

Model-Based Interpolation, Prediction, and Approximation

Antonio Possolo

Statistical Engineering Division
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, Maryland, USA
antonio.possolo@nist.gov
<http://www.nist.gov/itl/sed/possolo.cfm>

Abstract. Model-based interpolation, prediction, and approximation are contingent on the choice of model: since multiple alternative models typically can reasonably be entertained for each of these tasks, and the results are correspondingly varied, this often is a considerable source of uncertainty. Several statistical methods are illustrated that can be used to assess the contribution that this uncertainty component makes to the uncertainty budget: when interpolating concentrations of greenhouse gases over Indianapolis, predicting the viral load in a patient infected with influenza A, and approximating the solution of the kinetic equations that model the progression of the infection.

Keywords: interpolation, prediction, approximation, uncertainty, influenza, greenhouse gases, projection pursuit

1 Introduction

Three examples are described that illustrate the application of statistical methods to assess the contribution that model uncertainty makes to the overall uncertainty when doing interpolations, producing predictions, or building approximations.

The first example, described in §2, relates to the interpolation of concentrations of CO₂ measured in the course of a flight over Indianapolis: two interpolants are considered, local regression and kriging, among many others that could reasonably be entertained, and this suffices to make the point that model uncertainty can make a sizeable contribution to the overall uncertainty budget.

The second example, reviewed in §3, concerns a measurand (the quantity intended to be measured [15, 2.3], here the time that elapses after infection with the influenza A virus, until the viral load peaks in a particular patient) whose measurement involves the solution of a set of simultaneous differential equations. Since this depends on several parameters, each set of values assigned to them in fact defines a particular model. The example characterizes the dispersion of values that correspond to these multiple, alternative choices.

The third example, presented in §4, is about the construction of a data-driven approximant to an unknown function of several variables. The quantity of interest is the same that is considered in the example of §3, the time that elapses after infection with the influenza A virus until the viral load peaks, but here regarded as a function of four parameters. The approximants considered all belong to the class of projection pursuit regressions. In this case, the alternative models under consideration correspond to the different numbers of “ridge functions” used to define the structure of the approximant.

These examples were all worked out using the facilities for statistical computing available in the R programming environment for statistical computing, data analysis, and graphics [22], which is free and open source. Writing for the *New York Times* of January 6, 2009, under the title “Data Analysts Captivated by R’s Power”, Ashlee Vance explains that R is a popular programming language used by a growing number of data analysts, and suggests that “it is becoming their lingua franca”.

Some commercial instruments are identified in this paper accurately to represent the sources of some of the data that is used. Neither does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment identified is necessarily the best available for the purpose, or that it has been properly calibrated or used.

2 Interpolation

2.1 INFLUX Experiment

The Indianapolis Flux Experiment (INFLUX, <http://influx.psu.edu/>, [24]) serves to develop and assess methods for quantifying greenhouse gas emissions on an urban scale: it is a joint undertaking of Arizona State University, the Cooperative Institute for Research in Environmental Sciences (National Oceanic and Atmospheric Administration and the University of Colorado at Boulder), the Earth Science Research Laboratory (National Oceanic and Atmospheric Administration), the National Institute of Standards and Technology, Penn State University, and Purdue University.

Figure 1 shows the flight path of a small aircraft, instrumented with a Picarro greenhouse gas analyzer (http://www.picarro.com/gas_analyzers), which flew from Purdue University to Indianapolis on June 1, 2011, and then took measurements of CO₂ concentrations at about 17 000 points lying approximately on a vertical curtain about 80 km wide and 1.2 km tall, as illustrated in Figure 2.

We consider the problem of interpolating the observations to create a map of those concentrations covering the whole vertical flight curtain, which subsequently will be used to compute vertical CO₂ flux. Since such interpolation can reasonably be done in any of several alternative ways, the corresponding dispersion of results ought to be reflected in the map’s uncertainty budget.

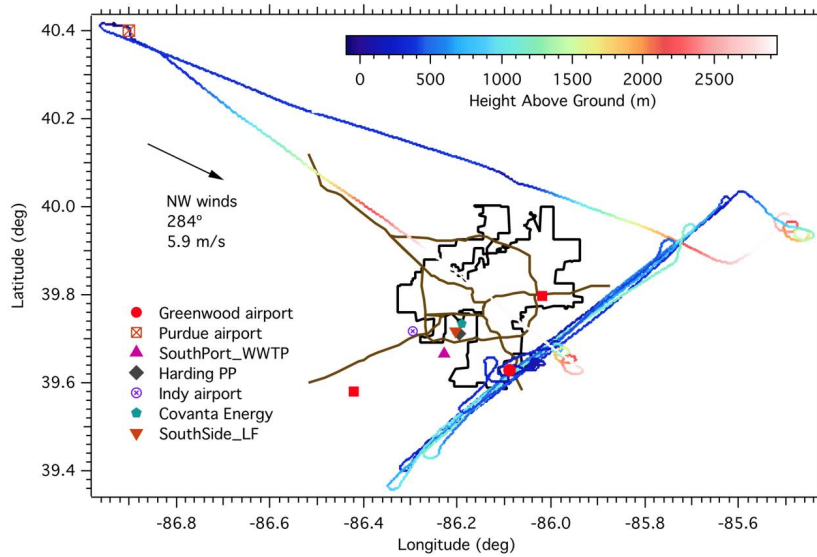


Fig. 1. Flight path of a small aircraft equipped with a Picarro greenhouse gas analyzer that made the measurements depicted in Figure 2; map drawn by Paul B. Shepson and M. Obiminda Cambaliza (Department of Chemistry, Purdue University) [24]

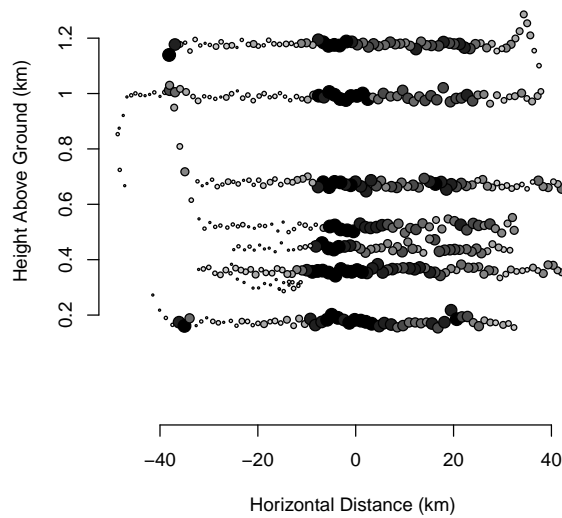


Fig. 2. Sample of locations on the vertical curtain where concentrations of CO₂ were measured on June 1, 2011, using a Picarro greenhouse gas analyzer, with circle size and gray level (the darker the higher the value) indicating the measured values.

2.2 Local Regression & Kriging

The data are a set of values y_1, \dots, y_m of a real-valued, albeit unknown function θ , which were measured at points x_1, \dots, x_m in a space \mathcal{X} on which there is defined a distance metric. The objective is to produce an estimate y of $\theta(x)$, for any x “in the middle” of the $\{x_i\}$, typically but not necessarily different from these $\{x_i\}$, and to characterize the uncertainty $u(y)$ associated with the estimate, conceived as an indication of how close to $\theta(x)$ one believes y to be.

According to the *Guide to the expression of uncertainty in measurement* (GUM) [13, 3.3.1] measurement uncertainty “reflects the lack of exact knowledge of the value of the measurand” (the value like $\theta(x)$, this parameter will be the standard deviation of a probability distribution that describes the dispersion of values that may reasonably be assigned to y).

The goal of estimating $\theta(x)$ is approached by modeling the observations $\{y_i\}$ probabilistically, so that the interpolation problem becomes a statistical estimation problem. In this conformity we represent the data as $y_i = \theta(x_i) + \epsilon_i$ for $i = 1, \dots, m$, where the $\{\epsilon_i\}$ are regarded as realized values of non-observable random variables (*measurement errors*). The objective, then, is to interpolate the signal $\theta(x_i)$, not the signal plus the noise, $\theta(x_i) + \epsilon_i$. (The assumption that the signal and the noise combine additively can often be satisfied by suitably re-expressing, or transforming, the data before the statistical analysis.)

This goal can be achieved in any one of many different ways, several of which are comparably reasonable under mild assumptions about θ and about the probability distribution of the measurement errors. We choose to illustrate and compare the results of two well-known, widely used procedures: local regression [2,3,17] and kriging [6,16,27]. In both, a model needs to be chosen for θ , which next needs to be calibrated in light of the data (that is, its adjustable parameters need to be estimated), the resulting $\hat{\theta}$ finally being used to compute the interpolated value as $y = \hat{\theta}(x)$.

To apply local regression we assume that θ is continuous and locally quadratic. To apply kriging we assume that θ is a realized value of a stationary Gaussian random function Θ whose covariance function is a member of a particular parametric family. Both methods can be employed under more general assumptions, but neither of these assumptions already is particularly restrictive. In fact, and for local regression, assuming that θ is locally constant, linear, or quadratic makes little difference in most practical applications; and for kriging, the collection of realizations of a Gaussian random function, even a stationary one, is sufficiently varied to be able to mimic the majority of continuous functions likely to be encountered in practice.

The methods available to build these interpolants from empirical data are amply documented in the literature: for example in [4], which discusses interpolation of spatial data at great length. We used function `locfit` from the R package with the same name [18] to fit the local regression model, and function `automap` from the R package `intamap` [21] to fit the kriging model. Before discussing the results for the CO₂ data, we illustrate them in the context of a very simple example.

Figure 3 shows differences between a local regression model and a kriging model fitted to the same data, for a set of simulated data where \mathcal{X} is the real line and θ is a real-valued function of a real argument. The kriging model is defined by a constant mean $\mu = \mathbb{E}(\Theta(x))$ for all x , and by a covariance function γ such that $\gamma(h) = \mathbb{E}[(\Theta(x+h) - \mu)(\Theta(x) - \mu)]$.

The assumption is made that γ depends only on h and not on x (where the increment h is a scalar when \mathcal{X} is the real line, and it is a vector when \mathcal{X} is the Euclidean plane), and that it belongs to the class of covariance functions introduced by Matérn [19], involving three adjustable parameters. This model was fitted to the data depicted in Figure 3 using facilities provided by the R package `geoR` [23].

Both the local regression and the kriging interpolants do some smoothing because the goal is to interpolate the signal — that is, values of $\theta(x)$ — not the signal plus the noise, $\theta(x) + \epsilon$. In practice, of course, since neither θ nor ϵ are known, a statistical procedure needs to be employed that “guesses” the most appropriate extent of smoothing: one such procedure is *cross-validation*, which will be discussed below.

The results obtained by the two methods are similar but different. Since either model could have reasonably been selected for the task at hand, these differences reflect an uncertainty component that is attributable to model selection. Another uncertainty component relates to model calibration (that is, estimation of model parameters), and derives from the fact that this calibration is based on only a finite amount of data.

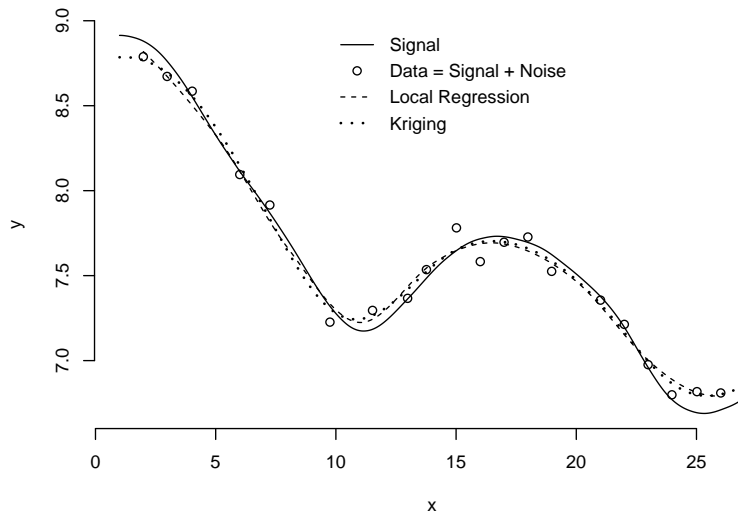


Fig. 3. Illustrative example of differences between a local regression model and a kriging model fitted to the same data. The solid line, marked “Signal” in the legend, depicts the “true” function θ : this is known in this case because the data are simulated.

Kriging is often heralded as providing assessments of uncertainty of its interpolations automatically. However, in many instances of application, kriging's built-in assessments *underestimate* uncertainty because one pretends that the estimate $\hat{\gamma}$ of the covariance function is identical to the covariance function γ itself. Bayesian kriging provides means to account for this often neglected component of uncertainty [6]. Here, we will rely on cross-validation to evaluate the corresponding uncertainty component.

2.3 Cross-Validation and Model Uncertainty

Cross-validation [11,20] is an established procedure to assess the performance of statistical procedures realistically. In the context of interpolation, the idea is to partition the data into two subsets, use one (*training subset*) to develop and calibrate the interpolant, and the other (*testing subset*) to gauge its performance, by comparing the interpolation results with the observations at those locations present in the testing subset. The process may be repeated multiple times, over many different partitions, and some figure of merit (for example, root mean square of the differences between interpolated and actual values) averaged.

The partitioning may be done at random, or it may include consideration for the particulars of the situation. For the CO₂ data whose measurement locations are depicted in Figure 2, a random partition may lead to overoptimistic assessments. Given the fairly high rate at which the measurements were made, the measurement locations are very closely spaced along the horizontal legs of the curtain portion of the flight: for this reason, if the testing subset had been obtained by selecting, say, about one third of the measurement locations chosen uniformly at random, then there would have been data in the training subset very close to each of the locations in the testing subset, and the interpolation accuracy would misleadingly appear to be very high.

The interpolation challenge in this case, however, is mostly in the vertical direction, at points whose height lies between the heights of any pair of adjacent, horizontal legs of the flight. A suitable testing subset should then be one of these horizontal legs (Figure 4), with the rest of the observations being used to build an interpolant to estimate concentrations of CO₂ at all the points in the leg that will have been set aside. And then to repeat this procedure for each leg in turn.

2.4 INFLUX Experiment — Uncertainty Budget

Figure 5 shows the results of local regression and kriging interpolations for the INFLUX data. The differences are generally small, being most pronounced along the edges of the CO₂ plume. The cross-validation assessment of their respective performance indicates that they are quite comparable, and there is hardly any reason to prefer one over the other.

The uncertainty budget in Table 1 lists the recognized uncertainty components and the contributions they make to the expanded combined measurement uncertainty in the last line of the table, computed as $U_{95\%} = 2\sqrt{0.36^2 + \dots + 0.7^2}$

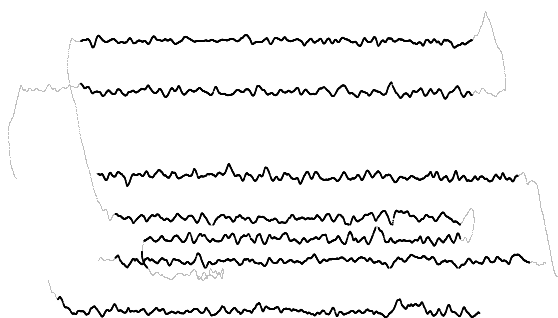


Fig. 4. Locations on the flight curtain where measurements of CO_2 were made (gray and black points), also depicted in Figure 2. The seven testing subsets used when cross-validating the performance of the two interpolation procedures are the horizontal legs of the flight path highlighted in black.

$= 2.5 \text{ mg/m}^3$ (ppmv), meaning that, with approximate 95 % probability, the true CO_2 concentrations are within $\pm U_{95\%}$ of the values shown on the map.

Table 1. Uncertainty budget for the map of CO_2 concentrations over the curtain flight plane. To within the single significant digit listed, the interpolation uncertainty is the same for the local regression and kriging models. The evaluations marked MANUF were drawn from the specification sheet for the Picarro CRDS Analyzer model G2301-m (http://www.picarro.com/gas_analyzers/flight_co2_ch4_h2o); CV indicates evaluation via cross-validation; and LAB+CERT indicates that the evaluation expresses laboratory calibration uncertainty and the uncertainty in the calibration standards.

SOURCE	EVALUATION	STD. UNCERT. mg/m^3 (ppmv)
Model selection	CV	0.36
Interpolation	CV	0.9
Instrument calibration	LAB+CERT	0.034
Instrument repeatability	MANUF	0.2
Instrument drift	MANUF	0.2
Atmospheric temperature	MANUF	0.0075
Atmospheric pressure	MANUF	0.7
Expanded Uncertainty		$U_{95\%} = 2.5 \text{ mg/m}^3$ (ppmv)

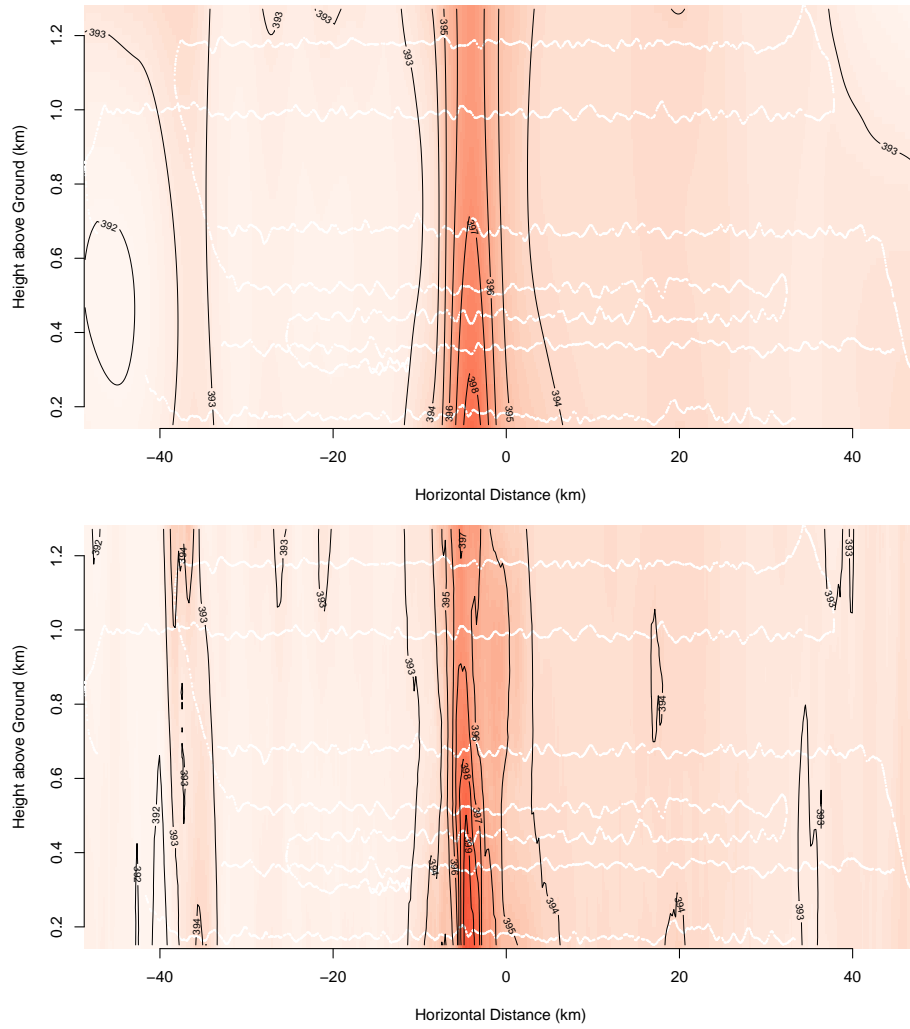


Fig. 5. Results of interpolating the CO₂ measured values (expressed in mg/m³) over a regular grid, then colored in shades of gray and contoured, using local regression (top panel) and kriging (bottom panel).

3 Prediction

3.1 Viral load in influenza A infection

[1] describe several mathematical models for the progression of influenza A infection within an individual patient. These models capture the principal features of the underlying viral kinetics: initially the viral load grows exponentially fast, peaks 2 to 3 days post-infection, and finally decreases, also at an exponential rate, to undetectable levels within 6 to 8 days post-infection.

Here we will consider the simplest model that [1] describe. The infection is assumed to be limited by the availability of susceptible epithelial cells rather than by the patient's immune response. The number T of uninfected target cells, the number I of productively infected cells, and the viral load V satisfy the following system of differential equations, where β is the infection rate, $1/\delta$ is the expected lifespan of an infected cell, ρ is the increment to viral load contributed by each infected cell, and γ is the viral clearance rate:

$$\frac{dT}{dt} = -\beta TV, \quad \frac{dI}{dt} = \beta TV - \delta I, \quad \frac{dV}{dt} = \rho I - \gamma V. \quad (1)$$

Figure 6 shows the values of the viral load in patient 4 (from Table 1 of [1]) as measured on seven consecutive post-infection days, and the fit this model achieves assuming that $\log_{10} V$ is Gaussian, and that V satisfies equations (1). The differential equations were solved by application of the Livermore Solver for Ordinary Differential Equations (LSODA) [12], as implemented in R package `deSolve` [26]. The model was fitted by minimizing the sum of squares of the differences between the logarithm of the mean viral load (which is a function of β , δ , ρ , and γ), and the logarithm of the corresponding measured load: the unit of measurement is TCID₅₀, that is 50% Tissue Culture Infective Dose per milliliter of nasal wash.

Consider predicting the post-infection time $\operatorname{argmax}_t V_t$ when the viral load peaks, and estimating the infection's *Basic Reproductive Number* $R_0 = \rho\beta T_0/(\gamma\delta)$, which is the average number of second-generation infections produced by a single infected cell placed among susceptible cells: if $R_0 > 1$ the infection progresses to full course, and if $R_0 < 1$ the infection dies out prematurely.

Predicting the post-infection time $\operatorname{argmax}_t V_t = \psi(\beta, \delta, \rho, \gamma)$ involves solving the foregoing system of differential equations and assessing the uncertainty of the result. Estimating R_0 involves estimating ρ , β , T_0 , γ and δ , and propagating their associated uncertainties to obtain $u(R_0)$, the uncertainty associated with the estimate of R_0 .

Both tasks can be accomplished using the parametric statistical bootstrap [7] (*cf.* [14]), by drawing samples from suitable, joint probability distributions for the participating quantities. The key ingredient for this is a numerical approximation to the Hessian $H(\beta, \delta, \rho, \gamma)$ of the negative log-likelihood used to fit the kinetic model to the data for Patient 4, which was obtained using function `hessian` of the R package `numDeriv` [10].

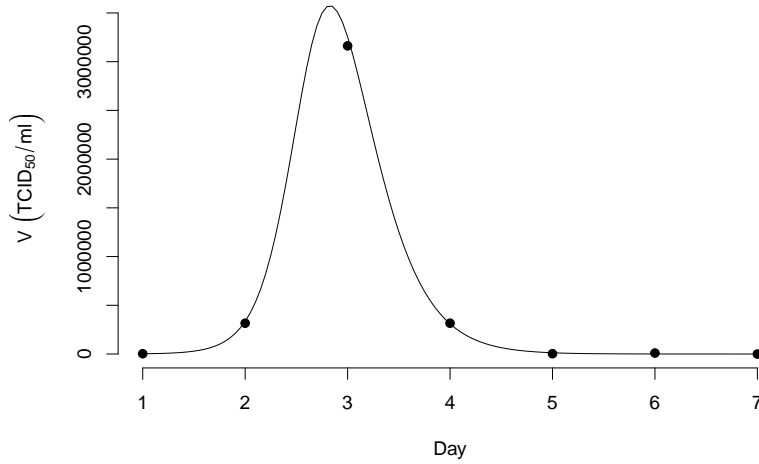


Fig. 6. Observed viral load of patient 4 [1, Table 1] at each of seven days (red dots), and fitted model (blue curve).

The parametric statistical bootstrap then amounts to repeating the following steps for $k = 1, 2, \dots, K$, where K is a suitably large integer: (i) draw a sample $(\beta_k, \delta_k, \rho_k, \gamma_k)$ from a multivariate Gaussian distribution with mean $(\hat{\beta}, \hat{\delta}, \hat{\rho}, \hat{\gamma})$ and covariance matrix $H^{-1}(\hat{\beta}, \hat{\delta}, \hat{\rho}, \hat{\gamma})$, where $\hat{\beta}$ denotes the estimate of β , and similarly for the other parameters; (ii) draw one sample from a uniform distribution for each initial condition $T_0 \pm 0.1T_0$, $I_0 \pm 0.1I_0$, and $V_0 \pm 0.1V_0$ (the 10% relative uncertainty assumed here is merely for purposes of illustration of the general idea that the uncertainty in the initial conditions also ought to be propagated); (iii) solve the kinetic model with perturbed parameters and compute $\psi(\beta_k, \delta_k, \rho_k, \gamma_k)$.

These steps produce K replicates of $\operatorname{argmax}_t V_t$ whose spread is indicative of the corresponding measurement uncertainty. Figure 7 summarizes the results for as an estimate of the probability density that encapsulates the state of knowledge about $\operatorname{argmax}_t V_t$.

These same K samples that were drawn from the joint probability distribution of β , δ , ρ , γ , and T_0 also yield K replicates of the Basic Reproductive Number R_0 , similarly depicted in Figure 8.

4 Approximation

Suppose that we observe values corrupted by non-observable errors, $(x_1, \psi(x_1) + \epsilon_1), \dots, (x_m, \psi(x_m) + \epsilon_m)$, of an unknown function $\psi : \mathcal{X} \mapsto \mathbb{R}$ that is “expensive” to evaluate, and that, based on such data, we wish to develop an approximant φ of ψ , and to assess its quality.

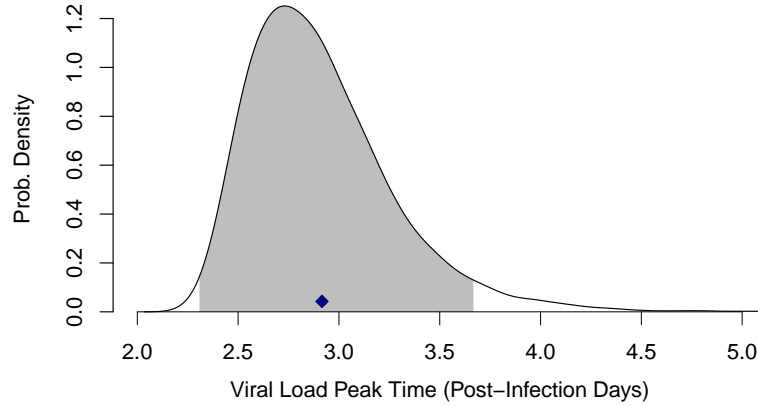


Fig. 7. Kernel probability density estimate [25] based on $K = 10\,000$ replicates of the values of $\operatorname{argmax}_t V_t$ obtained by application of the parametric statistical bootstrap. The average of these replicates, 2.9 post-infection days (marked with a diamond), is an estimate of $\operatorname{argmax}_t V_t$, and the associated uncertainty is the corresponding standard deviation, 0.4 post-infection days (PID). The shortest 95% probability interval (the footprint of the shaded area under the curve) ranges from 2.3 to 3.7 PID.

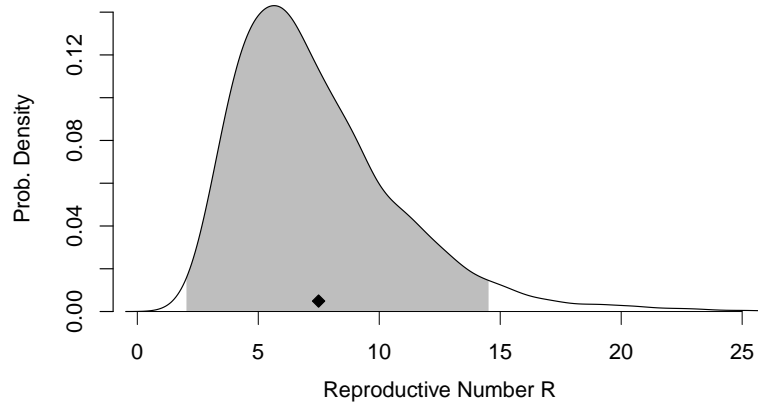


Fig. 8. Kernel probability density estimate [25] based on $K = 10\,000$ replicates of the values of R_0 obtained by application of the parametric statistical bootstrap. The average of these replicates is $\hat{R}_0 = 7.5$ with associated standard uncertainty $u(R_0) = 3.5$, and the shortest 95% probability interval is (2, 15), thus suggesting that the basic reproductive number is greater than 1 with very high probability: in fact, the same replicates allow estimating $\Pr(R_0 > 5) = 0.76$.

For the influenza example described in 3, the unknown function is $\psi(\beta, \delta, \rho, \gamma) = \operatorname{argmax}_t V_t$, and we wish to build φ such that $\varphi(\beta, \delta, \rho, \gamma) \approx \psi(\beta, \delta, \rho, \gamma)$, and such that the approximation is reasonably accurate, and less “expensive” to evaluate than ψ (whose evaluations involve solving a system of differential equations).

From among the several different methods available to build such approximant, we illustrate projection pursuit regression (PPR) [8,11], which builds on the idea of projection pursuit, an algorithm that “seeks to find one- and two-dimensional linear projections of multivariate data that are relatively highly revealing” [9]. PPR builds predictors out of these “highly revealing” projections, automatically sets aside variables with little predictive power, and bypasses the *curse of dimensionality* by focussing on functions of linear combinations of the original variables, at the price of intensive computation.

The approximant is of the form $\varphi(x) = \sum_{j=1}^J \varphi_j(\alpha_j \cdot x)$, where $\alpha_1, \dots, \alpha_J$ are vectors defining the projection directions, $\alpha_j \cdot x$ denotes the inner product of α_j and x , and $\varphi_1, \dots, \varphi_J$ are the so-called *ridge functions*. Considered for all values of J , and for appropriate choices of the ridge functions, these functions form a class of universal approximants capable of approximating any continuous function [5].

In practice one needs to choose a value for J , and both the projection directions and the ridge functions are derived from the data by means of a suitable smoothing procedure, which typically requires copious amounts of data. In this case we have 10 000 sets of observations wherefrom to build a PPR approximation to $\psi(\beta, \delta, \rho, \gamma) = \operatorname{argmax}_t V_t$.

The number of ridge functions, $J = 4$, was set equal to the number of predictors: in general, one may like to use cross-validation to select the optimal number. The R function `ppr` produces the ridge functions depicted in Figure 9, when instructed to select the best model with four terms among all such models with up to eighteen terms.

The cross-validated, relative approximation error is 3 % (this is the ratio between the standard deviation of the cross-validated approximation errors and the median of the values of the function ψ that is the target of the approximation). The component of uncertainty attributable to model choice amounts to 0.7 %, which is the mean value of the relative standard deviations of the predictions made by models with $J = 2, \dots, 6$, at each element of the testing subset: this is considerably smaller than the relative approximation error given above.

5 Conclusions

Models are necessary for measurement. And since a measurement result must comprise both an estimate of the value of the measurand and an assessment of the associated measurement uncertainty [15, 2.9], the models that are used to describe the interplay between all the quantities involved in measurement should facilitate such assessment. Since, furthermore, probability distributions are generally regarded as the best means to characterize measurement uncertainty [13, 3.3.4], measurement models ought to be statistical models.

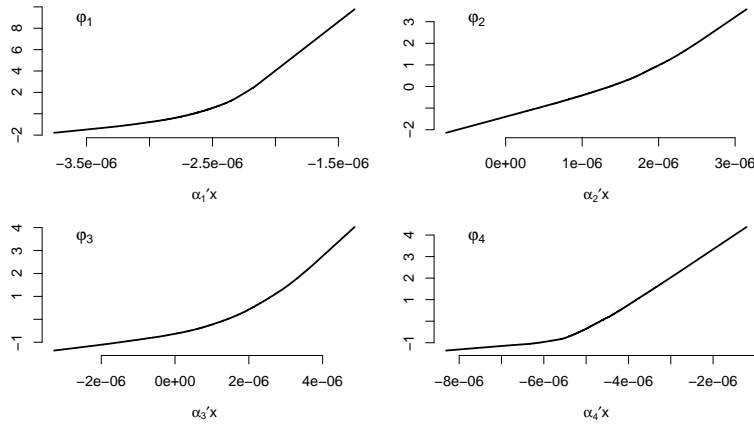


Fig. 9. Ridge functions of the projection pursuit approximant to $\psi(\beta, \delta, \rho, \gamma) = \operatorname{argmax}_t V_t$.

In many measurement situations, however, more than one model may reasonably be entertained. The three examples given above all should have sufficed to make this abundantly clear, each in a particular way that may be common to many other, similar applications. The dispersion of values that such multiplicity of choices entails ought to be evaluated and expressed in the assessment of measurement uncertainty, similarly to how the contributions from all the other recognized sources of uncertainty are so evaluated and expressed.

The assessment of the component of measurement uncertainty related to model choice generally involves inter-comparing the results corresponding to alternative, comparably tenable models, done in ways that insulate the results (of such inter-comparison) from the perils of over-fitting models to data. For this reason, cross-validation, judiciously employed, presents itself as a valuable tool, not only for model selection, but also to gauge the impact that such selection has upon the uncertainty of the results.

Acknowledgements. The dataset used in Section 2 comprises aircraft measurements of CO₂ concentrations made by Paul B. Shepson and M. Obiminda Cambaliza (Department of Chemistry, Purdue University), during a curtain flight over Indianapolis, Indiana, on June 1st, 2011, within the scope of the INFLUX project [24].

Jolene Splett (Statistical Engineering Division, Information Technology Laboratory, NIST) provided many useful suggestions that stimulated much improvement of an early draft of this paper. The author is particularly grateful to Andrew Dienstfrey and Ronald Boisvert (Applied and Computational Mathematics Division, Information Technology Laboratory, NIST) for the opportunity of presenting this material at the International Federation for Information Processing

(IFIP) Working Conference on Uncertainty Quantification in Scientific Computing (August 1-4, 2011, Boulder, Colorado).

References

1. Baccam, P., Beauchemin, C., Macken, C.A., Hayden, F.G., Perelson, A.S.: Kinetics of influenza a virus infection in humans. *Journal of Virology* 80(15), 7590–7599 (August 2006)
2. Cleveland, W.S.: Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74, 829–836 (1979)
3. Cleveland, W.S., Devlin, S.J.: Locally-weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association* 83, 596–610 (June 1988)
4. Cressie, N., Wikle, C.K.: *Statistics for Spatio-Temporal Data*. John Wiley & Sons, Hoboken, New Jersey (2011)
5. Diaconis, P., Shahshahani, M.: On nonlinear functions of linear combinations. *SIAM Journal on Scientific and Statistical Computing* 5(1) (March 1984)
6. Diggle, P.J., Ribeiro, P.J.: *Model-based Geostatistics*. Springer, New York, NY (2010)
7. Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Chapman & Hall, London, UK (1993)
8. Friedman, J.H., Stuetzle, W.: Projection pursuit regression. *Journal of the American Statistical Association* 73(376), 817–823 (December 1981)
9. Friedman, J.H., Tukey, J.W.: A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers* C-23(9), 881–890 (September 1974)
10. Gilbert, P.: numDeriv: Accurate Numerical Derivatives (2011), <http://CRAN.R-project.org/package=numDeriv>, r package version 2010.11-1
11. Hastie, T., Tibshirani, R., Friedman, J.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer-Verlag, New York, second edn. (2009)
12. Hindmarsh, A.: ODEPACK, a systematized collection of ODE solvers. In: Stepleman, R.S., Carver, M., Peskin, R., Ames, W.F., Vichnevetsky, R. (eds.) *Scientific Computing: Applications of Mathematics and Computing to the Physical Sciences*, IMACS Transactions on Scientific Computation, vol. 1, pp. 55–64. North-Holland, Amsterdam (1983)
13. Joint Committee for Guides in Metrology: Evaluation of measurement data — Guide to the expression of uncertainty in measurement. International Bureau of Weights and Measures (BIPM), Sèvres, France (September 2008), <http://www.bipm.org/en/publications/guides/gum.html>, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 100:2008, GUM 1995 with minor corrections
14. Joint Committee for Guides in Metrology: Evaluation of measurement data — Supplement 1 to the “Guide to the expression of uncertainty in measurement” — Propagation of distributions using a Monte Carlo method. International Bureau of Weights and Measures (BIPM), Sèvres, France (2008), <http://www.bipm.org/en/publications/guides/gum.html>, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 101:2008
15. Joint Committee for Guides in Metrology: International vocabulary of metrology — Basic and general concepts and associated terms (VIM). International Bureau of Weights and Measures (BIPM), Sèvres, France (2008), <http://www.bipm.org/en/publications/guides/vim.html>, BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML, JCGM 200:2008

16. Krige, D.G.: A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa* 52, 119–139 (1951)
17. Loader, C.: *Local Regression and Likelihood*. Springer-Verlag, New York (1999)
18. Loader, C.: *locfit: Local Regression, Likelihood and Density Estimation* (2010), <http://CRAN.R-project.org/package=locfit>, r package version 1.5-6
19. Matérn, B.: *Spatial Variation*. Lecture Notes in Statistics, Springer-Verlag, New York, NY, 2nd edn. (1987)
20. Mosteller, F., Tukey, J.W.: *Data Analysis and Regression*. Addison-Wesley Publishing Company, Reading, Massachusetts (1977)
21. Pebesma, E., Cornford, D., Dubois, G., Heuvelink, G., Hristopoulos, D., Pilz, J., Stoeckler, U., Morin, G., Skoien, J.: INTAMAP: the design and implementation of an interoperable automated interpolation web service. *Computers & Geosciences* 37, 343–352 (March 2011), <http://dx.doi.org/10.1016/j.cageo.2010.03.019>
22. R Development Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria (2011), <http://www.R-project.org>, ISBN 3-900051-07-0
23. Ribeiro, P.J., Diggle, P.J.: *geoR: a package for geostatistical analysis*. *R-NEWS* 1(2), 14–18 (June 2001), <http://CRAN.R-project.org/doc/Rnews/>, iISSN 1609-3631
24. Shepson, P., Cambaliza, M.O., Davis, K., Gurney, K., Lauvaux, T., Miles, N., Richardson, S., Sweeney, C., Turnbull, J.: The INFLUX project: Indianapolis as a case study for the accurate and high resolution determination of CO₂ and CH₄ emission fluxes from an urban center. Poster (December 2010), American Geophysical Union Fall 2010 Meeting, San Francisco, CA
25. Silverman, B.W.: *Density Estimation*. Chapman and Hall, London (1986)
26. Soetaert, K., Petzoldt, T., Setzer, R.W.: Solving differential equations in r: package deSolve. *Journal of Statistical Software* 33(9), 1–25 (2010), <http://www.jstatsoft.org/v33/i09>
27. Stein, M.L.: *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Verlag, New York, NY (1999)

DISCUSSION

Speaker: Antonio Possolo

Mark Campanelli : What can be said about extrapolating the influenza model to make predictions about higher risk populations (such as asthmatics, children, and the elderly), when data from lower risk populations were (presumably) used to generate the time series data from planned infections?

Antonio Possolo : The model reviewed in the presentation is generic in its formulation, and the simplest of those described by Prasith Baccam and collaborators in the 2006 article quoted in my presentation and in my corresponding contribution to these proceedings. Its purpose is to describe the progression of the disease in a particular patient, once its adjustable parameters have been estimated based on observations made of the viral load carried by the patient on different days post-infection.

Van Snyder : Are the measurements of CO₂ in Indianapolis operational *in situ* measurements, or for validation of a remote sensing system?

Antonio Possolo : The measurements of the concentration of CO₂ made in the Indianapolis area in the context of the INFLUX experiment, serve to demonstrate the operation of a multimodal, pilot network to monitor abundances of greenhouse gases, not to validate a remote sensing system.

Ruth Jacobsen : For the measurements of greenhouse gas emissions: what were the detection limits of the instrument?

Antonio Possolo : For the measurements of concentration of CO₂ that I mentioned in my presentation, my colleagues from Purdue University that participate in the INFLUX experiment, Paul Shepson and Obie Cambaliza, use a Picarro gas analyzer. The performance specifications of Picarro's analyzers are listed in data sheets published at www.picarro.com. Please note that the mention of specific products, trademarks, or brand names in the answer to this question is for purposes of identification only, and ought not to be interpreted in any way as an endorsement or certification of such products or brands by the National Institute of Standards and Technology.

Mladen Vouk : This presentation emphasizes R. Is there an issue with SAS, or Genstat, or SPSS (or other packages) regarding functions you used or customization?

Antonio Possolo : I have little first-hand knowledge of SAS and SPSS, and none of Genstat. My understanding, based on the opinions of colleagues who use them, is that all these products include reliable implementations of many statistical procedures. I find R very well-suited to my needs as a statistician, not only for the functionality that it offers, or for how well it facilitates prototyping new ideas or customizing existing procedures, but also for the scrutiny that a very large base of sophisticated, inquisitive users, constantly subject R to (including

its source code), and, in the process, ensure its quality. Please note that the mention of specific products, trademarks, or brand names in the answer to this question is for purposes of identification only, and ought not to be interpreted in any way as an endorsement or certification of such products or brands by the National Institute of Standards and Technology.

Maurice Cox : On interpolation, you have measurement variances associated with the data points. You also apply a weight function that decays away from the point of interest. How can you justify that metrologically/statistically?

Antonio Possolo : Local regression is a statistical procedure with a long history and commendable performance characteristics. The work of Bill Cleveland, Susan Devlin, and of Catherine Loader, that I reference in my contribution to these proceedings, among a large body of related work, include several results on the optimality of this class of procedures under fairly general conditions.

Jon Helton : An observation: additional examples of the use of nonparametric regression techniques in uncertainty/sensitivity analysis are given in a sequence of articles by Curt Storlie *et al.* in *Reliability Engineering and System Safety*.