Input Estimation from Sparse Measurements in LPV Systems and Isotopic Ratios in Polar Firns

Emmanuel Witrant, Patricia Martinerie

5th Symposium on System Structure and Control
Part of 2013 IFAC Joint Conference SSSC
Grenoble, France, February 4-6, 2013

Input Estimation from Sparse Measurements in LPV Systems and Isotopic Ratios in Polar Firns ⋆

Emmanuel WITRANT * **Patricia MARTINERIE** **

* *UJF-Grenoble 1/CNRS, Grenoble Image Parole Signal Automatique (GIPSA-lab), UMR 5216, B.P. 46, 38402 St Martin d'Hères, France (e-mail: emmanuel.witrant@ujf-grenoble.fr)*
** *UJF-Grenoble 1/CNRS, Laboratoire de Glaciologie et Géophysique de l'Environnement (LGGE) UMR 5183, Grenoble, 38041, France (e-mail: patricia@lgge.obs.ujf-grenoble.fr)*

Abstract: We adress the problem of inverse input reconstruction for linear parameter-varying (LPV) systems when only a limited amount of data (i.e. sparse measurements at final time) is available. We include the LPV property by deriving a time-varying Green's function that models the input/output behavior. The estimation is achieved by solving a least-squares optimization problem parameterized in terms of the input rugosity (regularization term) to take into account the under-constrained nature of the problem. Several automatic tuning methods for the rugosity are described, based on stochastic analysis of the data. A new LPV model is derived for the isotopic ratio of chemical species and our results are applied to the atmospheric history reconstruction of trace gases from polar firn measurements, a problem of prime interest in the environmental sciences community.

Keywords: LPV systems, input estimation, parameter optimization, environmental sciences, air pollution.

1. INTRODUCTION

Linear parameter-varying (LPV) models represent a wide class of systems. They are extensively used to capture the complexity of coupled dynamics, partially modeled transient phenomena and/or changes of operating points. Estimating an unknown input from a limited set of measurements is a challenging topic in inverse problems studies and the inclusion of LPV dynamics can be of prime interest for a large class of applications. While numerous identification methods have been proposed for estimating the varying parameters, the state or the LPV model architecture, few results (such as Kulcsár et al., 2010) aim at reconstructing the input variations. We further complicate the problem by considering that only the measurements at final time are available to compute the estimate. This approach is motivated by the analysis of processes for which the measurements are particularly difficult to obtain while an LPV model can be provided by a detailed physical analysis.

Isotopic ratios are widely used in environmental studies, especially as process tracers of geochemical cycles (e.g. Hoefs, 2009). They are measured with high precision by mass spectroscopy and usually expressed as a deviation with respect to a reference material ("delta unit") in per mil (‰):

$$\delta^{min}X = 1000 \times \left(\frac{[^{min}X]/[^{maj}X]}{R_{std}} - 1 \right) \quad (1)$$

where $[^{min}X]$ and $[^{maj}X]$ represent the concentrations of the minor and major isotopes, respectively, and R_{std} is the $[^{min}X]/[^{maj}X]$ ratio of a standard. Isotopic ratios are usually measured with a much better precision than concentrations of the individual isotopes ($[^{maj}X]$ and $[^{min}X]$). The specific application of isotopic geochemistry developed in this study is atmospheric time trends reconstruction from isotopic ratios in polar firns. Atmospheric trace gases are transported down slowly (in 15 to 100 years) in the interstitial air channels of polar firn (compacted snow). This process is ended at about 50 to 120 meters depth, where closed bubbles are formed and the air is no longer in contact with the atmosphere (e.g. Witrant and Martinerie, 2010; Buizert et al., 2012). Numerous air pumping operations at different depths (about 10 to 30 levels) in polar firns have been performed (e.g. Witrant et al., 2012), providing sparse final time measurements from which historical changes in atmospheric concentrations are estimated (e.g. Laube et al., 2010, 2012; Sturges et al., 2012, and references therein). In the case of isotopic ratios, the use of a linear inverse technique (Rommelaere et al., 1997) have been thought impossible due to nonlinearities associated with modeling two isotopes simultaneously (Sowers et al., 2005). Approximate methods based on separating the effects of the two isotopes (Trudinger et al., 1997; Sapart et al., 2012; Wang et al., 2012) or Monte-Carlo like testing

⋆ This work was partially supported by the CNRS/INSU LEFE program.

of numerous atmospheric scenarios (Bräunlich et al., 2001; Sowers et al., 2005) have been developed.

The paper is organized as follows. The inverse input estimation problem for LPV systems is formulated in Section 2. A specific input/output modeling approach based on Green's function is proposed in Section 3, where we include the time-varying property of the system. The sparsity of the measurements and their inclusion in an optimal input design scheme is considered in Section 4, where the solution of the inverse problem is parameterized in terms of the input rugosity as a tuning parameter. Several solutions to automatically tune this parameter are described in Section 5. Finally, our results are applied to finding the history of $\delta^{13}CH_4$ from polar firn measurements at DE08 (an ice core obtained at Law Dome, East Antarctica, at 66°43'S, 113°12'E) in Section 6.

2. PROBLEM FORMULATION

Consider the class of single-input multiple-outputs LPV systems that writes (in the discrete form) as, for $k = 1, \ldots, N_t$:

$$x_k = A_{D,k}x_{k-1} + B_{D,k}u_k + w_k, \quad x_{k=0} = x_0 \tag{2}$$

where $x \in \mathbb{R}^N$ is the state, $u \in \mathbb{R}$ is the unknown input, $w \in \mathbb{R}^N$ is an exogenous input, x_0 is the initial condition, and $A_{D,k}$ and $B_{D,k}$ are time-varying discrete state-space matrices of appropriate dimension. The sparsity of the measurements is considered in our case with the system output as:

$$y_{N_t} = Cx_{N_t} \quad \in \mathbb{R}^M \tag{3}$$

where C is the state to output matrix and M is the number of measurements. We thus consider that only the measurement of specific states at the final time is available.

Supposing that the time-evolution of $A_{D,k}$, $B_{D,k}$ and w_k is provided by an appropriate physical model and that the initial condition x_0 is known, our goal is to estimate the input history (u_k for $k = 1, \ldots, N_t$) from the sparse set of measurements y_m. This estimation problem can be formulated finding the optimal input history that minimizes the modeling error:

$$\epsilon \doteq y_m - y_{N_t}(u) \tag{4}$$

in the $\mathcal{L}_2$ sense. As such problem is clearly underconstrained, a specific care has to be taken for the choice of an appropriate regularization term and on the use of stochastic information on the measurements.

We also consider the case where the same input can affect different processes for which the measurements are taken at different final times (as it is the case in the firn problem where the same atmospheric history can be reconstructed from different polar sites). In this case, the dynamics and measurements are modeled as:

$$x_{i,k} = A_{Di,k}x_{i,k-1} + B_{Di,k}u_k + w_{i,k}, \quad x_{i,k=0} = x_{i,0} \tag{5}$$

$$y_{i,N_{ti}} = C_i x_{i,N_{ti}} \quad \in \mathbb{R}^{M_i} \tag{6}$$

where $i = 1 \ldots N_{proc}$ denotes the process i, and N_{ti} and M_i are the associated numbers of sampling times and measurements, respectively. Note that the same initial date has been chosen for the different dynamics, to simplify the notations as there is no technical difficulty to extend the proposed results to varying initial conditions.

Finally, as the initial condition is typically unknown for the processes for which only the final time measurement is available, we consider the steady-state approximation:

$$x_{i,0} \approx (I - A_{Di,0})^{-1}(B_{Di,0}u_0 + w_{i,0}) \tag{7}$$

For the polar firn example, this hypothesis provides a reasonable initial distribution for trace gases that exerienced transients later than the oldest age content in firn air and a reasonable order of magnitude otherwise.

3. SYSTEM MODELING USING GREEN'S FUNCTION

Green's function can be used to solve inhomogeneous differential equations with specific boundary conditions. For linear time-invariant (LTI) systems, Green's functions are equivalent to the impulse response (Bayin, 2006) and can thus be determined experimentally by measuring the system response to an impulse. Such principle can also be used to obtain a numerical input-output mapping for complex models, supposing a dominant LTI behavior. If such mapping can be inverted, then the system inputs can be inferred from the measured outputs and the Green's function of the system.

3.1 Input estimation for time-invariant systems

The previous method for trace gas reconstruction (Rommelaere et al., 1997) relied on generating a Green's function from the impulse response of the numerical model. An analytical equivalent method can be described as follows (trace gas transport being described by LTI systems, contrarily to the isotopic ratios), using the solution of the state equation.

Consider the linear dynamics:

$$\dot{x} = Ax + Bu, \quad x(t_0) = x_0 \tag{8}$$

$$y = Cx \tag{9}$$

where $x \in \mathbb{R}^N$, $u \in \mathbb{R}$, $y \in \mathbb{R}^M$, and A, B, C are state-space matrices of appropriate dimension. The solution of the state-space equations writes as:

$$y(t) = Ce^{A(t-t_0)}x(t_0) + C\int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$$

Considering a sampling time t_s and piecewise continuous inputs $u(t)$ for $t \in [t_0, t_f]$, the discretized version is obtained as:

$$y(t_k) = Ce^{A(t_k-t_0)}x(t_0) + t_s C\sum_{i=0}^{k} e^{A(t_k-t_i)}Bu(t_i)$$

which equivalently writes in the matrix form:

$$y(t_k) = G_0(t_k)x(t_0) + G(t_k)U(t_k) \tag{10}$$

where:

$$G_0(t_k) \doteq Ce^{A(t_k-t_0)} \in \mathbb{R}^{M\times N}$$
$$G(t_k) \doteq t_s[Ce^{A(t_k-t_0)}B,\, Ce^{A(t_k-t_1)}B, \ldots, CB]$$
$$U(t_k) \doteq [u(t_0), \ldots, u(t_k)]^T \in \mathbb{R}^k$$

Note that $G(t_k) \in \mathbb{R}^{M\times k}$ corresponds to the Green's function of the LTI system.

Equation (10) describes a linear mapping (algebraic relationship) between the vector of inputs $U(t_k)$ and the outputs $y(t_k)$, it can be directly used in an appropriate optimization scheme to infer $U(t_k)$ from the measurements $y_m(t_k)$. As the size of U is larger than the size of y_m, the problem is underconstrained.

3.2 Green's function for time-varying discrete systems with exogenous inputs

Consider the class of systems described by (2). In this case (e.g. for $k = 1, 2, 3$):

$$\begin{aligned}
x_1 &= A_{D,1}x_0 + B_{D,1}u_1 + w_1\\
x_2 &= A_{D,2}A_{D,1}x_0 + A_{D,2}B_{D,1}u_1 + B_{D,2}u_2\\
&\quad + A_{D,2}w_1 + w_2\\
x_3 &= A_{D,3}A_{D,2}A_{D,1}x_0\\
&\quad + A_{D,3}A_{D,2}B_{D,1}u_1 + A_{D,3}B_{D,2}u_2 + B_{D,3}u_3\\
&\quad + A_{D,3}A_{D,2}w_1 + A_{D,3}w_2 + w_3
\end{aligned}$$

and we can show (recursively) that at time t_k:

$$\begin{aligned}
y_k &= C\prod_{j=1}^{k} A_{D,k}x_0 + C\sum_{i=1}^{k}\left(\prod_{j=i+1}^{k} A_{D,j}\right)B_{D,i}u_i\\
&\quad + C\sum_{i=1}^{k}\left(\prod_{j=i+1}^{k} A_{D,j}\right)w_i
\end{aligned}$$

Equivalently:

$$y_k = G_{0,k}x_0 + G_kU_k + W_k \tag{11}$$

with $U_k = [u_1, \ldots, u_k]^T$ and

$$G_{0,k} = C\prod_{j=1}^{k} A_{D,k}, \quad W_k = C\sum_{i=1}^{k}\left(\prod_{j=i+1}^{k} A_{D,j}\right)w_i$$
$$G_k = C\left[\left(\prod_{j=2}^{k} A_{D,j}\right)B_{D,1}, \left(\prod_{j=3}^{k} A_{D,j}\right)B_{D,2}, \ldots, B_{D,k}\right]$$

4. COST FUNCTION AND OPTIMAL DESIGN

In order to minimize the estimation error (4) we considered the regularized weighted least-squares cost function:

$$\begin{aligned}
J(u) \doteq \frac{1}{2}\sum_{i=1}^{N_{proc}} ||y_{i,N_{ti}}(u) - y_{m,i}||^2_{Q_i}\\
+\frac{1}{2}\int_{t_0}^{t_{f,max}} ||u''(t)||^2_R\, dt
\end{aligned} \tag{12}$$

where $t_{f,max} \doteq \max\{t_{f,1}, \ldots, t_{f,N_{proc}}\}$ is the latest measurement time (t_{fi} being the measurement date for process i), u'' is the second derivative of u and $||x||^2_Q = x^TQx$ is the weighted squared $\mathcal{L}_2$ norm. The involved weights are such that $Q_i = Q_i^T \geq 0\ \forall i$ and $R = R^T > 0$ (classical optimality conditions). The use of u'' as a regularization term allows us to weight the rugosity of the optimal solution, as proposed by Menke (1989) for solving geophysical problems or by Hastie et al. (2009) for statistical learning.

The set of admissible solutions is constrained by the discretized dynamics (5)-(6). The optimization problem is directly formulated as finding the optimal time evolution $u^*(t)$ for $t \in [t_0, t_f]$ such that $J(u)$ is minimized.

4.1 Discrete cost function

The regularization term is discretized using:

$$u''(t_k) \approx \frac{u(t_{k+1}) - 2u(t_k) + u(t_{k-1})}{t_s^2}$$

The boundaries are approximated supposing a linear second order behavior calculated from the first (last) two points:

$$\begin{aligned}
u''(t_0) &\approx 2u''(t_1) - u''(t_2)\\
u''(t_N) &\approx 2u''(t_{N-1}) - u''(t_{N-2})
\end{aligned}$$

The dynamics (11) with initial condition (7) implies that:

$$y_{i,N_{ti}} = \bar{G}_{0i,N_{ti}}(B_{Di,0}u_0 + w_{i,0}) + G_{i,N_{ti}}U_{N_{ti}} + W_{i,N_{ti}}$$

where $\bar{G}_{0i,N_{ti}} \doteq G_{0i,N_{ti}}(I - A_{Di,0})^{-1}$. The measurements are corrected from the known exogenous inputs by defining:

$$\bar{y}_{m,i} = y_{m,i} - \bar{G}_{0i,N_{ti}}w_{i,0} - W_{i,N_{ti}} \in \mathbb{R}^{M_i} \tag{13}$$

Note that, ideally (if $y_{i,N_{ti}} = y_{m,i}$) $\bar{y}_{m,i} = \bar{G}_{0i,N_{ti}}B_{Di,0}u_0 + G_{i,N_{ti}}U_{N_{ti}}$.

The continuous cost (12) can then be expressed in terms of the discretized inputs $\{u_0, U_k\}$ as:

$$\begin{aligned}
J(U) = \frac{1}{2}\sum_{i=1}^{N_{proc}} ||\bar{G}_{0i,N_{ti}}B_{Di,0}u_0 + G_{i,N_{ti}}U_{N_{ti}} - \bar{y}_{m,i}||^2_{Q_i}\\
+\frac{t_s}{2}||F\bar{U}||^2_R
\end{aligned} \tag{14}$$

where $\bar{U} \doteq [u_0\ U_{N_{\bar{t}}}^T]^T$, $N_{\bar{t}} = \max\{N_{t1}, \ldots, N_{tN_{proc}}\}$ and:

$$F = \frac{1}{t_s^2}\begin{bmatrix}
2 & -5 & 4 & -1 & 0 & \cdots & 0\\
1 & -2 & 1 & 0 & \cdots & \cdots & \vdots\\
0 & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots\\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\
\vdots & \ddots & \cdots & 0 & 1 & -2 & 1\\
0 & \cdots & 0 & -1 & 4 & -5 & 2
\end{bmatrix}$$

The different time horizons and the inclusion of the initial condition in the input vector are handled by defining:

$$\begin{aligned}
\mathcal{G}_i \doteq \bar{G}_{0i,N_{ti}}B_{Di,0} \times [1\ 0_{1\times N_{\bar{t}}}]\\
+G_{i,N_{ti}} \times [0_{N_{ti}\times 1}\ I_{N_{ti}}\ 0_{N_{ti}\times(N_{\bar{t}}-N_{ti})}]
\end{aligned}$$

where I_α is the identity matrix of size α and $0_{\alpha\times\beta}$ is a zero matrix of size $\alpha \times \beta$. The cost function (14) thus equivalently writes in the compact form:

$$J(\bar{U}) = \frac{1}{2}\sum_{i=1}^{N_{proc}} ||\mathcal{G}_i\bar{U} - \bar{y}_{m,i}||^2_{Q_i} + \frac{1}{2}||F\bar{U}||^2_R \qquad (15)$$

where the multiplication of the second term by t_s is included in R to simplify the notation.

4.2 Optimal input

The optimization objective is formulated as finding:

$$\bar{U}^*(t) = \arg\min_{\bar{U}} J(\bar{U})$$

and the necessary conditions for optimality are:

$$\frac{\partial J(\bar{U}^*)}{\partial \bar{U}} = 0 \quad \text{and} \quad \frac{\partial^2 J(\bar{U}^*)}{\partial \bar{U}^2} > 0$$

Applying the first order optimality condition on (15) gives:

$$\sum_{i=1}^{N_{proc}} [\mathcal{G}_i\bar{U}^* - \bar{y}_{m,i}]^T Q_i\mathcal{G}_i + \bar{U}^{*T}F^T RF = 0$$

$$\Leftrightarrow \bar{U}^* = \left[F^T RF + \sum_{i=1}^{N_{proc}} \mathcal{G}_i^T Q_i \mathcal{G}_i\right]^{-1} \sum_{i=1}^{N_{proc}} \mathcal{G}_i^T Q_i \bar{y}_i \qquad (16)$$

The second order condition for optimality is satisfied by construction as $\mathcal{G}_i^T Q_i \mathcal{G}_i + F^T RF > 0$ for the indicated constraints on W_i and R.

4.3 Weights selection

The weights on the estimation error are obtained from the standard deviations of the measurements $\sigma_i(j)$, where $i = 1, \ldots, N_{proc}$ is the process index and $j = 1, \ldots, M_i$ is the measurement index, as $Q_i = \text{diag}\left(1/\sigma_i^2(j)\right)$. The input weight is chosen as:

$$R = \kappa^2\, t_s \times I_{N_{\bar{t}}} \qquad (17)$$

where κ is the rugosity tuning parameter. Note that the optimal solution of (16) depends only on κ as a free parameter.

5. AUTOMATIC RUGOSITY TUNING

One of the main difficulties in estimating a relevant inverse input is related to the proper weighting between the least-squares terms, parameterized in our case with the rugosity penalty κ in (17). Several available methods based on stochastic analysis are briefly described in this section.

5.1 Data prediction versus model resolution

The underdetermined property of the problem can be addressed as a trade-off between perfect data prediction and perfect model resolution (Menke, 1989; Rommelaere et al., 1997). Based on this idea, an optimization criterion using root mean square deviations (rmsd) can be derived as follows.

Defining the aggregated Green's functions, weighted Green's functions and measured outputs as, respectively, $\bar{\mathcal{G}} \doteq [\mathcal{G}_1^T \ \ldots \ \mathcal{G}_{N_{proc}}^T]^T$, $\bar{\mathcal{G}Q} \doteq [\mathcal{G}_1^T Q_1 \ \ldots \ \mathcal{G}_{N_{proc}}^T Q_{N_{proc}}]$ and $\bar{Y} \doteq [\bar{y}_{m,1}^T \ \ldots \ \bar{y}_{m,N_{proc}}^T]^T$, the generalized inverse A_{GI} describes the data-to-input mapping as $\bar{U}^* = A_{GI}(\kappa)\bar{Y}$ and is computed from (16)-(17) as:

$$A_{GI}(\kappa) \doteq \left[\kappa^2 t_s F^T F + \bar{\mathcal{G}Q}\bar{\mathcal{G}}\right]^{-1} \bar{\mathcal{G}Q}$$

Similarly, defining the predicted ouputs vector as $Y \doteq [y_1^T \ \ldots \ y_{N_{proc}}^T]^T$, the mapping between data and predicted outputs is $Y = S(\kappa)\bar{Y}$ where $S(\kappa) \doteq \bar{\mathcal{G}}A_{GI}(\kappa)$. S is typically referred to as the *data resolution* or *smoother* matrix.

The data prediction error $\epsilon_Y = \bar{Y} - Y = (I - S)\bar{Y}$ is included in the optimization problem by considering its root mean square deviation. The model resolution is included using the covariance matrix $\text{cov}(\bar{U}^*) = A_{GI}\,\text{cov}(\bar{Y})A_{GI}^T$. This matrix determines the degree of error amplification induced by the mapping between data and model parameters (Menke, 1989). The diagonal elements of $\text{cov}(\bar{U}^*)$ measure the width of the data distribution and these elements can be mapped into the output space using the Green's function. The rugosity parameter is thus determined as:

$$\kappa^* = \min_{\kappa}\left\{rmsd(\bar{Y} - Y) + rmsd(\tilde{Y}_u)\right\} \qquad (18)$$

where $\tilde{Y}_u = \bar{\mathcal{G}}\sqrt{\text{diag}(\text{cov}(\bar{U}^*))}$ reflects the model resolution impact on the output.

5.2 Bias versus variance

The rugosity tuning can also be considered as a trade-off between bias and variance, measured by the cross-validation (CV) curve. Such tool is extensively used for model selection (e.g. see the survey by Arlot and Celisse, 2010) and can be calculated, in the generalized cross-validation (GCV) form as:

$$\text{GCV}(\kappa) = \frac{1}{N_{data}}\left(\frac{||(I - S(\kappa))\bar{Y}||}{\text{tr}(I - S(\kappa))/N_{data}}\right)^2$$

where N_{data} is the number of data points, $||\cdot||$ is the Euclidean norm on $\mathbb{R}^{N_{data}}$ and $\text{tr}(\cdot)$ is the trace of the matrix. The optimization problem is in this case to find κ^* that minimizes the GCV.

While CV methods typically necessitate a large data set, some recent results by Lukas (2006, 2008) allow for considering small data sets and for adjusting the robustness of the criterion. The robust GCV function is given by:

$$\text{RGCV}(\kappa) = \gamma\text{GCV}(\kappa) + (1 - \gamma)\mu(\kappa)\text{GCV}(\kappa) \qquad (19)$$

where $\mu(\kappa) \doteq \text{tr}(S(\kappa)^2)/N_{data}$ and $\gamma \in [0; 1]$ is the robustness parameter (small for more robust results). The problem of correlated data has been considered in (Lukas, 2008), where the RGCV is improved to the R_1GCV by setting $\mu(\kappa) = [\text{tr}(S(\kappa)) - \text{tr}(S(\kappa)^2)]/[N_{data}\kappa]$ in (19).

6. INVERSION FOR ISOTOPIC RATIOS

6.1 Isotopic model in δ units

Consider the concentration dynamics of two isotopes (x_{min} and x_{maj}) and their isotopic ratio r, for $k = 1, \ldots, N_t$:

$$x_{min,k} = A_{Dmin} x_{min,k-1} + B_{Dmin} x^{atm}_{min,k} \quad (20)$$

$$x_{maj,k} = A_{Dmaj} x_{maj,k-1} + B_{Dmaj} x^{atm}_{maj,k} \quad (21)$$

$$r_k = \mathrm{diag}(1/x_{maj,k}) x_{min,k} \quad (22)$$

where x_{min}, x_{maj} and $r \in \mathbb{R}^N$, N being the number of discretization depths used to simulate the reference model described in the Appendix, and the state-space matrices are given by an appropriate transport model (in our case by Witrant et al. (2012) for trace gas transport in firns). The system inputs are the atmospheric concentrations x^{atm}_{min} (unknown) and x^{atm}_{maj} (known). The initial condition of the major isotope is calculated according to (7) while the one of the minor isotope results from the optimization process. The time-evolution of x_{maj} is calculated using (21) and x^{atm}_{maj}. The LPV model of the ratio is obtained by combining (20) and (22) as:

$$\begin{aligned} r_k &= \mathrm{diag}(1/x_{maj,k}) \left(A_{Dmin}\, x_{min,k-1} + B_{Dmin}\, x^{atm}_{min,k} \right) \\ &= A_{D,k} r_{k-1} + \mathrm{diag}(1/x_{maj,k}) B_{Dmin}\, x^{atm}_{min,k} \end{aligned} \quad (23)$$

with $A_{D,k} \doteq \mathrm{diag}(1/x_{maj,k}) A_{Dmin}\, \mathrm{diag}(x_{maj,k-1})$.

From definition (1) we can use the changes of variables $x^{atm}_{min} = (\delta^{atm}/1000 + 1)\, R_{std}\, x^{atm}_{maj} \in \mathbb{R}$ and $\delta_k = (r_k/R_{std} - \mathbf{1}) \times 1000 \in \mathbb{R}^N$ to get, using (23):

$$\begin{aligned} \frac{\delta_k}{1000} &= \frac{1}{R_{std}} \left(A_{D,k} r_{k-1} + \mathrm{diag}(1/x_{maj,k}) B_{Dmin}\, x^{atm}_{min,k} \right) - \mathbf{1} \\ &= A_{D,k} \left(\frac{\delta_{k-1}}{1000} + \mathbf{1} \right) + B_{D,k} \left(\frac{\delta^{atm}_k}{1000} + \mathbf{1} \right) - \mathbf{1} \end{aligned}$$

where $B_{D,k} \doteq \mathrm{diag}(x^{atm}_{maj,k}/x_{maj,k}) B_{Dmin}$ and $\mathbf{1} \in \mathbb{R}^{\mathbf{N}}$ is a column vector of ones. Hence:

$$\begin{aligned} \delta_k = {} & A_{D,k} \delta_{k-1} + B_{D,k} \delta^{atm}_k \\ & + 10^3 \left(A_{D,k} \times \mathbf{1} + B_{D,k} - \mathbf{1} \right) \end{aligned} \quad (24)$$

Supposing a steady-state initial condition, it follows that:

$$\begin{aligned} \delta_0 = {} & (I_N - A_{D,0})^{-1} B_{D,0}\, \delta^{atm}_0 \\ & + 1000 \left(I_N - A_{D,0} \right)^{-1} \left(A_{D,0} \times \mathbf{1} + B_{D,0} - \mathbf{1} \right) \end{aligned} \quad (25)$$

The δ isotopic ratio model (24)-(25) is thus a LPV system (the dominant isotope being the varying parameter) that belongs to the general class of systems (2)-(7). Note that a physical modeling approach based on the PDE transport model also leads to the same conclusion, as described in the Appendix.

6.2 $\delta^{13}CH_4$ atmospheric scenario estimation from firn air measurements at DE08

Our inverse scenario method is evaluated on the isotopic ratio $\delta^{13}\mathrm{CH}_4$ of $^{13}\mathrm{CH}_4$ (x_{min}) versus $^{12}\mathrm{CH}_4$ (x_{max}) at the DE08 polar site. Atmospheric measurements for $\delta^{13}\mathrm{CH}_4$ are available since 1978, as described by Francey et al. (1999), and we can thus evaluate the inverse method efficiency. DE08 is an ice core for which the measurements are particularly sparse. However it is considered as the most reliable site for isotopic scenario reconstruction because isotopic fractionation in firn is minimal (Sapart et al., 2012). Undergoing a high snow accumulation rate, the DE08 firn air covers a short period of time with a high time resolution (Witrant et al., 2012).

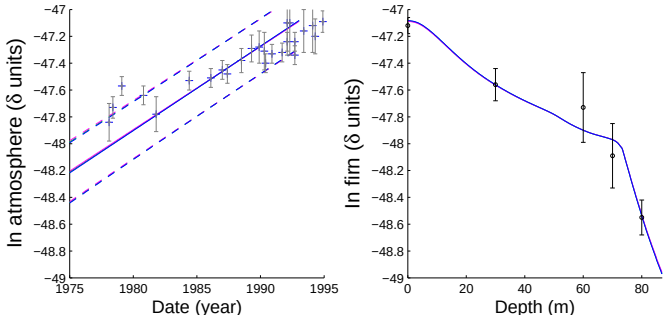

Fig. 1. Efficiency of the inverse scenario strategy for $\delta^{13}\mathrm{CH}_4$ at DE08 (all data are from Francey et al., 1999). Left pannel: estimated scenarios (plain lines), 2 σ limit (dashed lines) and atmospheric data ('+'). Right pannel: final time distributions (plain lines) and ice core measurements ('o')

The input estimation results are presented on Figure 1, where the estimated scenarios (top) and their corresponding final time distribution (bottom) are calculated for the different rugosity tuning strategies discussed in Section 5. The time range is set according to the drilling date (1993) and the gas transport time at DE08 ($\approx$ 20 years). For this specific isotopic ratio/ice core combination the tuning methods are consistent and provide almost identical results (superposed curves in the figure). While the final time distributions remain in the error bars, a strong weight has been set on the rugosity by the tuning methods (due to the very limited number of measurements). The optimal rugosity values for RMSD, GCV, RGCV and R1GCV are 2850, 3772, 3808 and 4444, respectively. The robustness weight in the RGCV is set as $\gamma = 7 \times 10^{-4}$ (based on entensive model validation on multiple sites and for multiple gases) and has to be increased up to 0.9 to observe an increased robustness (and hence an increased rugosity, which does not appear as a necessary issue in our case).

7. CONCLUSION

We formulated the input estimation problem for LPV systems as an inverse problem where the input/output relationship is described by a known time-varying mapping, formalized with a Green's function. This allowed us to find an analytical solution for optimal input estimation, in the least-squares sense, parameterized in terms of the input rugosity (regularization coefficient). Several stochastic methods were investigated to automatically tune this coefficient. A new LPV model was derived to describe the time-variation of isotopic ratio and our method was shown to effectively estimate the atmospheric history of chemical species from polar firn measurements. Future works will consider the accuracy and sensitivity of the automatic tuning methods when more measurements are available (e.g. in the multi-sites case). The impact of data calibration (which sometimes differ depending on the data) is also a topic of interest.

REFERENCES

Arlot, S. and Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79.

Bayin, S. (2006). *Mathematical Methods in Science and Engineering*. Wiley.

Bräunlich, M., Aballain, O., Marik, T., et al. (2001). Changes in the global atmospheric methane budget over the last decades inferred from ^{13}C and D isotopic analysis of antarctic firn air. *J. Geophys. Res.*, 106(D17), 20465–20481.

Buizert, C., Martinerie, P., Petrenko, V., et al. (2012). Multiple-tracer firn air transport characterisation and model intercomparison for neem, northern greenland. *Atmos. Chem. Phys.*, 12, 4259–4277. doi:10.5194/acp-12-4259-2012.

Francey, R.J., Manning, M.R., Allison, C.E., et al. (1999). A history of δ^{13}C in atmospheric methane from the cape grim air archive and antarctic firn air. *J. Geophys. Res.*, 104, 23 631–23 643.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Statistics. Springer, 2^{nd} edition.

Hoefs, J. (2009). *Stable isotope geochemistry*. Springer-Verlag, Berlin, Germany, 6^{th} edition.

Kulcsár, B., Bokor, J., and Shinar, J. (2010). Unknown input reconstruction for LPV systems. *Int. J. of Robust and Nonlinear Control*, 20(5), 579–595.

Laube, J.C., Hogan, C., Newland, M.J., et al. (2012). Distributions, long term trends and emissions of four perfluorocarbons in remote parts of the atmosphere and firn air. *Atmos. Chem. Phys.*, 12, 4081–4090. doi:10.5194/acp-12-4081-2012.

Laube, J., Martinerie, P., Witrant, E., et al. (2010). Accelerating growth of HFC-227ea (1,1,1,2,3,3,3-heptafluoropropane) in the atmosphere. *Atmos. Chem. Phys.*, 10, 5903–5910. doi:doi:10.5194/acp-10-5903-2010.

Lukas, M. (2006). Robust generalized cross-validation for choosing the regularization parameter. *Inverse Problems*, 22, 1883.

Lukas, M. (2008). Strong robust generalized cross-validation for choosing the regularization parameter. *Inverse Problems*, 24, 034006.

Menke, W. (1989). *Geophysical data analysis: Discrete inverse theory*. Academic Press, New York.

Rommelaere, V., Arnaud, L., and Barnola, J. (1997). Reconstructing recent atmospheric trace gas concentrations from polar firn and bubbly ice data by inverse methods. *J. Geophys. Res.*, 102(D25), 30069–30083.

Sapart, C.J., Martinerie, P., Chappellaz, J., et al. (2012). Reconstruction of the carbon isotopic composition of methane over the last 50 yr based on firn air measurements at 11 polar sites. *Atmos. Chem. Phys. Discuss.*, 12, 9587–9619. doi:10.5194/acpd-12-9587-2012.

Sowers, T., Bernard, S., Aballain, O., et al. (2005). Records of the δ^{13}C of atmospheric CH_4 over the last 2 centuries as recorded in antarctic snow and ice. *Global Biogeochem. Cy.*, 19, GB2002. doi:10.1029/2004GB002408.

Sturges, W.T., Oram, D.E., Laube, J.C., et al. (2012). Emissions halted of the potent greenhouse gas SF_5CF_3. *Atmos. Chem. Phys.*, 12, 3653–3658. doi:10.5194/acp-12-3653-2012.

Trudinger, C., Enting, L., Etheridge, D., et al. (1997). Modeling air movement and bubble trapping in firn. *J. Geophys. Res.*, 102(D6), 6747–6763.

Wang, Z., Chappellaz, J., Martinerie, P., et al. (2012). The isotopic record of northern hemisphere atmospheric carbon monoxide since 1950, implications for the CO budget. *Atmos. Chem. Phys.*, 12, 4365–4377. doi:10.5194/acp-12-4365-2012.

Witrant, E. and Martinerie, P. (2010). A variational approach for optimal diffusivity identification in firns. In *Proc. of the 18th Med. Conf. on Control and Automation*, 892–897. Marrakech, Morocco.

Witrant, E., Martinerie, P., Hogan, C., et al. (2012). A new multi-gas constrained model of trace gas non-homogeneous transport in firn: evaluation and behavior at eleven polar sites. *Atmos. Chem. Phys.*, 12, 11465–11483. doi:10.5194/acp-12-11465-2012.

Appendix A. ISOTOPIC RATIOS IN FIRN AIR

Considering two trace gases concentrations ρ_1 and ρ_2, the physical models are given by the trace gas model proposed in (Witrant et al., 2012) as, for $i = 1, 2$:

$$[\rho_i(z,t)f(z)]_t + \rho_i(z,t)\tau(z) + \Big[\rho_i(z,t)f(z)(v(z) + w_{air}(z)) - D_i(z)\left([\rho_i(z,t)]_z - \rho_i(z,t)\frac{M_i g}{RT}\right)\Big]_z = 0$$

where f is the volume fraction of open porosity, τ the rate of mass exchange between open and closed air networks, v the firn sinking velocity, w_{air} the air velocity, D_i the effective diffusivity, M_i the gas molar mass, g the gravitational acceleration, R the ideal gas constant and T the temperature. The δ ratio definition implies that:

$$\rho_1 = \left(\frac{\delta}{1000} + 1\right) R_{std}\, \rho_2$$

The resulting linear model in $\delta(z,t)$ is obtained as:

$$\begin{aligned}
&[\delta f]_t + \left(f(v + w_{air}) + \frac{D_2 M_2 g}{RT}\right)[\delta]_z - [D_2\delta_z]_z \\
&-2\frac{\rho_{2,z}}{\rho_2} D_2\delta_z \\
&+\frac{1}{\rho_2}\left[(D_2 - D_1)[\delta\rho_2]_z - (D_2M_2 - D_1M_1)\frac{g}{RT}\delta\rho_2\right]_z \\
&= -\frac{1000}{\rho_2}\left[(D_2 - D_1)[\rho_2]_z - (D_2M_2 - D_1M_1)\frac{g}{RT}\rho_2\right]_z
\end{aligned}$$

where ρ_2 acts as the varying parameter.

The upper boundary condition is $\delta(0,t) = \delta^{atm}(t)$. For the bottom boundary condition (at $z = z_f$), we start from the definition:

$$\delta = \left(\frac{\rho_1/\rho_2}{R_{std}} - 1\right) \times 1000$$

to derive:

$$\delta_z = \left[\frac{\rho_1/\rho_2}{R_{std}}\right]_z \times 1000 = \frac{g}{RT}(M_1 - M_2)(\delta + 1000)$$

Hence we have a non-homogeneous boundary condition:

$$\delta_z(z_f,t) - \frac{g}{RT}(M_1 - M_2)\,\delta(z_f,t) = 1000\frac{g}{RT}(M_1 - M_2)$$