Trajectory Prediction : a Functional Regression Approach

S. Puechmorel
ENAC Dept. MI LMA
7, Avenue Edouard Belin
31055 TOULOUSE FRANCE
email: puechmor@recherche.enac.fr

D. Delahaye
ENAC Dept. MI LMA
7, Avenue Edouard Belin
31055 TOULOUSE FRANCE
email: delahaye@recherche.enac.fr

L. Boussouf
INSA
157, Avenue de Rangueil
31400 TOULOUSE FRANCE
email: lboussou@etud.insa-toulouse.fr

***Index Terms*—Functional regression, Functional data, Trajectory Prediction**

***Abstract*—Accurate trajectory prediction is an important issue for decision support tools in the field of ATM. This paper presents a new approach that considers trajectories as points in a functional space. By finding a expansion of observed trajectories on a suitable basis and truncating the expansion to a finite number of terms, standard regression algorithms can be used. Within this framework, full segments of trajectories can be forecasted up to 10-15 minutes.**

I. Introduction

Functional data analysis is an active branch of statistics in which relevant objects are mappings belonging to a well defined space, most of the time an Hilbert space. It has been proven very efficient for problems where preserving the functional nature of data is of great importance : curves classification, functional dependence learning and similar problems [15]. In the recent literature an increasing attention has been paid to linear functional regression [2],[6] and some of its generalization [23],[14]. In this setting, either a scalar value or a mapping (the *response*), possibly contaminated by an independent measure noise is assumed to be linearly dependent on a mapping (the *predictor*). In the functional framework, the equivalent of the slope coefficient in the classical finite dimensional linear model is a kernel function that has to be estimated. Solving the associated least square problem leads to the well known Wiener-Hopf equation that generally admits no unique solution. One of the main issue in functional regression is thus to add some extra assumption on the regressor kernel so that the original ill-posed problem can be solved. On the other hand, in the field of ATM, the need for accurate trajectory predictor has appeared as a prerequisite for Decision Support Tools (DST). Air traffic management research and development has produced a substantial collection of decision support tools that provide automated conflict detection and resolution [4], [1], [22],trial planning [10], controller advisories for metering and sequencing [20], [3], traffic load forecasting [11], [9], weather impact assessment [8], [19], [5]. The ability to properly forecast future aircraft trajectories is central in many of those decision support tools. As a result, trajectory prediction (TP) and the treatment of trajectory prediction uncertainty continue as active areas of research and development (eg [17], [21], [12], [16], [18]. In this paper we will present an innovative approach based on functional regression for solving the short to mid-term trajectory prediction (TP) problem. Long-term prediction is yet beyond the scope of this study, but considering a database of trajectories and taking into account intents of aircraft through flight plans may allow an extension of our methodology to encompass it. The first part of the paper will be devoted to a short compendium of available trajectory goodness-of-fit metrics, then the main results on functional regression will be exposed, with the potential applications and improvements of existing algorithms for the specific trajectory prediction problem.

II. Trajectory prediction metrics

When an aircraft flies from a city A to a city B, it has to be managed by air traffic controllers in order to avoid collisions with others aircraft. Everyday, about 8000 aircraft fly in the French airspace, inducing a huge amount of control workload. Such a workload, is then spread by the mean of the airspace sectoring. The airspace is divided into geometrical sectors, each of them being assigned to a controller team. When a conflict between two (or more) aircraft is detected, the controller changes their routes (heading, speed, altitude) in order to keep a minimum distance between them during the crossing. All flying aircraft are then monitored during their navigation and so from the departure till the destination.

When a controller observes its traffic on the radar screen, he tries to identify convergent aircraft which may be in conflict in a near future, in order to apply manoeuvers that will separate them. The problem is to estimate where the aircraft will be located in this near future (5-10-20 minutes); this process is call trajectory prediction. This prediction may be also very useful in order to estimate the workload level in control sector to prevent over capacity event. As a mater of fact, it is very useful to estimate when an aircraft will enter a sector in order to compute the associated sector workload and to apply regulation if necessary. When a sector is expected to be overloaded, the aircraft involved in such a process will be speeded up or slow down by the controller in order to adapt the demand to the actual capacity as much as possible.

The trajectory prediction depends mainly on the residual noise after filtering which are the weight of the aircraft, the

ISBN: 978-0-615-20720-9

temperature and the wind. The residual noise is integrated with time with a growing covariance matrix indicating that the estimated position is less and less accurate. The weight of the aircraft is relevant in the flight dynamics model but is still a raw data. The engines of aircraft are sensitive to the air temperature and such a data is very useful to model the trust of the aircraft but it is also very difficult to measure on real time. Finally, the wind influences strongly the cinematic of the aircraft and limits also the trajectory prediction. Based on the available accuracy, the actual limit of the trajectory prediction is about 15 minutes for the conflict detection. It means that after 15 minutes the uncertainty is so big that the estimated position is no more useful for such application.

One of the issues in trajectory prediction is to measure how accurately a model will fit to a target trajectory. Unfortunately, many different metrics can be proposed, each of them focusing on a specific aspect of accuracy. Most of the time, the proposed metrics fall into one of these categories [13] :

- Time coincidence. The time difference between a predicted event and a real event is used as a measure of TP accuracy. Time coincidence is relevant in applications where synchronizing is important, like sequencing traffic, or when the DST uses time information to instruct controller about the order in which actions have to be taken.
- Spatial coincidence. Similar to the previous one except that spatial distance at specified time (or more generally at events that can be predicted with the knowledge of aircraft positions up to a given time) between the model and the real aircraft is computed. Spatial coincidence can be refined by further splitting into altitude and horizontal error. Furthermore, for some applications, mainly conflict predictors and/or solvers, spatial difference is projected onto a vector normal to the real trajectory (cross-track error) and onto a vector tangent to the real trajectory (along-track error).
- 4D coincidence. Trajectories are considered to be 4D curves, and distance between such curves is computed. Most of the metrics derived for spatial coincidence can be extended to the 4D setting, with the benefit of including a kind of time coincidence, thus generalizing in some sense the previous two aspects.
- Morphological similarity. Different in nature from the previous metrics, an intrinsic distance between trajectories considered as curves in a 3D space can be derived from Riemannian geometry. Since only the shape of the trajectory is taken into account, this metric is relevant mainly for trajectory design tools.

Except for the last one, all those basic metrics can be integrated along trajectories to produce a mean value indicator (the classical L^2 distance is for example obtained by integrating the standard spatial coincidence metric over time). Within the frame of functional regression, the standard choice is to consider L^2 distance as goodness-of-fit measure. In the following, we will use this spatial coincidence metric along with a specific 4D distance. Investigation of different kind of accuracy evaluations is planned for future work.

III. Functional regression

A. The functional nature of the trajectory prediction problem

An aircraft trajectory is by definition a mapping from a time interval $[a, b]$ to the space $\mathbb{R}^3$ (sometimes, it is convenient to add speed, so that the resulting expanded state space is $\mathbb{R}^6$). Such a trajectory is indeed the observed result of a complex evolution process that involves flight dynamics, external actions (pilot, ATC) and atmospheric factors (wind, temperature ...). The complete description of the trajectory using all these factors is generally not possible, because many influencing factors are unknown (aircraft mass, local wind, etc ...), so a less accurate but tractable model has to be chosen. For the purpose of short term prediction, a linear controlled model is accurate enough. The main assumption made is that the derivative of acceleration is zero (in a weak sense, since in most models commands are piecewise constant functions). Based on this observation, we will focus on trajectories belong to the Sobolev space of square integrable mappings with square integrable derivatives (in distributional sense) up to order 3. From now, we will assume that all trajectories belong to this space.

B. Linear regression with a functional predictor

The linear functional regression problem can be stated as follows :

- The predictor and the response are square integrable mappings from respective compact time intervals $[a, b]$, $[c, d]$ to $\mathbb{R}^n$ (resp. $\mathbb{R}^m$).
- The data set consists of pairs $(X_i, Y_i)_{i=1...N}$ of predictor/response. It is assumed that the X_i, Y_i are sample trajectories of two underlying smooth Hilbert random processes (for a general account on these processes, see [7]), with unknown smooth mean μ_X, μ_y and covariance kernels B_X, B_Y.
- The functional linear model on the predictor X has the general form :

$$\hat{Y} : t \mapsto f(t) + \int_{[a,b]} K(s,t)X(s)ds$$

  with $f : [c, d] \to \mathbb{R}^n$ a smooth square integrable mapping and $K : [a, d] \times [c, d] \to M_{m,n}(\mathbb{R})$ a smooth square integrable (m, n)-matrix valued kernel.
- The solution of the functional regression problem is the optimal couple (f, K) that minimize the mean square error between Y and $\hat{Y}$.

Most of the time, the related literature on the subject addresses the problem with $n = 1$, that is for real valued trajectories. In our setting, this clearly means that it enforces the fact that the x, y, z components of a trajectory can be treated as independents scalar valued mappings. Although we will see later that functional regression models satisfying some invariance properties must fall into this category, there is no reason to limit ourselves to kernels with values in the

ISBN: 978-0-615-20720-9

set of diagonal matrices. It worth notice too that the basic theory makes the assumption that trajectories are continuously observed, which is clearly not the case of ATC data. The reference [23] extends the least square criterion to irregularly spaced samples on predictor and response, providing us with the right framework for trajectory prediction applications and will be the starting point of our work. The solution of the functional regression problem is known to satisfy a Wiener-Hopf equation :

$$E[XY] = \int_{[a,b]} K(s,t)B_x(s,t)ds \quad (1)$$

Unfortunately, this equation has generally not a unique solution. Furthermore, solving 1 from sampled trajectories yields an even more ill-posed problem. Solutions to this problem, mainly by using regularization, have been proposed in [15]. Most of the time, an expansion of the predictor on an Hilbert basis is used to solve the functional regression (note that at some point of the process, all expansion must be truncated in order to obtain something computable). Several choices exist for such a basis. In many papers it is recommended to use Karhunen-Loeve expansion, with the eigenfunctions of the covariance operator as basis :

$$\psi_i(t) = \lambda_i \int_{[a,b]} B_X(s,t)\psi_i(s)ds$$

with λ_i the eigenvalue corresponding to ψ_i. Since the covariance and mean functions are unknown, there are to be estimated from the data. The procedure used in [23] is to use a weighted sum kernel approximation. A complete treatment can be found in [24]. It has to be noted that this particular choice is essentially heuristic : since the basis functions depend only on the predictor and not on the response, it cannot be guaranteed that the first q eigenfunctions, associated to the q largest eigenvalues, are the q most predictive (the integer q represents the truncation index used when solving the functional regression problem).

C. Solving the functional regression

The framework in this part will be the one chosen in [23]. Under the general assumptions of the previous section, we will further assume that the data set consists in a finite number of sparsely sampled predictor/response pairs. Let X_i (resp. Y_i) be the realization of predictor process X (resp. response process Y) corresponding to observation i in the data set. Let M_i (resp. N_i) be the number of samples available for this observation and let $X_{i,j}, j = 1 \dots M_i$ (resp. $Y_{i,j}, j = 1 \dots N_i$) be the actual samples along trajectories X_i (resp. Y_i) with corresponding sample times $\tau_{i,j}$ (resp. $\tilde{\tau}_{ij}$). The number of samples M_i, N_i and the sampling times are assumed to be random variables independent from the processes X, Y. As mentioned before, the first step towards solving the regression problem is to find an expansion of the predictor and the response on respective (infinite but countable) basis, respectively $(\phi_i)_{i\in\mathbb{N}}$ and $(\psi_i)_{i\in\mathbb{N}}$. A widely used procedure in the field of functional data analysis is to find a smooth approximation to covariance function of X (resp. Y) then find estimators of the eigenvalues/eigenvectors of the covariance operators (such procedure is known as Functional Principal Component Analysis FPCA). As mentioned before, it is not guaranteed that this will result in an optimal representation for regression purpose, but it has proved quite efficient and robust in practice. However, finding the eigenvalue/eigenvectors from the empirical covariance function obtained from the measurements is quite a lengthy process. First of all, a smooth estimator of mean and covariance function has to be obtained. A first approach is to use a local linear smoother . Given a kernel K and a bandwidth parameter h, the local linear smoother for the (t, X) scatterplot is obtained by minimization over a, b of :

$$\sum_{i=1}^{n}\sum_{j=1}^{M_i} K\left(\frac{t-\tau_{ij}}{h}\right)(X_{ij} - a - b(t-\tau_{ij}))^2$$

with t being a fixed time. The optimal values a, b obtained for a fixed t give a local linear model, so that the estimated mean at time t is $\hat{\mu}_X(t) = a$. For covariance estimation, the procedure is roughly the same, but instead of considering the samples X_{ij} of X, the empirical covariance function :

$$C'_X(\tau_{i,j_1}, \tau_{i,j_2}) = (X_{i,j_1} - \hat{\mu_X}(\tau_{i,j_1}))(X_{i,j_2} - \hat{\mu_X}(\tau_{i,j_2}))$$

is used. The corresponding local linear model is bivariate and corresponds to the minimum over a, b, c of :

$$\sum_{i=1}^{n}\sum_{j=1}^{M_i}\sum_{k=1,k\neq j}^{M_i} K\left(\frac{t-\tau_{ij}}{h}, \frac{s-\tau_{ik}}{h}\right) (C'_X(\tau_{i,j}, \tau_{i,k}) - a - b(t-\tau_{i,j}) - c(s-\tau_{i,k}))^2 \quad (2)$$

The estimated covariance $\hat{C}_X(t,s)$ at time t, s is thus the term a in the previous expression. Several packages exist for such estimation. For our purpose, an ad-hoc algorithm has been developed : the general principle will be described below.

It must be noted that local linear smoother is not the only usable procedure for fitting a smooth curve or surface to scatterplot data. Smoothing splines seem to be a good choice too. The difference in terms of performance between those two approaches has not been investigated yet, but at first sight there is no evidence why the first one will be better : it has been chosen in our application only because of its availability and some asymptotic approximation results.

Next step is to estimate eigenvalues and eigenvectors of the covariance function by solving the functional equation :

$$\int_{[a,b]} \hat{C}_X(t,s)\psi(s)ds = \lambda\psi(t)$$

Several numerical procedures can be found in the numerical analysis literature for solving such problems. We have applied a Nyström method on a regular grid for the trajectory application (see details in the next section). The result is two finite sets of pairs eigenvalues/eigenvectors $(\lambda_i, \psi_i)_{i=1\dots P}$ and $(\mu_i, \theta_i)_{i=1\dots Q}$ for X and Y respectively. The number of representing functions (the P and Q integers) has to be chosen

ISBN: 978-0-615-20720-9

on either a leave-one-out or an AIC : the later gives good results in the trajectory case.

Given the eigenvalues/eigenfunctions pairs for X and Y respectively, it is possible to compute an estimate of the covariance of X and Y components respectively :

$$\hat{\sigma}_{XY}(i,j) = \int_{[c,d]} \int_{[a,b]} \theta_j(s)\psi_i(t)\hat{C}_X(t,s)dtds$$

The resulting optimal kernel solving (approximately) the functional regression problem is then obtained as :

$$K(t,s) = \sum_{i=1}^{P}\sum_{j=1}^{Q} \frac{\hat{\sigma}_{XY}(i,j)}{\lambda_i}\psi_i(t)\theta_j(s)$$

IV. Application to trajectory prediction

Applying functional regression to trajectories implies :

- Extending all previous estimators to vector valued ones (thus replacing the covariance function by a 3×3 matrix valued function).
- Find the right predictor and response.

A. Principal components in the vector case

Recall that the chosen representation basis is obtained by functional principal component analysis. For trajectory prediction purpose, all random processes have values in $\mathbb{R}^3$, so that canonical procedures have to be extended. Estimation of mean and covariance functions can be used readily since the proposed local linear estimator extends componentwise to the 3-dimensional case. It should be noted however that computing the mean function involve 3 times more computation than for the scalar case and computational task is scaled by a 6 factor for the covariance (due to the symmetry of the covariance matrix). An important step in the design of a linear smoother is the choice of weighting kernel and bandwidth. The problem has been addressed in the field of non parametric statistics and it is known that the kernel has less influence than the bandwidth. The Epanechnikov kernel :

$$K_e(t) = \frac{3}{4}(1-t^2)1_{[-1,1]}(t)$$

has some interesting optimality properties and is easy to compute. Another choice is the Gaussian kernel :

$$K_g(t) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{t^2}{2})$$

For very fast computation, it is still possible to use a uniform kernel :

$$K_u(t) = \frac{1}{2}1_{[-1,1]}(t)$$

Since the data set is usually large (around 1000 trajectories sampled at 10s), a compactly supported kernel in the local linear smoother allows a reduced computational load and an complexity mostly independent of the number of samples in a trajectory. The Gaussian kernel is not compactly supported, but decreases very fast at infinity so that practically it can be set to 0 outside an compact interval. The optimal bandwidth can be found in the limit of large samples. If the kernel is K, and the target function f is supported in the interval $[a,b]$ then the asymptotically optimal bandwidth is :

$$h^5 = \frac{\sigma^2(b-a)^2\int K^2(x)dx}{N\left(\int x^2K(x)dx\right)^2\int_a^b f''^2(x)dx}$$

with σ^2 the variance of the noise, N the number of samples. Since f is unknown, one has to estimate both σ^2 and f''. Since f has to be estimated, it is clear that only some rough guess on plausible value can be done. A first approach is simply to increase the order of the model as :

$$\sum_{i=1}^{n}\sum_{j=1}^{M_i} K\left(\frac{t-\tau_{ij}}{h}\right)(X_{ij} - a - b(t-\tau_{ij}) - c(t-\tau_{ij})^2)^2$$

the c coefficient will then give an estimate of $f''(\tau_i j)/2$. This method is appealing for trajectory modelling since curvature can be obtained directly from c. However, experiments on real data show that little is gained since the second order model itself has a bandwidth that must be set heuristically. A second approach is simply to compute using a finite difference operator. It works surprisingly well for such a naive approach, probably because noise is not dominant in ATM data. Assuming that the bandwidth is known, solving for the parameters of the local linear model can be done easily. For a given vector or matrix valued function X sampled at times τ_{ij} with values A_{ij}, the mean value at time t can be obtained using the following formula :

$$\hat{X}(t) = -dm_X + ef_X$$

with :

$$d = \sum_{i=1}^{n}\sum_{j=1}^{N_i} K_e\left(\frac{t-\tau_{ij}}{h}\right)(t-\tau_{ij})^2$$

$$e = \sum_{i=1}^{n}\sum_{j=1}^{N_i} K_e\left(\frac{t-\tau_{ij}}{h}\right)(t-\tau_{ij})$$

and :

$$m_X = \sum_{i=1}^{n}\sum_{j=1}^{N_i} K_e\left(\frac{t-\tau_{ij}}{h}\right)X_{ij}$$

$$f_X = \sum_{i=1}^{n}\sum_{j=1}^{N_i} K_e\left(\frac{t-\tau_{ij}}{h}\right)(t-\tau_{ij})X_{ij}$$

Since K_e is compactly supported (or approximately such), the inner sum in the previous expression will only involve a number of terms bounded by a constant so that the computational task for computing $\hat{X}(t)$ at a given t scales linearly with the number of trajectories in data set but independent from the size of the trajectories themselves. Once the covariance functions have been obtained for both the predictor and the response, the eigenvalues/eigenfunctions are computed by solving an homogeneous Fredholm equation of the second kind :

$$\int_{[a,b]} C(s,t)\phi(s)ds = \lambda\phi(t)$$

ISBN: 978-0-615-20720-9

with C the covariance function of interest. Several numerical methods can be found in the literature for such a problem. In our setting, the Nyström method based on Gauss-Legendre quadrature formula has proved very well fitted. The method proceeds by approximating the integral operator as :

$$\sum_{i=1}^{n} w_i C(s_i, t)\phi(s_i)$$

with w_i and s_i respectively the weights and the abscissa of the Gauss-Legendre quadrature approximation. For n points, the abscissa are the roots of the n-th Legendre polynomial P_n while the weights are computed as :

$$w_i = \frac{2}{(1-s_i^2)P_i'(s_i)^2}$$

with the Legendre polynomials P_i evaluated by the recursion :

$$(i+1)P_{i+1} = (2i+1)xP_i - iP_{i-1}$$

with conventionally$P_{-1} = 0, P_0 = 1$. The Fredholm equation of the second kind is approximated by the following eigenvalues problem :

$$\sum_{i=1}^{n} w_i C(s_i, s_j)\phi(s_i) = \lambda\phi_(s_j),\ j = 1 \dots n$$

The result is a set of eigenvectors $\phi_(s_j, i), j = 1 \dots n, i = 1 \dots n$ and eigenvalues λ_i. The approximation to the eigenfunction of the original covariance operator associated to eigenvalue λ_i is then :

$$\phi_i(t) = \frac{1}{\lambda_i}\sum_{i=1}^{n} w_i C(s_i, t)\phi(s_i)$$

(the case $\lambda_i = 0$ will not occur in our application). Since our original processes X, Y take their values in $\mathbb{R}^3$, all previous equations are to be taken as vector ones. The main consequence is that the approximate eigenvalue has to be solved with a full $3n \times 3n$ system. Fredholm equation solving is quite a critical step in the overall algorithm, so attention has been paid to its accuracy. In fact, Gaussian quadrature is not the only procedure that can be used, but any quadrature formula based on samples will work. A comparison has been made between low accuracy rectangle method for approximating the integral and the complete Nyström algorithm with Gauss-Legendre quadrature. For that purpose, the test set has be obtained by generating trajectories of a simple random process (namely a square root function with an additive gaussian noise and a random scaling).

The error obtained with the two methods on the test set is, as a function of the number of eigenfunctions : Nyström algorithm appears to be more accurate than rectangle quadrature, at least for small number of eigenfunctions. However, method tends to yield ill conditioned matrices when a large number of eigenfunctions is required : this phenomenon explains why Nyström is outperformed by rectangle quadrature in such cases. The difference between the two procedures remains anyway quit low, indicating that a high order quadrature formula is not of great importance for our purpose.

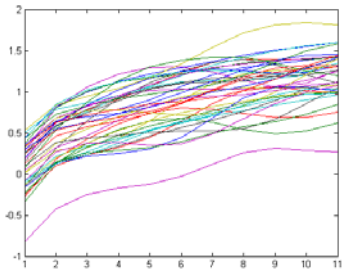

Fig. 1. Test functions for fredholm equation

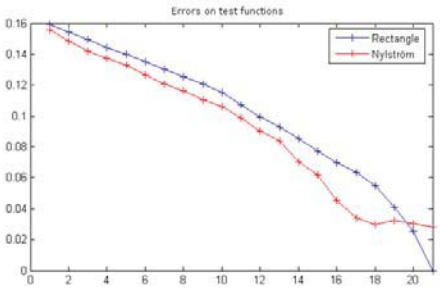

Fig. 2. Error comparison

B. Predictor

Finding the right predictor is a critical task in applying functional regression. For trajectory prediction purpose, it is natural to consider a part of the observed trajectory as the predictor, and a part of the future trajectory as target. The learning database has thus been chosen by selecting homogeneous segments of 20 radar plots from a day of traffic, then for each segment cut into two 10 plots pieces. The first piece will be used as predictor and second one as target. A total of 3200 trajectories has been considered, with a final database of 100 segments. In the traffic, a test database with the same number of segments and similar characteristics has been selected too. Since the random process associated to trajectories has no reason to be stationary, the Karhunen-Loeve basis is a priori different for the predictor and the target. These two basis will be denoted respectively by $(\phi_i)_{i\in\mathbb{N}}$ and $(\psi_i)_{i\in\mathbb{N}}$. Let $(X_k, Y_k)_k$ be the k-th sample from the learning base (that is X_k is the first half on segment k while Y_k is the second half), the regression problem is to find an optimal K such that :

$$\sum_{k=1}^{N} \int \|Y_k(t) - \int K(t,s)X_k(s)ds\|^2 dt$$

is minimal. The kernel K can be expressed using basis $(\phi_i)_{i\in\mathbb{N}}$, $(\psi_i)_{i\in\mathbb{N}}$ as :

$$K(t,s) = \sum_i \sum_j K_{ij}\psi_i(t)\phi_j(s)$$

Using orthogonality of the Karhunen-Loeve eigenfunctions, the original problem is reduced to find optimal an optimal

ISBN: 978-0-615-20720-9

sequence (K_{ij}) minimizing :

$$\sum_{k=1}^{N} \int \left\| \sum_i \psi_i(t) \left(b_{ik} - \sum_j K_{ij} a_{jk} \right) \right\|^2 dt$$

assuming that the expansions of X_k, Y_k are :

$$X_k(s) = \sum_j a_{jk} \phi_j(s), Y_k(t) = \sum_i b_{ik} \psi_k(t)$$

using the orthonormality of $(\psi_i)_{i \in \mathbb{N}}$, the minimization problem reduces further to :

$$\min_{(K_{ij})} \sum_{k=1}^{N} \sum_i \left(b_{ik} - \sum_j K_{ij} a_{jk} \right)^2$$

In practice, expansions are truncated to a fixed rank. Let P be the corresponding integer. The approximate finite dimensional mimization problem is :

$$\min_{(K_{ij})} \sum_{k=1}^{N} \sum_i^P \left(b_{ik} - \sum_j^P K_{ij} a_{jk} \right)^2$$

which is nothing but a linear least mean square problem that can be solved with the help of normal equations or QR factorization. The result of prediction for first the test functions (same as in previous sections) is summarized below :

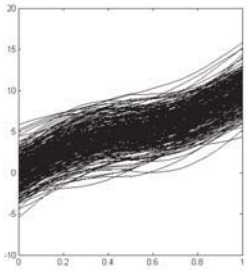

Fig. 3. Test functions for prediction

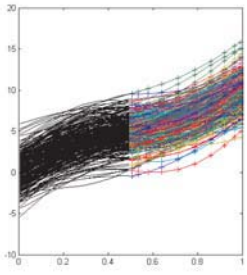

Fig. 4. Predicted second half of trajectories

In the case of real traffic, the obtained results are : The relative error for in the horizontal plane is kept low and increases as expected with prediction time. In the case of

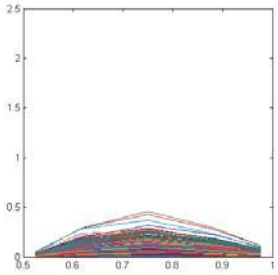

Fig. 5. Relative prediction error

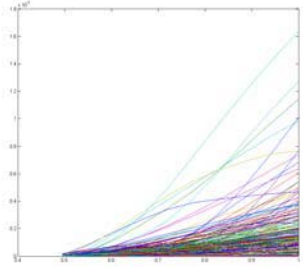

Fig. 6. Relative prediction error (x-y plane)

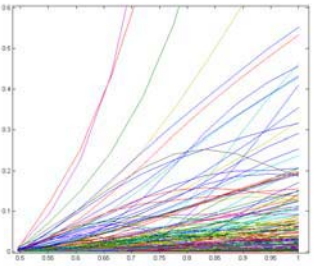

Fig. 7. Relative prediction error (z component)

z component, relative error is high, but measurements are discrete (flight levels), thus there is an intrinsic noise coming from quantization. For this component, prediction is indeed built-in with smoothing : a comparison with a smoothed trajectory yields much lower error.

C. Clustered regression

From now, only the case of trajectories originating from the same underlying stochastic process has been considered. While several operational situations fall in this category (especially when the predictor consists of small parts of trajectories), it is known that this assumption is false when applied to large areas of the airspace. To deal with this problem, it may be necessary to introduce a cluster regression. The data set of trajectories is partitioned into homogeneous classes (clusters) based on a relative distance criterion (most of the time, a L^2 norm or a Sobolev norm is used). Once the clustering has been done,

ISBN: 978-0-615-20720-9

functional regressors are computed cluster by cluster, assuming that all trajectories in the same cluster are close enough to be modeled as samples of the same stochastic process. When applying functional regression to a new trajectory, the closest cluster is chosen, then the corresponding regressor is used. This way of applying functional regression has proved to be very efficient for inhomogeneous 1D curves ; however, application to 3D trajectories has not been done yet.

V. Conclusion and future work

The functional regression is a promising approach on simulated situations. Besides of producing an estimation of the future positions of aircraft in the short to mid term prediction range, it is possible to derive confidence regions for the actual position, thus yielding a better control on the quality of the produced solution. The computational task involved is heavy, but has to be done only once : as soon as the Karhunen-Loeve has been produced, it can be used at low cost. The future work will be first experiments on selected learning data sets so that functional regression can be compared with other methods. A second aspect will be to introduce other kind of decompositions (namely wavelets and curvlets basis) that are known to perform much like Karhunen-Loeve but at a much lower computational cost and to investigate the cluster functional regression. Furthermore, a functional regression software adapted to trajectory prediction is currently under development.

References

[1] D.J. Brudnicki and A.L. McFarland, *User request evaluation tool (uret) conflict probe performance and benefits assessment*, FAA/Eurocontrol ATM Seminar Saclay France, 1997.

[2] H. Cardot, F. Ferraty, and P. Sarda, *Functional linear model*, Statist. Probab. Letters **45** (2003), 11–22.

[3] R. Coppenbarger, R. Lanier, D. Sweet, and S. Dorsky, *Design and development of the enroute descent advisor (eda) for conflict-free arrival metering*, AIAA-2004-4875 AIAA GNC Conference Providence RI, 2004.

[4] H. Erzberger, R.A. Paielli, D.R. Isaacson, and M.M. Eshowl, *Conflict detection in the presence of prediction error*, FAA/Eurocontrol ATM Seminar Saclay France, 1997.

[5] J. Evans and al, *Reducing severe weather delays in congested airspace with weather support for tactical air traffic management*, FAA/Eurocontrol ATM Seminar BUdapest Hungary, 2003.

[6] F. Ferraty and P. Vieu, *Functional nonparametric statistics : A double infinite dimensional framework*, North-Holland, 2003.

[7] I.I. Gikhman and A.V. Skorokhod, *Introduction to the theory of random processes*, Dover, 1996.

[8] D.B. Kirk and al, *Problem analysis resolution and ranking (parr) development and assessment*, FAA/Eurocontrol ATM Seminar Santa Fe NM, 2004.

[9] A. Masalonis and al, *Using probabilistic demand prediction for traffic flow management decision support*, AIAA-2004-5231 AIAA GNC Conference Providence RI, 2004.

[10] B.D. McNally, R.E. Bach, and W. Chan, *Field test evaluation of the ctas conflict prediction and trial planning capability*, AIAA-1998-4480 AIAA GNC Conference Boston MA, 1998.

[11] C. Meckiff, R. Chone, and J.P. Nicolaon, *The tactical load smoother for multi-sector planning*, FAA/Eurocontrol ATM Seminar Orlando FL, 1998.

[12] S. Mondoloni, S.M. Paglone, and S. Green, *rajectory modeling accuracy for air traffic management decision support tools*, ICAS Congress Toronto ON, 2002.

[13] S. Mondoloni, Swierstra S., and M. Paglione, *Assessing trajectory prediction performane – metrics definition*, 24th Digital and Avionics Systems Conference, 2005.

[14] H. G. Müller and U. Stadtmüller, *Generalized functional linear models*, Annals of statistics **33** (2005), no. 2, 774–805.

[15] J. Ramsay and B. W. Silverman, *Functional data analysis*, Springer-Verlag, 1997.

[16] H.F. Ryan, M. Paglione, and S. Green, *Review of trajectory accuracy methodology and comparison of error measure metrics*, AIAA-2004-4787 AIAA GNC Conference Providence RI, 2004.

[17] R. Slattery and Y. Zhao, *Trajectory synthesis for air traffic automation*, AIAA Journal Guidance Control and Dynamics, vol. 20, 1997, pp. 232–238.

[18] S.Mondoloni and I. Bayraktutar, *mpact of factors, conditions and metrics on trajectory prediction accuracy*.

[19] V. Sud and al, *Air traffic flow management collaborative routing coordination tools*, AIAA-2001-4112 AIAA GNC Conference Montreal PQ, 2001.

[20] H.N Swensen and al, *Design and operational evaluation of the traffic management advisor at the forth worth air route traffic control center*, FAA/Eurocontrol ATM Seminar Saclay France, 1997.

[21] S. Swierstra and S. Green, *Common trajectory prediction capability for decision support tools*, FAA/Eurocontrol ATM Seminar BUdapest Hungary, 2003.

[22] A. Vink, *Eatchip medium term conflict detection: Part 1 eatchip context*, FAA/Eurocontrol ATM Seminar Saclay France, 1997.

[23] F. Yao, H. G. Müller, and J. L. Wang, *Functional linear regression for longitudinal data*, Annals of statistics **33** (2005), no. 6, 2873–2903.

[24] ———, *Functional linear regression for sparse longitudinal data*, Journal of Amer. Statist.Assoc. **100** (2005), 577–590.

ISBN: 978-0-615-20720-9