Anisotropic functional Laplace deconvolution

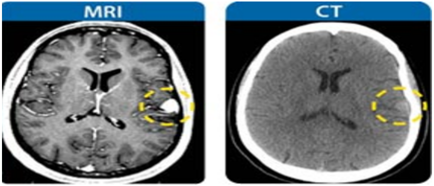
Rasika Rajapakshage and Marianna Pensky
Department of Mathematics,University of Central Florida
Rida Benhaddou
Department of Mathematics,Ohio University

May 23, 2018

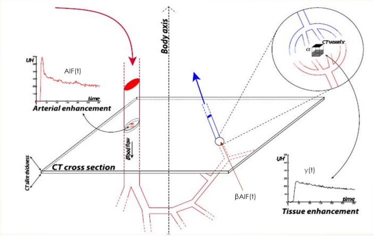
Content

Motivation

- This study is motivated by the analysis of Dynamic Contrast Enhanced (DCE) imaging data, E.g MRI,CT scan images
- DCE imaging provides a non-invasive measure of tumor angiogenesis and cancer detection and characterization, monitoring.
- Used in various medical assessments such as brain flows, strokes or cancer angiogenesis

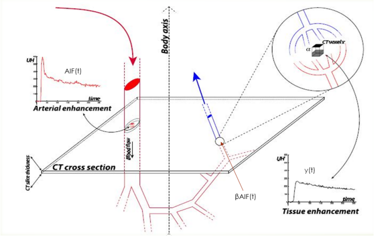


Motivation



- The particles of the contrast agent entering a tissue voxel $\mathbf{x}$ satisfies the following equation.
- $$Y(t,\mathbf{x}) = \int_0^{t-\delta} g(t-z)\,\beta(\mathbf{x})(1-F(z,\mathbf{x}))dz + \varepsilon\xi(t,\mathbf{x}) \qquad (1)$$
- The errors $\xi(t,\mathbf{x})$ are independent for different t and $\mathbf{x} = (x_1, x_2)$, $g(t) = \text{AIF}(t)$, a positive coefficient $\beta(\mathbf{x})$ is related to a fraction of the contrast agent entering the voxel $\mathbf{x}$ and δ is the time delay that can be easily estimated from data. The function of interest is $f(z,\mathbf{x}) = \beta(\mathbf{x})(1-F(z,\mathbf{x}))$.

Motivation



- The DCE imaging experiments can be described by a collection of Laplace convolution equations based on noisy observations, one equation per unit volume (voxel)
- $Y(t, \mathbf{x}) = \int_0^{t-\delta} g(t-z)\,\beta(\mathbf{x})(1 - F(z, \mathbf{x}))dz + \varepsilon\xi(t, \mathbf{x})$

Motivation

- $Y(t, \mathbf{x}) = \int_0^{t-\delta} g(t-z)\, \beta(\mathbf{x})(1 - F(z, \mathbf{x}))dz + \varepsilon\xi(t, \mathbf{x})$
- At Present due to high level of noise in the left hand side of above equation, the curves for each voxel $\mathbf{x}$ are roughly clustered and averaged
- For each of the clusters, the problem appears as location-independent version
- Analysis is done for the secondary data
- It is impossible to accurately assess the clustering errors,the results are unreliable in estimation errors

Our Method

- Our objective is solve the functional Laplace deconvolution problem directly.
- We assume that for each voxel "x",function $f(z, \mathbf{x})$ is a smooth in t moreover for each t it is piecewise smooth.
- We assume that the unknown function belongs to an anisotropic Laguerre-Sobolev space and recover it using a combination of wavelet and Laguerre functions expansion.
- After time measurement are appropriately shifted, use $\delta = 0$
- $Y(t, \mathbf{x}) = q(t, \mathbf{x}) + \varepsilon \xi(t, \mathbf{x}) \quad \text{with} \quad q(t, \mathbf{x}) = \int_0^t g(t - z) f(z, \mathbf{x}) dz.$

Estimation Algorithm

- $$Y(t, \mathbf{x}) = q(t, \mathbf{x}) + \varepsilon \xi(t, \mathbf{x}) \quad \text{with} \quad q(t, \mathbf{x}) = \int_0^t g(t-z) f(z, \mathbf{x}) dz. \quad (2)$$

  where $\mathbf{x} = (x_1, x_2)$, $(t, x_1, x_2) \in U = [0, \infty) \times [0, 1] \times [0, 1]$ and $\xi(z, x_1, x_2)$ is the three-dimensional Gaussian white noise such that

  $$\text{Cov}\,(\xi(z_1, x_{11}, x_{12}), \xi(z_2, x_{21}, x_{22})) = \mathbb{I}(z_1 = z_2)\,\mathbb{I}(x_{11} = x_{21})\,\mathbb{I}(x_{12} = x_{22}).$$

- Choose finitely supported periodized r_0-regular wavelet basis (e.g., Daubechies) $\psi_{j,k}(x)$. On $[0, 1]$, form a product wavelet basis $\Psi_{\boldsymbol{\omega}}(\mathbf{x}) = \psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2)$ on $[0, 1] \times [0, 1]$ where $\boldsymbol{\omega} \in \Omega$ with
- $\Omega = \left\{ \boldsymbol{\omega} = (j_1, k_1; j_2, k_2) : j_1, j_2 = 0, \cdots, \infty;\ k_1 = 0, \cdots, 2^{j_1 - 1},\ k_2 = 0, \cdots, 2^{j_2 - 1} \right\}.$

Estimation Algorithm

- Obtain functional wavelet coefficients of $f(t, \mathbf{x})$, $q(t, \mathbf{x})$, $Y(t, \mathbf{x})$ and $\xi(t, \mathbf{x})$ by, respectively, $f_\omega(t)$, $q_\omega(t)$, $Y_\omega(t)$ and $\xi_\omega(t)$. Then, for any $t \in [0, \infty)$
- $Y_\omega(t) = q_\omega(t) + \varepsilon \xi_\omega(t)$ with $q_\omega(t) = \int_0^t g(t-s) f_\omega(s) ds$
- The function $f(t, \mathbf{x})$ can be written as $f(t, \mathbf{x}) = \sum_{\omega \in \Omega} f_\omega(s) \Psi_\omega(\mathbf{x})$, $\mathbf{x} = (x_1, x_2)$, $\omega = (j_1, j_2, k_1, k_2)$

Estimation Algorithm

- Consider the orthonormal basis $\varphi_l(t)$ $l = 0, 1, 2, \ldots,$ that consists of a system of Laguerre functions:
  $\varphi_l(t) = e^{-t/2} L_l(t), \quad l = 0, 1, 2, \ldots,$
  where $L_k(t)$ are Laguerre polynomials $L_k(t) = \sum_{j=0}^{k} (-1)^j \binom{k}{j} \frac{t^j}{j!}, \quad t \geq 0.$
- Expand the wavelet coefficients $f_\omega(\cdot)$, $Y_\omega(\cdot)$, $q_\omega(\cdot)$ and kernal $g(\cdot)$ over the Laguerre basis, and obtained the coefficients $\theta_{l;\omega}$ for $f_\omega(\cdot)$, $Y_{l;\omega}$ for $Y_\omega(\cdot)$, $q_{l;\omega}$ for $q_\omega(\cdot)$ and g_l for $g(\cdot)$, $l = 1, \ldots, \infty$
- Plugging these expansions into formula (2), obtain the following equation
  $\sum_{l=0}^{\infty} q_{l;\omega}\, \varphi_l(t) = \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \theta_{l;\omega}\, g_k \int_0^t \varphi_k(t-s)\varphi_l(s)ds.$

Estimation Algorithm

- Use formula

$$\int_0^t \phi_k(x)\phi_j(t-x)dx = e^{-t/2}\int_0^t L_k(x)L_j(t-x)dx = \phi_{k+j}(t) - \phi_{k+j+1}(t).$$

- Rewrite the equation above

$$\sum_{k=0}^{\infty} q_{k;\omega}\,\varphi_k(t) = \sum_{k=0}^{\infty}\left[\theta_{k;\omega} g_0 + \sum_{l=0}^{k-1}(g_{k-l} - g_{k-l-1})f_{l;\omega}\right]\varphi_k(t).$$

- Equating coefficients for each basis function, obtain an infinite triangular system of linear equations.

Estimation Algorithm

- Cut Laguerre expansions to M coefficients
- Obtain the following expansions for f and q:

$$f_M(t,\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega} \sum_{l=0}^{M-1} \theta_{l;\boldsymbol{\omega}} \varphi_l(t) \Psi_{\boldsymbol{\omega}}(\mathbf{x}), \quad q_M(t,\mathbf{x}) = \sum_{\boldsymbol{\omega}\in\Omega} \sum_{l=0}^{M-1} q_{l;\boldsymbol{\omega}} \varphi_l(t) \Psi_{\boldsymbol{\omega}}(\mathbf{x}).$$

- We need to estimate coefficients $\theta_{l;\boldsymbol{\omega}}$ and we need to decide which of the estimated coefficients to "keep" or "kill"

Estimation Algorithm

- Let $\boldsymbol{\theta}_{\omega}^{(M)}$, $\mathbf{g}^{(M)}$ and $\mathbf{q}_{\omega}^{(M)}$ be M-dimensional vectors with elements $\boldsymbol{\theta}_{l;\boldsymbol{\omega}}$, g_l and $q_{l;\boldsymbol{\omega}}$, $l = 0, 1, \ldots, M-1$, respectively.
- Then, for any M and any $\boldsymbol{\omega} \in \Omega$, one has $\mathbf{q}_{\omega}^{(M)} = \mathbf{G}^{(M)} \boldsymbol{\theta}_{\omega}^{(M)}$ where $\mathbf{G}^{(M)}$ is the lower triangular Toeplitz matrix with elements $G_{i,j}^{(M)}$, $0 \leq i, j \leq M-1$
- $G_{i,j}^{(M)} = \begin{cases} g_0, & \text{if } i = j, \\ (g^{(i-j)} - g^{(i-j-1)}), & \text{if } j < i, \\ 0, & \text{if } j > i. \end{cases}$

Estimation Algorithm

- We estimate coefficients $q_{l;\boldsymbol{\omega}}$ by
  $\widehat{q}_{l;\boldsymbol{\omega}} = \int_0^\infty Y_{\boldsymbol{\omega}}(t)\,\varphi_l(t)\,dt, \quad l = 0, 2, \ldots,$
- Obtain an estimator $\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}}$ of vector $\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}$ of the form $\widehat{\boldsymbol{\theta}_{\boldsymbol{\omega}}^{(M)}} = (\mathbf{G}^{(M)})^{-1}\widehat{\mathbf{q}_{\boldsymbol{\omega}}^{(M)}}$.
- Finally, construct a hard thresholding estimator for the function $f(t, \mathbf{x})$ as
  $\widehat{f}(t, \mathbf{x}) = \sum_{l=0}^{M-1} \sum_{\boldsymbol{\omega} \in \Omega(J_1, J_2)} \widehat{\theta}_{l;\boldsymbol{\omega}}\, \mathbb{I}\left(|\widehat{\theta}_{l;\boldsymbol{\omega}}| > \lambda_{l,\varepsilon}\right) \varphi_l(t)\, \Psi_{\boldsymbol{\omega}}(\mathbf{x})$

Estimation Algorithm

Denote by $\Omega(J_1, J_2)$ a truncation of a set Ω

$$\Omega(J_1, J_2) = \left\{\omega = (j_1, k_1; j_2, k_2) : 0 \le j_i \le J_i - 1,\ k_i = 0, \cdots, 2^{j_i - 1}; i = 1, 2\right\}.$$

We choose J_1, J_2, M and $\lambda_{l,\varepsilon}$ such that

$$2^{J_1} = 2^{J_2} = A^2 \varepsilon^{-2}, \quad M = \max\{m \ge 1 : ||(\mathbf{G}^{(m)})^{-1}|| \le \varepsilon^{-2}\},$$

and thresholds $\lambda_{l,\varepsilon}$ of the forms

$$\lambda_{l,\varepsilon} = 2\varepsilon \sqrt{2\,\nu \log(\varepsilon^{-1})\ l^{-1}}||(\mathbf{G}^{(l)})^{-1}||,$$

where ν is a large enough constant ($\nu = 8$)

Estimation error Assumption

- In order to evaluate the precision of the estimator $\widehat{f}(t, \mathbf{x})$, we need some assumptions on the function g.
- Let $r \geq 1$ be such that

$$\left.\frac{d^j g(t)}{dt^j}\right|_{t=0} = \begin{cases} 0, & \text{if } j = 0, \dots, r-2, \\ B_r \neq 0, & \text{if } j = r-1, \end{cases}$$

  where $g(0) = B_1 \neq 0$ for $r = 1$.
- We assume that function $g(x)$ and its Laplace transform $G(s) = \int_0^\infty e^{-sx} g(x) dx$ satisfy the following conditions:
- **Assumption A1.** $g \in L_1[0, \infty)$ is r times differentiable with $g^{(r)} \in L_1[0, \infty)$.
- **Assumption A2.** Laplace transform $G(s)$ of g has no zeros with non-negative real parts except for zeros of the form $s = \infty + ib$.

Degree of ill-posedness

Lemma

(Lemma 4, Comte *et al.* **(2017)), Lemma 5.4, Vareschi (2015).**
Let conditions A1 and A2 hold. Denote the elements of the last row of matrix $(\mathbf{G}^{(m)})^{-1}$ *by* υ_j, $j = 1, \cdots, m$. *Then, there exist absolute positive constants* C_{G1}, C_{G2}, $C_{\upsilon 1}$ *and* $C_{\upsilon 2}$ *independent of m such that*

$$C_{G1} m^{2r} \leq ||(\mathbf{G}^{(m)})^{-1}||^2 \leq ||(\mathbf{G}^{(m)})^{-1}||_F^2 \leq C_{G2} m^{2r},$$

$$C_{v_1} m^{2r-1} \leq \sum_{j=1}^{m} \upsilon_j^2 \leq C_{\upsilon 2} m^{2r-1}.$$

- The lemma shows that the degree of ill-posedness is r, the eigenvalues of the inverse matrix $(\mathbf{G}^{(m)})^{-1}$ grow as m^r.
- We do not know r and are not planning to use it

Laguerre-Sobolev ball

- We assume that the unknown function f belongs to the generalized three-dimensional Laguerre-Sobolev ball of radius A, characterized by its wavelet-Laguerre coefficients $\theta_{l;\boldsymbol{\omega}} = \theta_{l;j_1,j_2,k_1,k_2}$ as follows:

$$\mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A) = \left\{ f : \sum_{l=0}^{\infty}\sum_{j_1=0}^{\infty}\sum_{j_2=0}^{\infty} 2^{2js_1+2j's_2}(l\vee 1)^{2s_3} e^{2\gamma\, l^{\beta}} \sum_{k_1=0}^{2^{j_1}-1}\sum_{k_2=0}^{2^{j_2}-1} \theta^2_{l;\boldsymbol{\omega}} \le A^2 \right\},$$

where we assume that $\beta = 0$ if $\gamma = 0$ and $\beta > 0$ if $\gamma > 0$.

Minimax upper bounds for the risk

In order to construct upper bounds for the risk , we define the maximum risk of an estimator $\widehat{f}$ over a set V as

$$R_\varepsilon(V, \widehat{f}) = \sup_{f \in V} \mathbb{E}||\widehat{f} - f||^2$$

Denote

$$\Delta = \begin{cases} A^2 \ \left[A^{-2}\varepsilon^2\right]^{\frac{2s_1}{2s_1+1}}, & \text{if } \ s_1 \leq \min(s_2, s_3/(2r)), \ \gamma = \beta = 0 \\ A^2 \ \left[A^{-2}\varepsilon^2\right]^{\frac{2s_2}{2s_2+1}}, & \text{if } \ s_2 \leq \min(s_1, s_3/(2r)), \ \gamma = \beta = 0 \\ A^2 \ \left[A^{-2}\varepsilon^2\right]^{\frac{2s_3}{2s_3+2r}}, & \text{if } \ s_3 \leq \min(2rs_1, 2rs_2), \ \gamma = \beta = 0 \\ A^2 \ \left[A^{-2}\varepsilon^2\right]^{\frac{2s_1}{2s_1+1}}, & \text{if } \ s_1 \leq s_2, \ \gamma > 0, \beta > 0 \\ A^2 \ \left[A^{-2}\varepsilon^2\right]^{\frac{2s_2}{2s_2+1}}, & \text{if } \ s_2 \leq s_1, \ \gamma > 0, \beta > 0 \end{cases}$$

Minimax upper bound for the risk

Theorem

Let $\min\{s_1, s_2\} \geq 1/2$ *and* $s_3 \geq 1/2$ *if* $\gamma = \beta = 0$. *Let* $\widehat{f}(t, \mathbf{x})$ *be the wavelet-Laguerre estimator of* f. *If* $\nu \geq 12 C_{\upsilon 2}/C_{G1}$, *then, under Assumptions A1 and A2, if* ε *is small enough, for some absolute constant* $\overline{C} > 0$ *independent of* ε, *one has*

$$R_\varepsilon(B^{s_1,s_2,s_3}_{\gamma,\beta}(A), \widehat{f}) \leq \overline{C}\, \Delta\, [\log(1/\varepsilon)]^d$$

d depends on the parameters of the Laguerre-Sobolev ball and on r,$d \leq 3$.

Minimax lower bounds for the risk

In order to ensure that the estimator $\widehat{f}$ is asymptotically optimal, we evaluate the minimax lower bounds for the risk of any estimator $\tilde{f}$ over the Sobolev ball:

$$R_\varepsilon(\mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A)) = \inf_{\tilde{f}} \sup_{f \in \mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A)} \mathbb{E}||\tilde{f} - f||^2$$

Theorem

Let $\min\{s_1, s_2\} \geq 1/2$ *and* $s_3 \geq 1/2$ *if* $\gamma = \beta = 0$. *Then, if* ε *is small enough, under Assumptions A1 and A2, for some absolute constant* $\underline{C} > 0$ *independent of* ε*, one has*

$$R_\varepsilon(\mathcal{B}^{s_1,s_2,s_3}_{\gamma,\beta}(A)) \geq \underline{C}\,\Delta$$

where Δ *was defined before*

Simulation Studies

- Consider $Y(t, \mathbf{x}) = q(t, \mathbf{x}) + \varepsilon\xi(t, \mathbf{x})$ with $q(t, \mathbf{x}) = \int_0^t g(t-z)f(z, \mathbf{x})dz$.
- We considered and n equally spaced observations on the time interval $[0; T]$
- For each test function $f(t, \mathbf{x})$ and a kernel $g(t)$, we obtained exact values of $q(t, \mathbf{x})$ in the equation above equation.

Simulation Studies

- We created an uniform grid $\{x_{1,i}, x_{2,j}\}$ with $i = 1, \cdots, n_1$ and $j = 1, \cdots, n_2$, and obtained the three-dimensional array of values $q(x_{1,i}, x_{2,j}, t_k)$, $k = 1, \cdots, n$
- Finally, we obtained a sample $Y_{i,j,k}$ of the left-hand side of the equation above by adding independent Gaussian $\mathbb{N}(0, \sigma^2)$ noise to each value $q(x_{1,i}, x_{2,j}, t_k)$, $i = 1, \cdots, n_1$, $j = 1, \cdots, n_2$, $k = 1, \cdots, n$.

Simulation Studies

- We constructed a system of M Laguerre functions obtained an estimator

$$\widehat{f}(t,\mathbf{x}) = \sum_{l=0}^{M-1} \sum_{\boldsymbol{\omega}\in\Omega(J_1,J_2)} \widehat{\theta}_{l;\boldsymbol{\omega}}\, \mathbb{I}\left(|\widehat{\theta}_{l;\boldsymbol{\omega}}| > \lambda_{l,\varepsilon}\right) \varphi_l(t)\, \Psi_{\boldsymbol{\omega}}(\mathbf{x})$$

with the thresholds $\lambda_{l,\widehat{\varepsilon}}$, $l = 0, \cdots, M-1$, given by

$$\lambda_{l,\varepsilon} = 2\varepsilon \sqrt{2\,\nu \log(\varepsilon^{-1})\; l^{-1}}||(\mathbf{G}^{(l)})^{-1}||,$$

- $\widehat{\varepsilon}$ is estimated using the standard deviations of the wavelet coefficients at the highest resolution level.
- The precision of the estimator $\widehat{f}$ was measured by the relative risk $\Delta(\widehat{f}) = ||\widehat{f} - f||_2/||f||_2$.

Simulation Studies

We used $n_1 = n_2 = n = 32$, $M = 8$ and $T = 5$, and choose $g(x) = \exp(-x/2)$ and four test functions for f:

$f_1(t, \mathbf{x}) = t\, e^{-t}(x_1 - 0.5)^2)\,(x_2 - 0.5)^2$, $f_2(t, \mathbf{x}) = e^{-t/2} \cos(2\pi x_1 x_2)$,

$f_3(t, \mathbf{x}) = t\, e^{-t}(x_1 - 0.5)^2)\,(x_2 - 0.5)^2 + e^{-t/2} \cos(2\pi x_1 x_2)$

$f_4(t, \mathbf{x}) = e^{-t/2} \cos(2\pi x_1 x_2) + (x_1 - 0.5)^2)\,(x_2 - 0.5)^2$

Function	SNR=3	SNR=5	SNR=7
$f_1(t, \mathbf{x})$	0.1107 (0.0110)	0.0694 (0.0066)	0.0511 (0.0049)
$f_2(t, \mathbf{x})$	0.1224 (0.0100)	0.0761 (0.0071)	0.0567 (0.0051)
$f_3(t, \mathbf{x})$	0.1107 (0.0112)	0.0680 (0.0068)	0.0511 (0.0048)
$f_4(t, \mathbf{x})$	0.1080 (0.0117)	0.0690 (0.0058)	0.0519 (0.0046)

Table: The average values of the relative errors (with the standard errors of the means in parentheses) evaluated over 100 simulation runs.

Real Data example

- CE-CT (Computerized Tomography) images of a participant of the REMISCAN cohort study , who underwent anti-angiogenic treatment for renal cancer.
- The data consist of the arterial images and images of the area of interest (AOI) at 37 time points over approximately 4.6 minute interval.
- The first 15 time points (approximately the first 30 seconds) correspond to the time period before the contrast agent reached aorta, so we just used them for the evaluation of the base intensity.

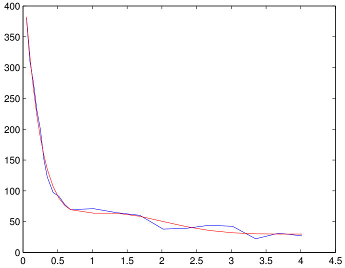
Real Data example

- Since the images of the aorta are extremely noisy, we evaluated the average values of the grey level intensity at each time point and then used Laguerre functions smoothing in order to obtain AIF
- The images of AOI contain 49×38 pixels.

Real Data example

- Since our technique is based on periodic wavelets and hence application of the method to a non-periodic function is likely to produce Gibbs effects, we cut the image to the size of 32×32 pixels.
- We obtained symmetric versions of the the images (reflecting the images over the two sides) and applied our methodology to the resulting spatially periodic functions.
- Consequently, the estimator obtained by the technique is spatially symmetric, so we record only the original part as the estimator $\widehat{f}$.

Real Data example



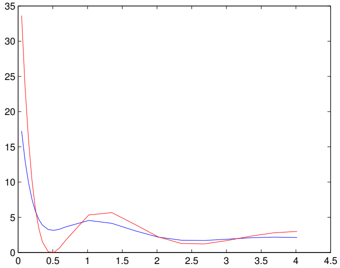
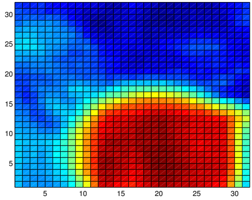
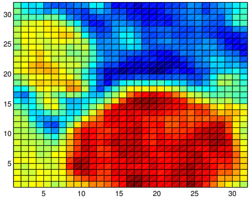


Figure: Left: the averages of the aorta intensities (blue) and the estimated Arterial Input Function AIF(t) (red). Right: two curves for distinct spatial locations.

Real Data example





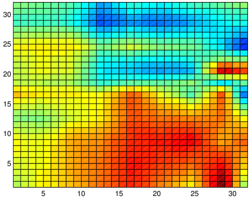


Figure: The values of $\widehat{f}$ at 34 seconds (corresponds to the first time point), 95 seconds (the 12th time point) and 275 seconds (the last time point).

Thank you !!

(udara@knights.ucf.edu)