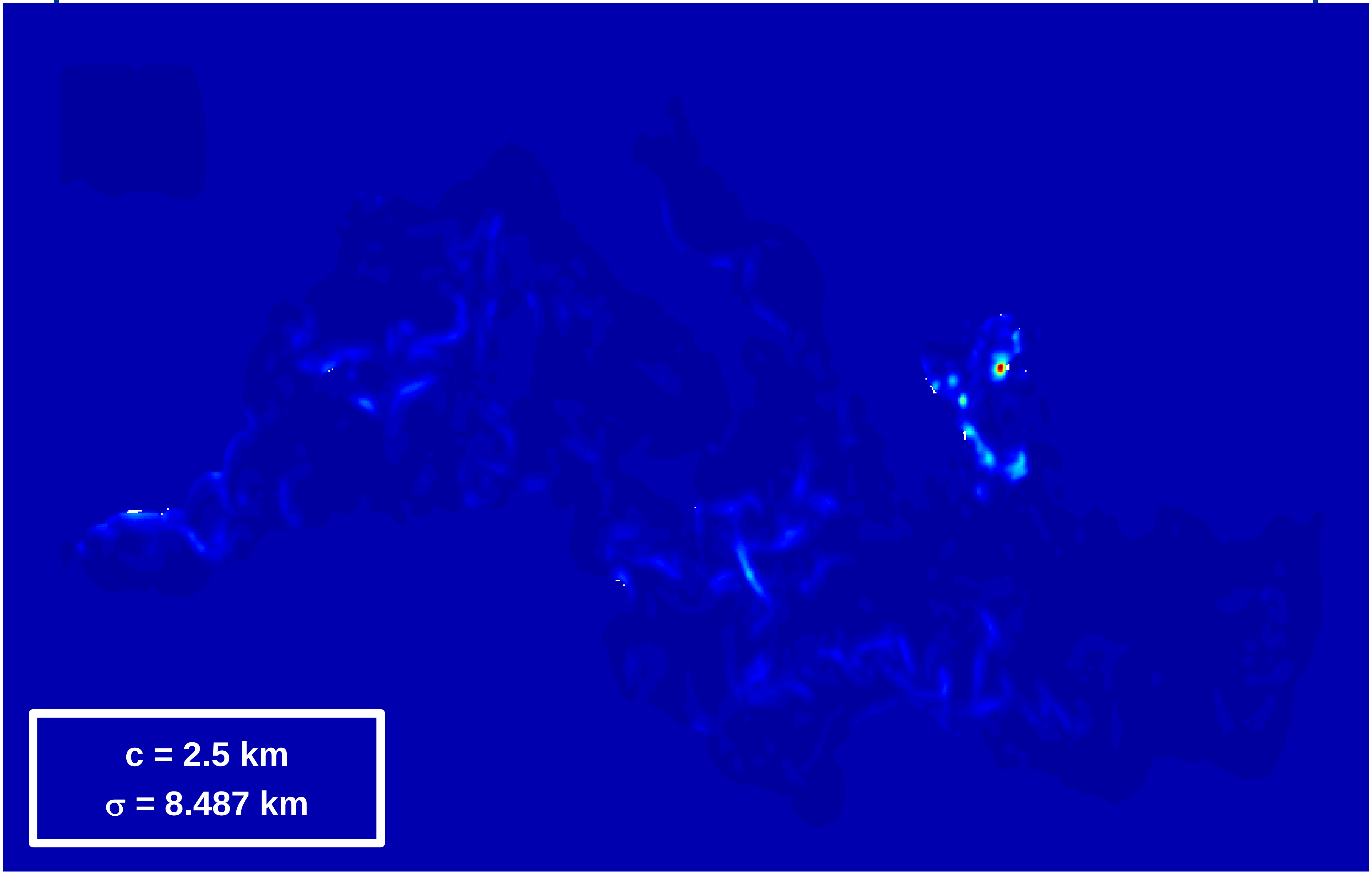
4 Kernel DM+V on Salinity Data



4 Kernel DM+V on Salinity Data



4 Kernel DM+V on Salinity Data

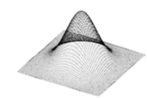


Summary and Conclusions

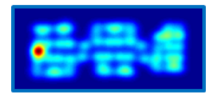
5 Summary

- Kernel DM+V

$$\omega_{k,i} = Gauss\left(\left|\vec{x}_i - \vec{x}_k\right|, \sigma\right)$$



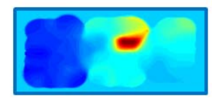
$$\Omega_k = \sum_{i=1}^{|D|} \omega_{k,i}$$



$$\alpha_k = 1 - \exp\left[-\Omega_k^2 / \sigma_\Omega^2\right]$$

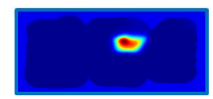
$$r_k = \alpha_k \cdot R_k / \Omega_k + \{1 - \alpha_k\} \cdot r_0$$

$$R_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot r_i \qquad r_0 = \frac{1}{|D|} \sum_{i=1}^{|D|} r_i$$



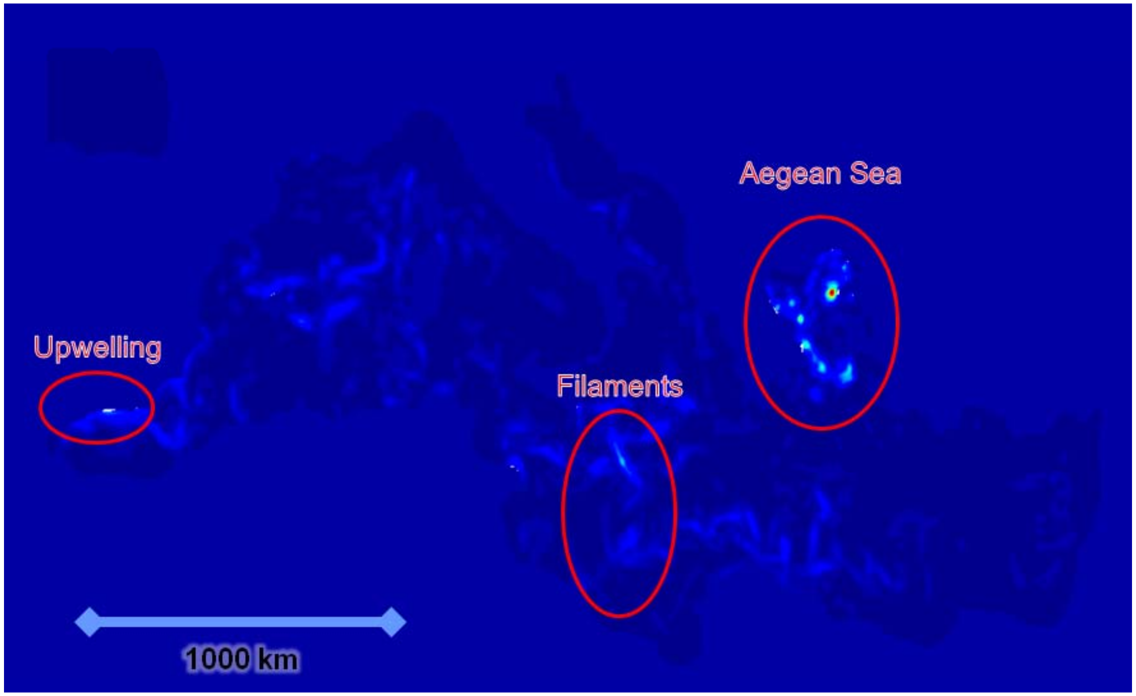
$$v_k = \alpha_k \cdot V_k / \Omega_k + \{1 - \alpha_k\} \cdot v_0$$

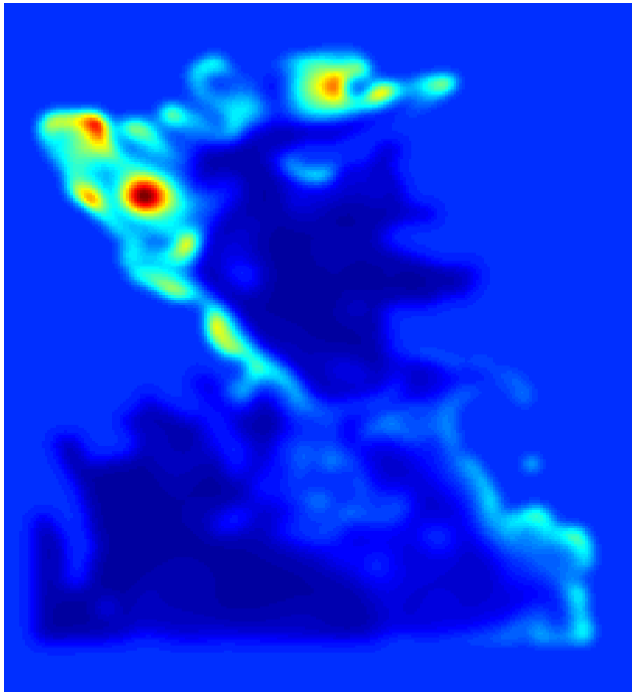
$$\tau_i = \left(r_i - r_{k(i)}\right)^2 \quad V_k = \sum_{i=1}^{|D|} \omega_{k,i} \cdot \tau_i \quad v_o = \frac{1}{|D|} \sum_{i=1}^{|D|} \tau_i$$



5 Summary

- Kernel DM+V
- Application of Kernel DM+V to Salinity Data

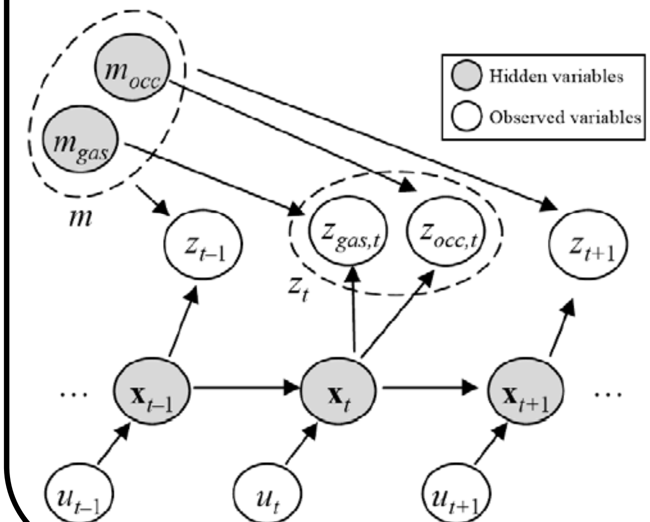


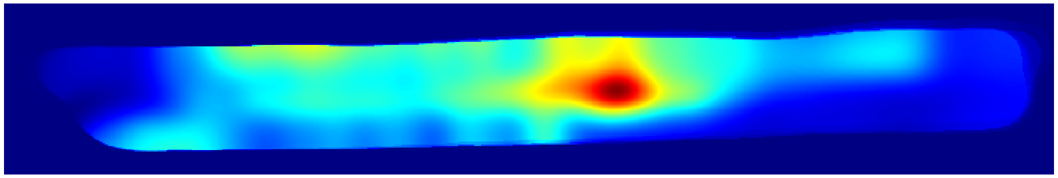


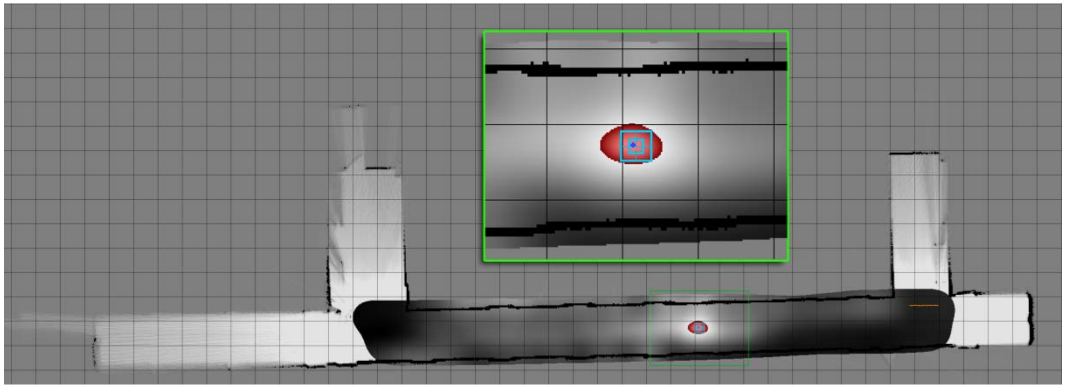
Future Work,

Kernel DM+V Extensions

ECMR 2007, Lilienthal et al., A Rao-Blackwellisation Approach to GDM-SLAM – Integrating SLAM and Gas Distribution Mapping



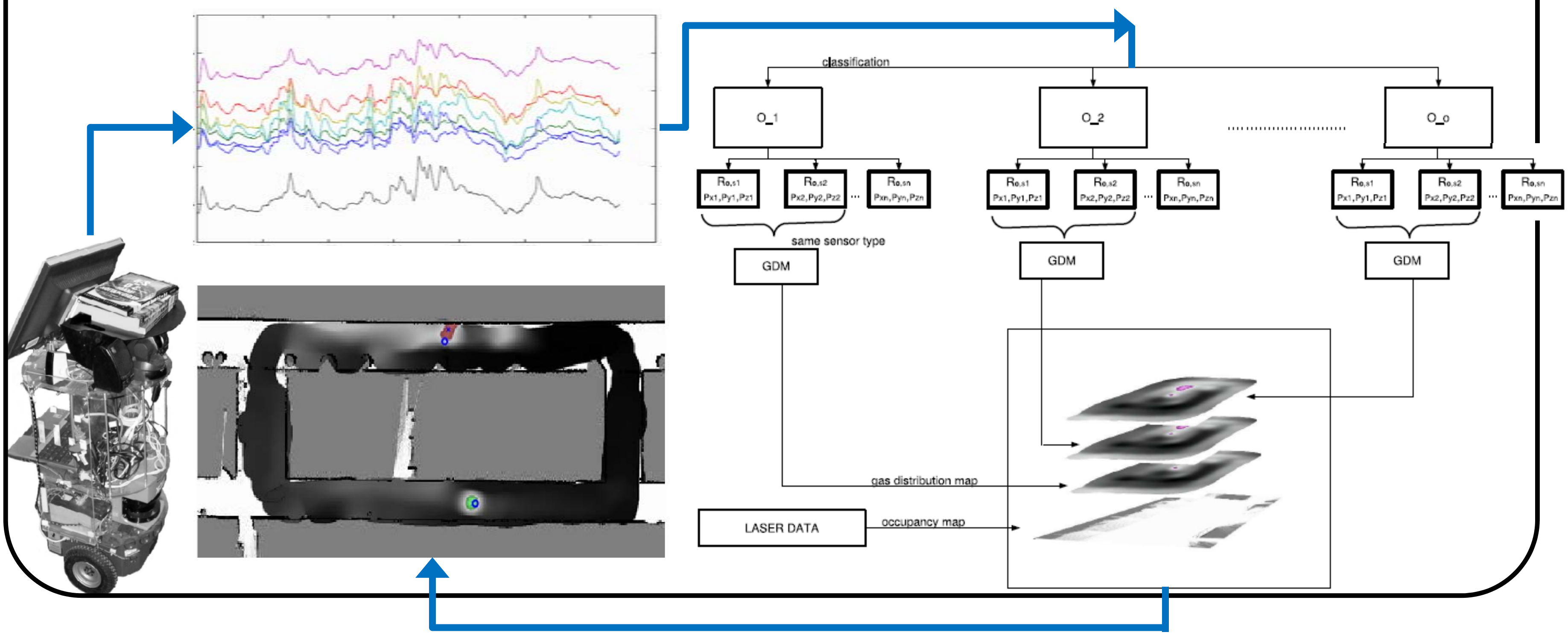




2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011

6 Results – Gas Distribution Modelling

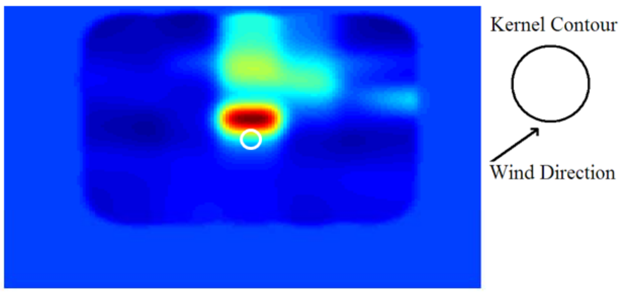
Robotics 2009, Loutfi et al., Gas Distribution Mapping of Multiple Odour Sources using a Mobile Robot

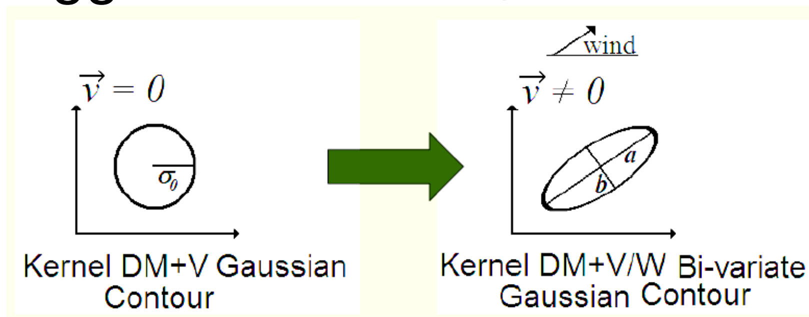


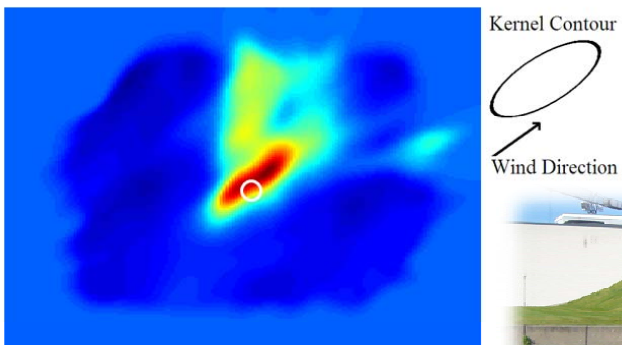
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011

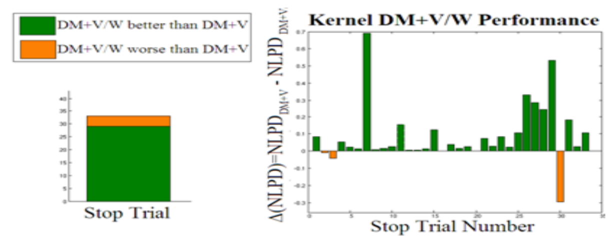
6 Results – Gas Distribution Modelling

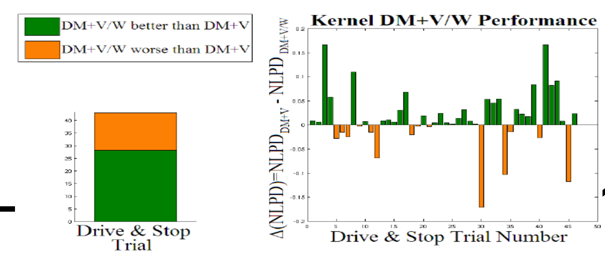
IEEE Sensors 2009, Reggente/Lilienthal, Kernel DM+V/W*







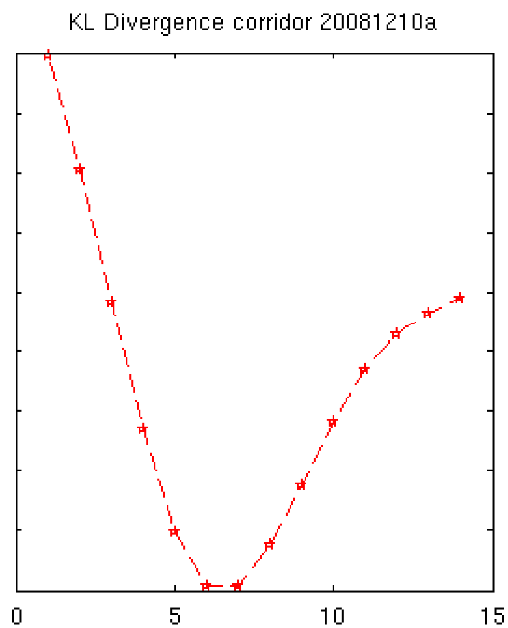


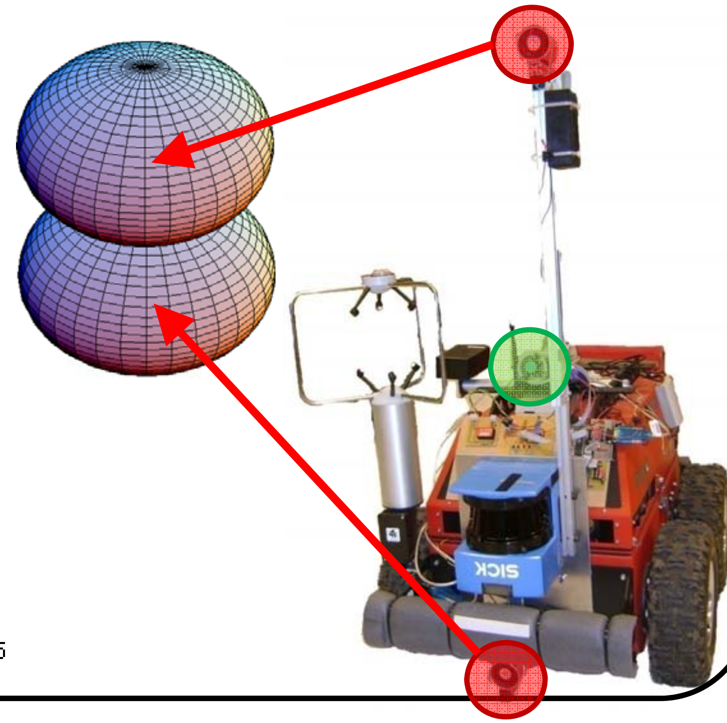


2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011

6 Results – Gas Distribution Modelling

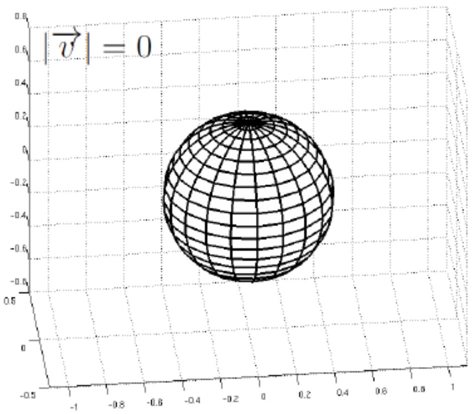
ISOEN 2009, Reggente/Lilienthal, 3D Statistical Gas Distribution Mapping in an Uncontrolled Indoor Environment



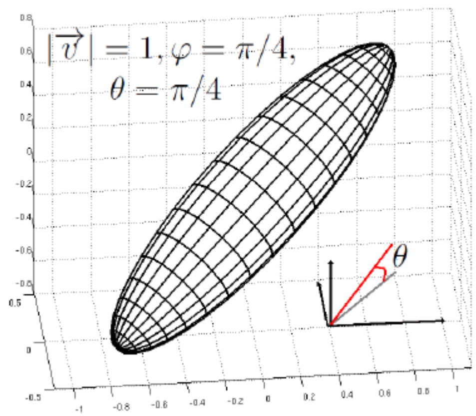


2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011

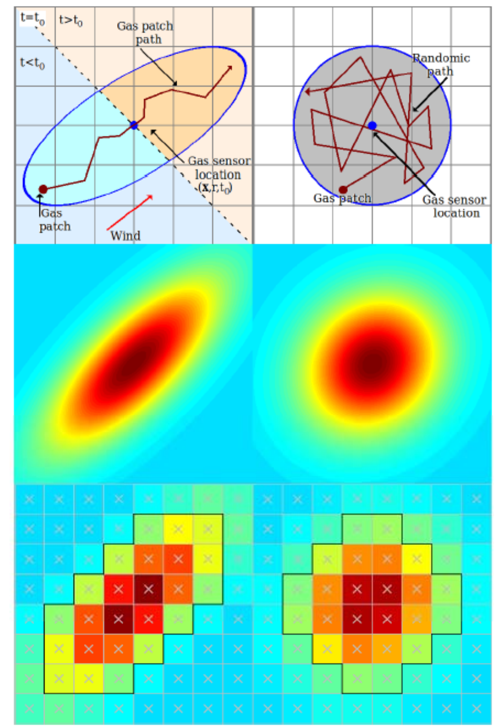
IEEE Sensors 2010, Reggente/Lilienthal, 3D-Kernel DM+V/W*



$$\Sigma = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.2 \end{bmatrix}$$



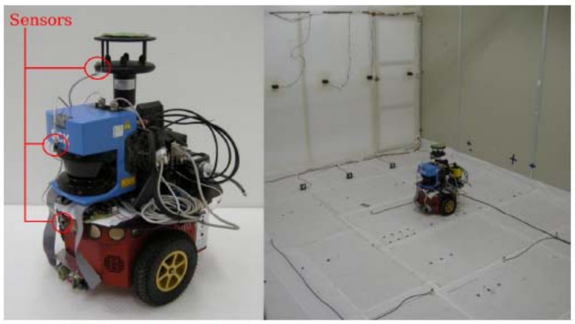
$$\Sigma_{R(\theta)} = \begin{bmatrix} 0.64 & 0.39 & 0.39 \\ 0.39 & 0.36 & 0.27 \\ 0.39 & 0.27 & 0.36 \end{bmatrix}$$

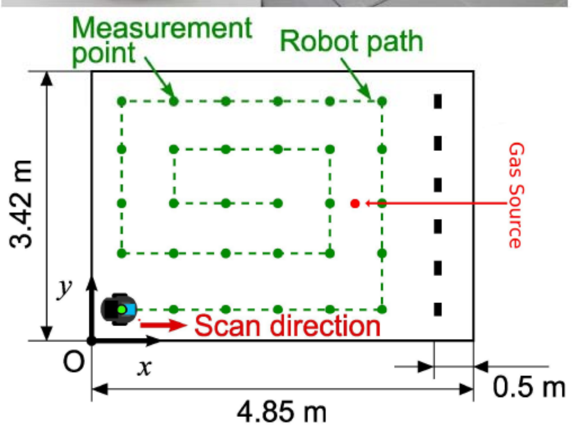


2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011

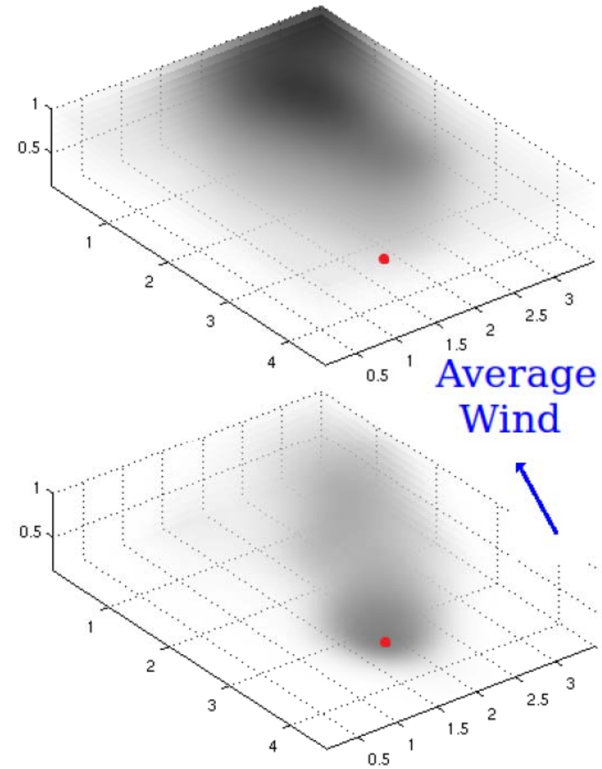
6 Results – Gas Distribution Modelling

IEEE Sensors 2010, Reggente/Lilienthal, 3D-Kernel DM+V/W*





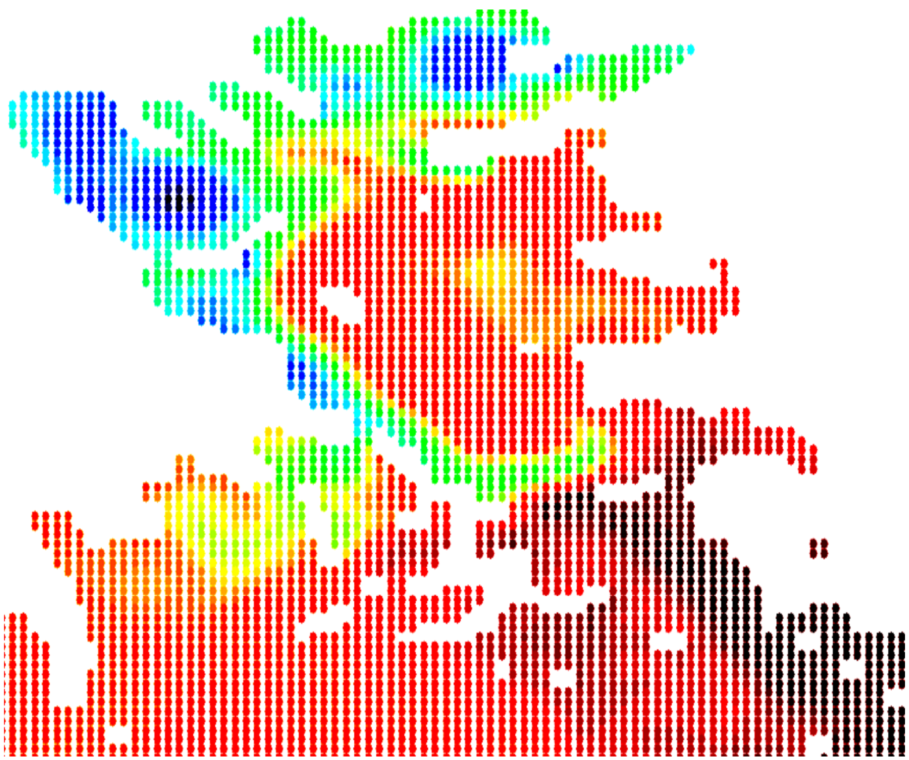
Trial	NLPD Wind	NLPD no Wind
1	-1.76	-1.72
2	-1.48	-1.44
3	-1.11	-0.87
4	-1.44	-1.37
5	-1.02	-0.99
6	-1.28	-1.25
7	-0.99	-0.90
8	-0.93	-0.92
9	-0.91	-0.89
10	-1.46	-1.50
11	-1.17	-1.13
12	-1.24	-1.06
13	-1.18	-1.14
14	-1.37	-1.32
15	-1.04	-1.03

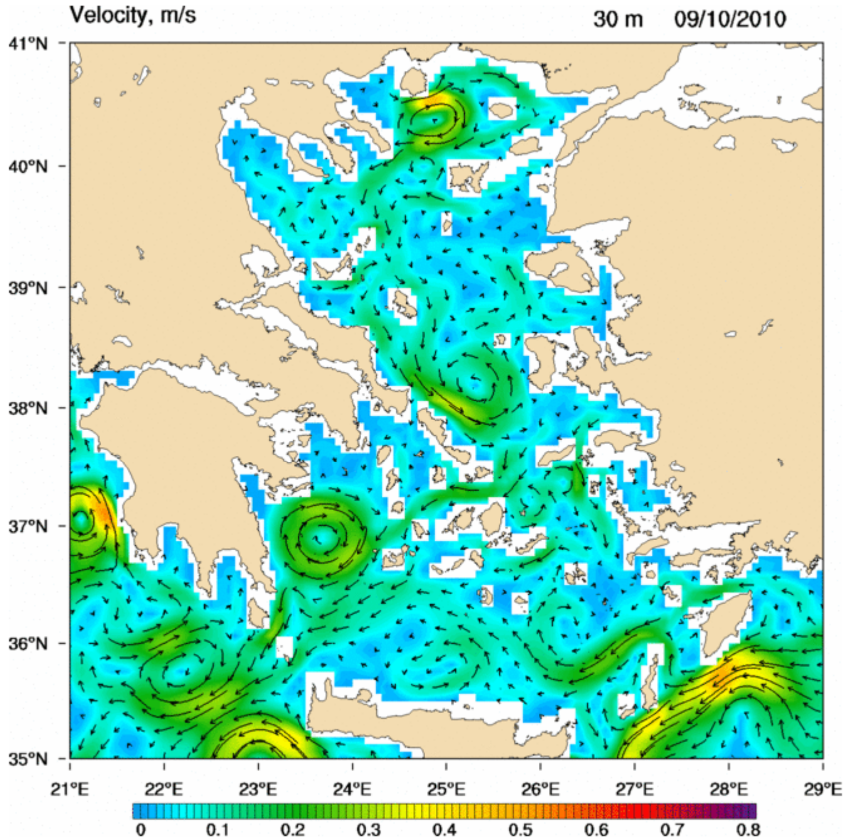


2001 2002 2003 2004 2005 2006 2007 2008 2009 **2010** 2011

6 Future Work, Kernel DM+V/W

- Salinity Distribution, Mean
  - input: salinity + flow data





6 Future Work

- Multi-Scale Extensions
- Model Tracer Distributions
  - oil spills
  - chlorophyll
  - other quantities that can be sensed remotely

Thanks For Your Attention!

Statistical Distribution Modelling with the Kernel DM+V Algorithm – Application to Oceanographic Data

Achim J. Lilienthal

suggested by Andrea Caiti and with help from Alberto Alvarez

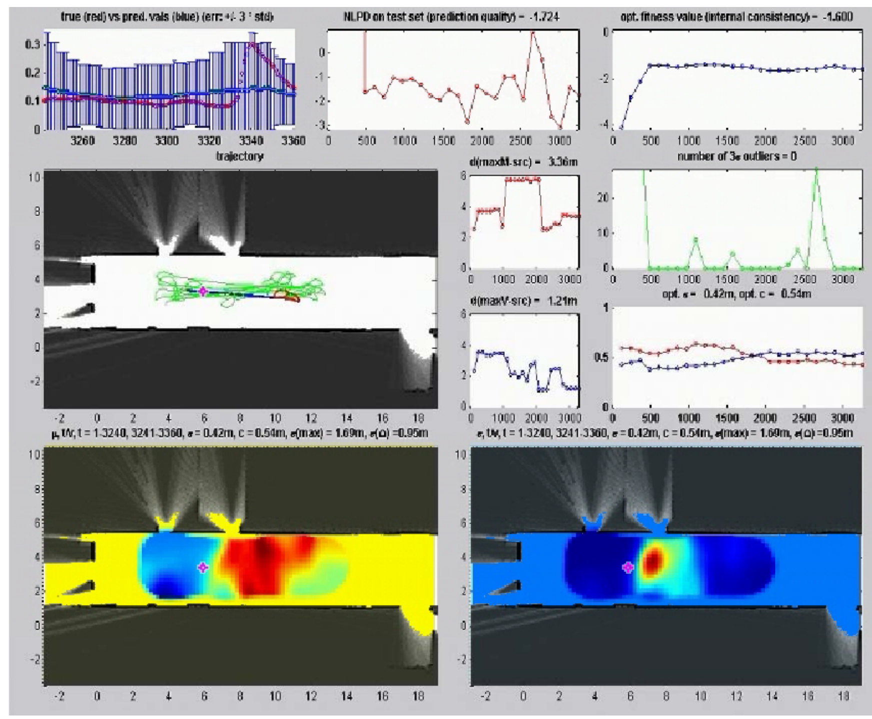
5

Agenda

GARBAGE

5 Statistical GDM with Pred. Variance – Summary

- Estimating the Pred. Variance Entails a Significant Step Forward for Statistical GDM!
  1. the models better fit the structure of gas distributions



5 Statistical GDM with Pred. Variance – Summary

- Estimating the Pred. Variance Entails a Significant Step Forward for Statistical GDM!
  1. the models better fit the structure of gas distributions
  - allows model evaluation in terms of the data likelihood
    2. learning meta parameters
    3. comparison of different approaches to statistical GDM

Dataset	GP	GPM	Kernel DM+V
3-rooms	-0.90	-1.54	-1.44
corridor (avg)	-0.98	-1.60	-1.81
outdoor	-0.94	-1.77	-1.75

5 Statistical GDM with Pred. Variance – Summary

- Estimating the Pred. Variance Entails a Significant Step Forward for Statistical GDM!
  1. the models better fit the structure of gas distributions
  - allows model evaluation in terms of the data likelihood
    2. learning meta parameters
    3. comparison of different approaches to statistical GDM
  - provides the means for
    4. sensor planning
       (suggest new measurement locations based on the current model)
    5. lazy update mechanisms
       (determine when the model should be updated or re-initialised)