On the Estimation of Missing Data in Incomplete Databases: Autoregressive Bayesian Networks

Pablo H. Ibargüengoytia, Uriel A. García
Electric Research Institute
Reforma 113, Palmira,
Cuernavaca, Mor., 62490, México
{pibar,uriel.garcia}@iie.org.mx

Javier Herrera-Vega, Pablo Hernández-Leal,
Eduardo F. Morales, L. Enrique Sucar, Felipe Orihuela-Espina
National Institute for Astrophysics, Optics and Electronics
Luis Enrique Erro 1,
Tonantzintla, Puebla, México
{vega,pablohl,emorales,esucar,f.orihuela-espina}@ccc.inaoep.mx

***Abstract*—Missing data can be estimated by means of interpolation, time series modelling, or exploiting statistically dependent information. The limits of when one approach is preferable to the alternatives have not been explored, but are likely to be a compromise between a signal autoregressive information, availability of future observations, stationary behaviour and the strength of the dependence with concomitant information. This paper takes a first step towards clarifying dataset characteristics delimiting the realm of application for each technique. In addition, this paper introduces autoregressive Bayesian networks (AR-BN), a variant of Dynamic Bayesian Networks for completing databases which exploits latent variable relations while still benefitting from autoregressive information of the variable being filled. Using AR-BN, new estimated values are calculated using inference in the dynamic model. Our results unveil how the interplay between the variable autoregressive information and the variable relationship to others in the dataset is critical to selecting the optimal data estimation technique. AR-BN appears as a good candidate ensuring a consistent performance across scenarios, datasets and error metrics.**

***Keywords*-dynamic probabilistic graphical models; incomplete data series; value estimation; knowledge discovery; autoregressive models.**

I. INTRODUCTION

Instrumentation failure, human error or interferences during data storing give rise to a number of disagreements between the real data and the repository. A wealth of literature is available on data validation methods for addressing any of the more common issues; outliers [1], [2], [3], [4], [5], [6], [7], sudden changes [8], [9], [4], [5], [6], [7], rogue values [10] and missing data [11], [12], [13]. Moreover, full data validation suites can be envisaged [14], [15], [16]. This paper concentrates in estimation of missing data.

Missing data can be estimated by means of interpolation, time series modelling, or exploiting statistically dependent information. Arguably the most widely applied methods are the various interpolation techniques [17] together with classical time series modelling [18] such as ARMA or ARIMA. Interpolation and time series modelling are appropriate for isolated time series. In isolated time series missing data is reconstructed exploiting within-variable information. The simplest method will replace missing data with the distribution average [12]. However, in complex databases, statistical dependencies among variables can be further exploited to fill information gaps. Hence a number of techniques have been developed to make the most out of this dependent information. Vagin and Fomina [12] proposed a method based on nearest neighbour. This consists in the definition of a metric that relates the similarity between different variables in a database. Lamrini et al [11] applied self organizing maps for reconstructing data from monitoring a water treatment process with a complex sensor configuration. In another example, a virtual sensor estimates the value of the fuel oil viscosity using related variables of a combustion process in power plants [19]. It can be argued, that these later approaches exploit information present in adjacent variables at the cost of ignoring any signal own information. It is yet unknown when the dataset characteristics will favour application of one technique over another. Moreover, it can be hypothesized that a method that utilizes both sources of information, the signal internal information and the related information present in the repository, will achieve data reconstruction with high accuracy.

This paper aims at demarcating the dataset characteristics advocating for the application of a particular approach for estimating missing data in incomplete dataset. Towards better exploiting the presence of both aforementioned sources of information, we further introduce the autoregressive Bayesian networks (AR-BN), a variant of Dynamic Bayesian Networks (DBN) for incomplete data estimation based upon dynamic probabilistic modelling. AR-BN exploits statistical dependencies among related variables as well as the variables' autoregressive information. The main contribution of this work is twofold. First, a preliminary exploration of dataset features hypothesized to guide the decision on missing data estimation process, as well as a comparative in missing data error reconstruction from a range of techniques. Second, the enrichment of a static probabilistic model with previous and posterior values of the variable set to afford an autoregressive probabilistic model. The reconstruction capabilities of the technique are demonstrated in two datasets from different domains.

II. METHODS

A beautiful mathematical formulation of the problem of incomplete data can be found in [13].

 ISBN: 978-1-61208-246-2

A. Autoregressive models

Autoregressive models are linear systems for predicting future values of a time series based on previous observations. The general autoregressive model of order n -denoted AR(n)- is defined as:

$$X_t = c + \sum_{i=1}^{n} \alpha_i X_{t-i} + \epsilon_t \quad (1)$$

where $\alpha_{i|i=1...n}$ are the model parameters, c is a constant and ϵ_t is noise. In an off-line repository, future data may also be available and can be easily incorporated:

$$X_t = c + \sum_{i=1}^{n} \alpha_i X_{t-i} + \sum_{j=1}^{m} \beta_j X_{t+j} + \epsilon_t \quad (2)$$

where $\alpha_{i|i=1...n}$ and $\beta_{j|j=1...m}$ are the AR(n,m) model parameters.

B. Probabilistic modelling

A Bayesian network is a directed acyclic graph (DAG) representing the joint probability distribution of all variables in a domain [20]. The topology of the network conveys direct information about the dependency between the variables. In particular, it represents which variables are conditionally independent given another variable.

Given the knowledge represented as a Bayesian network, it can be used to reason about the consequences of specific input data, by what is called probabilistic reasoning. This consists of assigning a value to the input variables, and propagating their effect through the network to update the probability of the hypothesis variables. The updating of the certainty measures is consistent with probability theory, based on the application of Bayesian calculus and the dependencies represented in the network. Several algorithms have been proposed for this probability propagation [20]. Bayesian networks can use historical data to acquire knowledge but may additionally assimilate domain experts' input.

Dynamic Bayesian Networks (DBN) are an attempt to add temporal dimension into the BN model [21], [22]. Often a DBN incorporates two models; an initial net B_0 learned using information at time 0, and the transition net $B_{\rightarrow}$ learned with the rest of the data. Together B_0 and $B_{\rightarrow}$ conform the DBN [23]. An important assumption is made for DBNs; the process is Markovian, this is, the future is conditionally independent of the past given the present. This assumption allows the DBN to use only the previous time stage information in order to obtain the next stage. DBN can be unfolded over as many stages as necessary and the horizontal structure can change from stage to stage. The resulting network is highly expressive but often unnecessarily complicated. Alternatives have been proposed to reduce this complexity [24]. In datasets arising from physical processes, statistical dependencies among variables can be expected to be stable across time. That is, if two variables X and Y are statistically dependent at time t_i they will likely be also statistically dependent at time t_{i+j} for any arbitrary samples i and j, and similar reasoning can be made for independencies. This can be exploited to simplify the network topology.

C. Autoregressive Bayesian Networks

Autoregressive Bayesian Networks are a simplified variant of DBNs. They incorporate the temporal dimension by observing time-shifted versions of the variables, whether past or future, therefore the Markovian assumption is not needed. Conceptually they can be regarded as bringing an autoregressive model AR(n,m) to the BN domain.

Suppose a time series dataset. Figure 1 illustrates the proposed probabilistic model. Variable X represents the variable to be estimated, variables Y and Z represent pieces of Bayesian network corresponding to all the related variables to X. X_post represents the value of variable X at the time $t+1$, and X_ant represents the value of variable X at the time $t-1$.

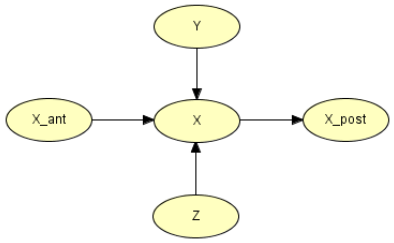


Fig. 1. Dynamic probabilistic model proposed for data estimation.

This proposed model represents a dynamic model which provides accurate information for estimating the variable in two senses. First, using related information identified by automatic learning algorithms or experts in the domain, or both. Second, using information of the previous and incoming values. This information includes the change rate of the variable according to the history of the signal.

In this approach the horizontal (inter-stage) topology of the network is fixed. The persistency arcs among a variable and its shifted versions are enforced, whereas those between different variables at different stages are forbidden.

D. Estimating missing data from incomplete databases using AR-BN

Summarizing, the proposed procedure for estimating missing data from incomplete databases is in Algorithm 1. The first 4 steps build the model, the last 3 propagate knowledge to estimate data holes.

III. Experiments and Results

Simulations were carried out to reconstruct missing data from 2 different industrial datasets of different nature (variables have been enumerated for confidentiality). The first dataset comprises 10 variables. It corresponds to a manufacturing process. The second dataset comprises 3 variables. It corresponds to an energy domain. Intrinsic dimensionality of the datasets as found by Principal Component Analysis is 7 and 1 respectively (99% of variance included). For the

 ISBN: 978-1-61208-246-2

Algorithm 1 Estimation of missing data

1: Obtain a complete data set that includes information from the widest operational conditions of the target process.
2: Clean the outliers and discretize the data set.
3: Utilize a learning algorithm that produces the static Bayesian network relating all the variables in the process. During the learning process a complete train set with data from all variables is needed as indicated in step 1.
4: Modify the static model to include previous and posterior values of every variable.
5: For all registers in an incomplete database, if one value is missing, instantiate the rest of the nodes in the model.
6: Propagate to obtain a posterior probability distribution of the missing value given the available evidence.
7: Return the estimated value with the value of the highest probability interval, or calculate the expected value of the probability distribution.

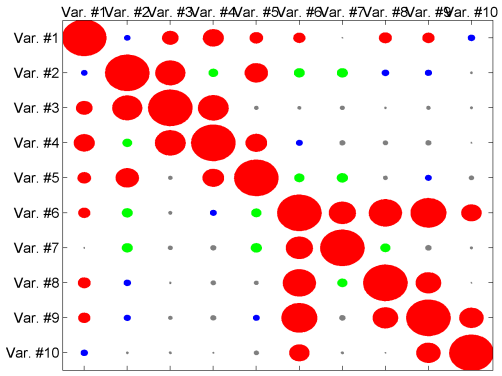
dataset 2, the scale of one of the variables is 5 orders of magnitude larger than the remaining 2 variables. Hence, the global intrinsic dimensionality is perceived to be 1 by PCA, but local dimensionality of the dataset remains 3, which can be determined by the Fukunaga and Olsen's algorithm [25]. The pairwise Pearson correlations among variables for the datasets in Fig. 2 hint about the dependencies among variables. Variables autoregressive order n was estimated using the Akaike Information Criterion [26] providing an indication of the signal own predictability. The autoregressive orders found with this criterion are summarised in Table I. Stationarity of the time series was estimated using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test for stationarity and is summarised in Table II.

TABLE I
AUTOREGRESSIVE ORDERS AS CALCULATED WITH THE AKAIKE INFORMATION CRITERION.

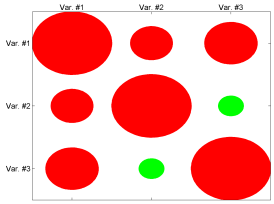
Var.#	1	2	3	4	5	6	7	8	9	10
Dataset 1	2	2	1	2	2	9	7	9	9	9
Dataset 2	25	1	25							

TABLE II
KWIATKOWSKI-PHILLIPS-SCHMIDT-SHIN (KPSS) STATIONARITY TESTS. ** INDICATES A HIGHLY SIGNIFICANT VALUE ($p < 0.01$). * INDICATES A SIGNIFICANT VALUE ($p < 0.05$).

Var.#	**Dataset 1**	**Dataset 2**
1	$p < 0.01$**	$p = 0.01$*
2	$p < 0.01$**	$p = 0.014$*
3	$p < 0.01$**	$p = 0.01$*
4	$p < 0.01$**	
5	$p < 0.01$**	
6	$p < 0.01$**	
7	$p = 0.04061$*	
8	$p = 0.05843$	
9	$p = 0.04314$*	
10	$p = 0.02301$*	



(a) Dataset 1



(b) Dataset 2

Fig. 2. Pairwise Pearson correlations among variables for the datasets. Circle size is proportional to correlation coefficient r. Circle color indicates significance: gray, non significant; blue, $p < 0.05$; green, $p < 0.001$; red, $p < 0.0001$;

From the datasets, specific samples were hidden to simulate missing values in three different fashions:

- Random Missing Data (**RMD**): Ghosted samples were chosen at random. Ghosted data accounts for 10% of each variable.
- Random Missing Blocks (**RMB**): Ghosted samples were chosen in blocks to have consecutive subseries of missing data. Ghosted data accounts for 10% of each variable. However, the location of the ghosted block and the number of blocks is random.
- All Missing Data (**AMD**): One full variable was ghosted at a time. Reconstruction can only occur from related information.

For each fashion, 10 train/test pairs were prepared for a 10-fold cross-validation. Note that the AMD has d test for each train case where d correspond to the number of variables in the dataset. After preparation of the ghosted test datasets, reconstruction was attempted by means of the following techniques:

- Static Bayesian Network (BN). Discretization was set to 5 equidistant intervals. Structure was learned using the PC algorithm [27].
- Autoregressive Bayesian Network (AR-BN). Autoregression order was fix to $< p, q >=< 1, 1 >$. Vertical (intra-stage) structure was learned using the PC algorithm. Equidistant intervals were used at all times, with the number of intervals being either 4 or 5 as bounded by memory limitations. The exemplary network for the Dataset 2 is illustrated in Fig. 3

 ISBN: 978-1-61208-246-2

- Linear interpolation (LI).
- Cubic spline interpolation (CSI).
- Autoregressive Models (AR(1))
- Autoregressive Models (AR(n)). Order n was chosen according to Table I. Notwithstanding, during the preparation of the train/test sets, some of the test sets did contain a number of samples lower than the autoregressive order i.e. AR order 25 for Dataset 1 variables 1 and 3. In those cases, the highest possible order was chosen based on the number of available samples.

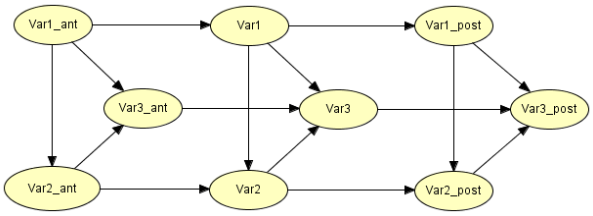


Fig. 3. Autoregressive Bayesian network proposed for data estimation for the Dataset 2.

In order to establish the accuracy of the estimation of the missing value the following error metrics were computed [28]:

Let E_i be the relative deviation of an estimated value x_i^{est} from an experimental value x_i^{obs}:

$$E_i = \left[\frac{x_i^{obs} - x_i^{est}}{x_i^{obs}}\right] \times 100 \qquad i = 1, 2, \ldots, n \tag{3}$$

with n being the number of missing data.

- **Root Mean Square Error**:

$$e_{rms} = \left[\frac{1}{n}\sum_{i=1}^{n} E_i^2\right]^{1/2} \tag{4}$$

- **Average Percent Relative Error**:

$$e_r = \frac{1}{n}\sum_{i=1}^{n} E_i \tag{5}$$

- **Average Absolute Percent Relative Error**:

$$e_a = \frac{1}{n}\sum_{i=1}^{n} |E_i| \tag{6}$$

- **Minimum and Maximum Absolute Percent Relative Error**:

$$e_{min} = min_{i=1}^{n}|E_i| \tag{7}$$

$$e_{max} = max_{i=1}^{n}|E_i| \tag{8}$$

As indicated above, for each reconstruction technique and ghosting fashion, a 10-fold validation was made. Since AMD scenario can only be reconstructed from related information, this scenario cannot be resolved by interpolation or autoregressive models. In total, 280 simulations were executed using MATLAB and Hugin [29]. For 3 simulations mistakes in pipeline from training to test were detected and their results not included for further analysis. Statistical analysis was carried out in R. R is a language and environment for statistical computing and graphics. See http://www.r-project.org/.

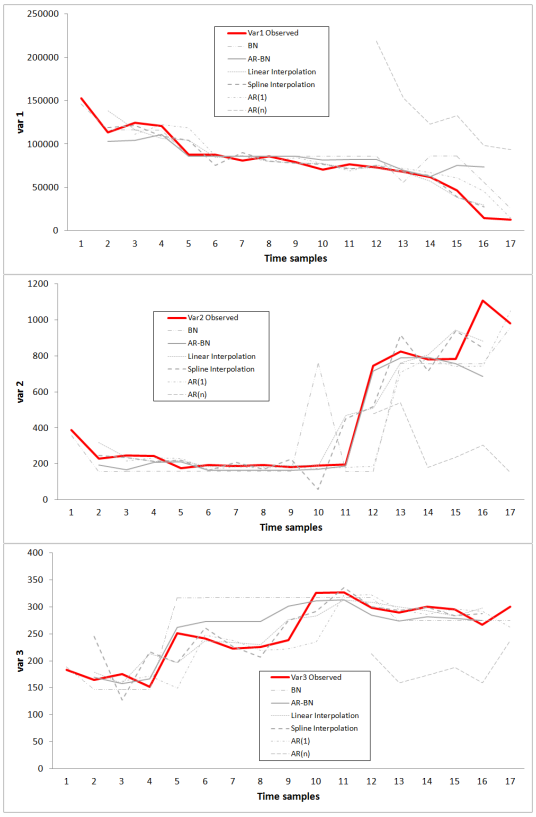


Fig. 4. Example of the missing data estimation using the different techniques. The example corresponds to the 3 different variables of Dataset 2 respectively for an RMD scenario. For this example, each sample of the time series is hidden one at a time, and the missing sample is estimated using the rest of the series as necessary by the different estimation techniques.

A. AR-BN performance

An example of the reconstruction with the different techniques is illustrated in Fig. 4. Fig. 5 summarizes the errors incurred by each technique, and Fig. 6 provides a more detailed view by dataset and error metric. From this detailed view, it can be appreciated that the proposed AR-BN achieves a good compromise in the reconstruction across different scenarios, datasets and error metrics. Unexpectedly, linear interpolation achieves better overall reconstruction than the more advanced spline interpolation. Classical autoregressive models achieve a reasonable performance but are highly unstable in their predictions as demonstrated by the large standard deviations coupled with disparate differences between e_{min} and e_{max}.

B. Limits of missing data estimation approaches

Fig. 7 relates the variable feature space given by the variable autoregressive order and its average relation to all other variables in its dataset (avg_r) against the dominant technique. The dominant technique is that which affords the lowest error in a particular region of the variable feature space. Regions are

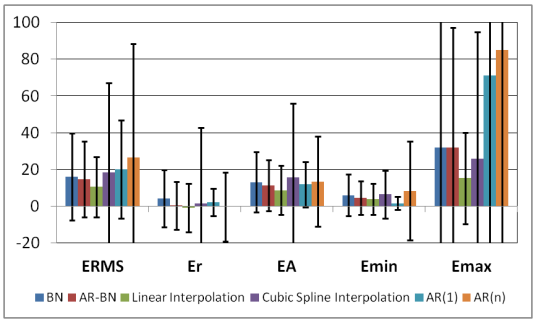


Fig. 5. Reconstruction errors incurred by each technique across datasets, scenarios, folds and variables. Bars and error lines correspond to average values and standard deviation respectively.

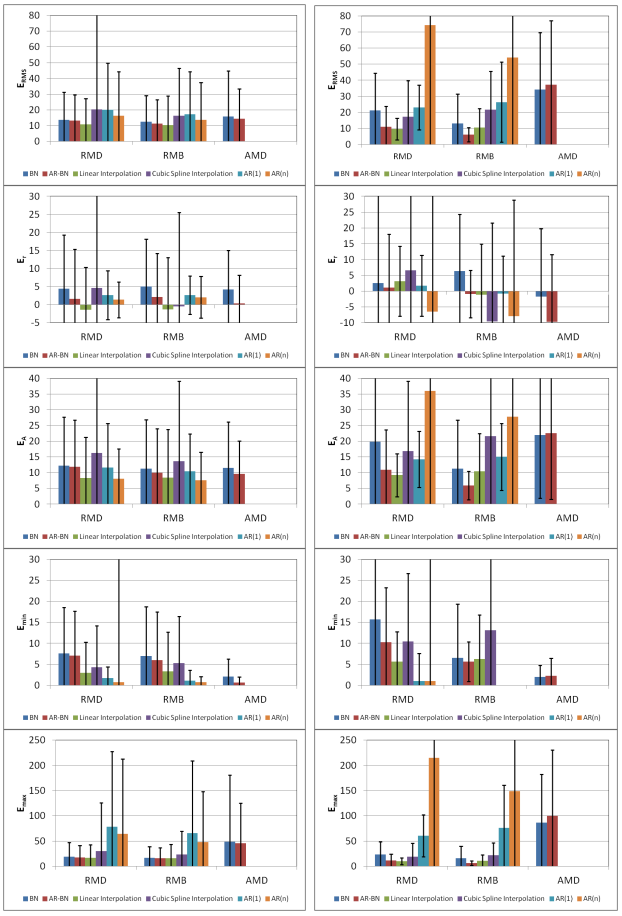


Fig. 6. Reconstruction errors incurred by each technique across folds and variables. Columns correspond to dataset; Left: Dataset 1; Right: Dataset 2; Rows correspond to different error metric: From top to bottom: e_{rms}, e_r, e_a, e_{min} and e_{max}. Bars and error lines correspond to average values and standard deviation respectively.

calculated using the Voronoi partition. It can be appreciated how the use of one technique over the other is subjected to the characteristics of the variable in terms of its autoregressive information as well as the amount of dependency that the variable share with fellow variables in the dataset as hypothesized. In particular, linear interpolation performs particularly well in these examples when the estimated autoregressive orders of the variables are low. When a full variable needs to be reconstructed from related information, it is obvious that the AR-BN dominance of the variable feature space grows as the autoregressive information does so.

IV. Conclusions and Future Work

We have explored the relation between a variable feature space represented by its autoregressive order and its relation to other variables in its dataset against different reconstruction techniques. Our results suggest that the interplay between the variables characteristics in the dataset dictates the most beneficial reconstruction option.

Unfortunately, the two datasets used in these simulations are not large enough to allow us exploring the variable feature space in detail. Yet some patterns start to be discernible. In particular, we have shown that the proposed AR-BN achieves a particularly competitive reconstruction regardless of the scenario, dataset and error metric used. Although we have reported signals stationarity for reproducibility, it has not further been considered for this paper. We believe signal stationarity will also be a critical element in the variable feature space supporting the decision over which estimation technique to use. Consequently, we plan to explore its effect.

The AR-BN model can be trivially extended to any level auto-regression and can be easily adapted for non-numerical data. In this sense, different autoregressive stages whether past or future must be added "in parallel" rather than "in series" so that these observations can be appreciated through the Markov blanket. We believe the proposed AR-BN profits from both within-variable information and statistical dependencies across variables, thus representing a valuable tool for the estimation of missing data in incomplete databases.

Acknowledgments

This research work is supported by the grant 146515 from CONACYT-SENER Sectorial Found.

References

[1] K. Hoo, K. Tvarlapati, M. Piovoso, and R. Hajare, "A method of robust multivariate outlier replacement," *Computers and Chemical Engineering*, vol. 26, pp. 17–39, 2002.

[2] B. Walczak, "Outlier detection in multivariate calibration," *Chemometrics and Intelligent Laboratory Systems*, vol. 28, pp. 259–272, 1995.

[3] J. Peng, S. Peng, and Y. Hu, "Partial least squares and random sample consensus in outlier detection," *Analytica Chimica Acta*, vol. 719, pp. 24–29, 2012.

[4] C. Muirhead, "Distinguishing outlier types in time series," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 48, no. 1, pp. 39–47, 1986. [Online]. Available: http://www.jstor.org/stable/2345636

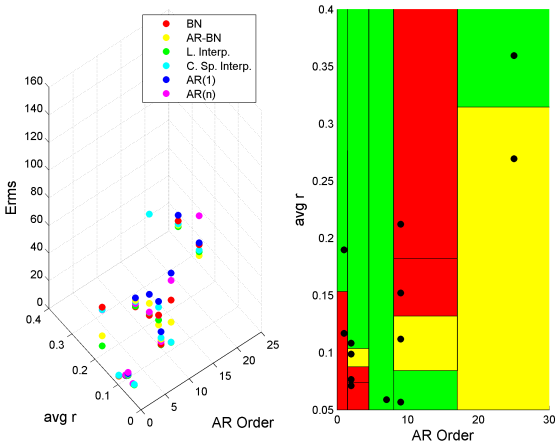
[5] R. S. Tsay, "Outliers, level shifts, and variance changes in time series," *Journal of Forecasting*, vol. 7, pp. 1–20, 1988.

[6] N. S. Balke, "Detecting level shifts in time series," *Journal of Business & Economic Statistics*, vol. 11, no. 1, pp. 81–92, 1993. [Online]. Available: http://www.jstor.org/stable/1391308

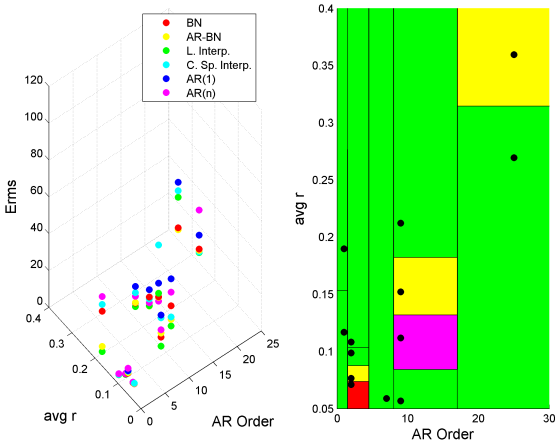
[7] B. Abraham and G. E. P. Box, "Bayesian analysis of some outlier problems in time series," *Biometrika*, vol. 66, no. 2, pp. 229–236, AUG 1979. [Online]. Available: http://www.jstor.org/stable/2335653

[8] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society London B*, vol. 207, pp. 187–217, 1980.

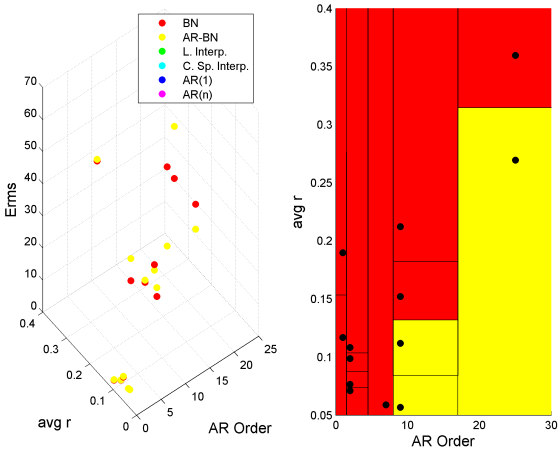
 ISBN: 978-1-61208-246-2



(a) RMD



(b) RMB



(c) AMD

Fig. 7. Relation between the variable feature space and the techniques for the three different scenarios. Left: Scatter plots of the variable feature space versus the error for each variable reconstructed through different techniques. The technique that achieves the lowest error is considered to dominate the region of the variable feature space. In order to determine the region of the variable feature space the different feature vectors for each of the variables for both datasets are used as seed vector quantizers for establishing a Voronoi partition. Each region of the Voronoi parcellation is then coloured according to the dominant technique.

[9] H. Sato, N. Tanaka, M. Uchida, Y. Hirabayashi, M. Kanai, T. Ashida, I. Konishi, and A. Maki, "Wavelet analysis for detecting body movement artifacts in optical topography signals," *NeuroImage*, vol. 33, pp. 580–587, 2006.

[10] J. Herrera-Vega, F. Orihuela-Espina, P. H. Ibargüengoytia, E. F. Morales, and L. E. Sucar, "On the use of probabilistic graphical models for data validation and selected application in the steel industry," *International Journal of Approximate Reasoning, In revision.*, 2012.

[11] B. Lamrini, E.-K. Lakhal, M.-V. Le Lann, and L. Wehenkel, "Data validation and missing data reconstruction using self-organizing map for water treatment," *Neural Computing and Applications*, vol. 20, pp. 575–588, 2011.

[12] V. Vagin and M. Fomina, "Problem of knowledge discovery in noisy databases," *International Journal Machine Learning & Cyber*, vol. 2, pp. 135–145, 2011.

[13] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.

[14] J. Herrera-Vega, F. Orihuela-Espina, E. Morales, and L. Sucar, "Validation data system based on a bayes network approach," in *Research Meeting on Dynamic Probabilistic Graphical Models and Applications*. Puebla, Mexico: FONCICYT, 3-4 JUN 2011.

[15] J. Herrera Vega, F. Orihuela-Espina, E. F. Morales, and L. E. Sucar, "A framework for oil well production data validation," in *Workshop on Operations Research and Data Mining (ORADM'2012)*, L. Villa-Vargas, L. Sheremetov, and H.-D. Haasis, Eds., Cancún, Mexico, 12-14 MAR 2012, pp. 226–235.

[16] P. H. Ibargüengoytia, S. Vadera, and L. Sucar, "A probabilistic model for information and sensor validation," *The Computer Journal*, vol. 49, no. 1, pp. 113–126, January 2006.

[17] P. Lancaster and K. Salkauskas, *Curve and surface fitting: an introduction*. Academic Press, 1986.

[18] C. Chatfield, *The analysis of time series: An introduction*. Boca Ratn: Chapman & Hall/CRC, 2004.

[19] P. H. Ibargüengoytia and M. A. Delgadillo, "On-line viscosity virtual sensor for optimizing the combustion in power plants," in *Advances in Artificial Intelligence - IBERAMIA 2010, LNAI 6433*, A.Kuri-Morales and G.Simari, Eds. Berlin Heidelberg: Springer-Verlag, 2010, pp. 463–472.

[20] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, San Francisco, CA., 1988.

[21] D. T. and K. K., "A model for reasoning about persistence and causation." *Computational Intelligence*, vol. 5, no. 3, pp. 142–150, 1989.

[22] V. Mihajlovic and M. Petkovic, "Dynamic bayesian networks: A state of the art," University of Twente. Dept. of Electrical Engineering, Mathematics and Computer Science (EEMCS), CTIT technical report series TR-CTI 36632, 2001. [Online]. Available: http://doc.utwente.nl/36632/

[23] D. Koller and N. Friedman, *Probabilistic graphical models: principles and techniques*. MIT press, 2009.

[24] P. Hernández-Leal, L. E. Sucar, and J. A. González, "Learning temporal nodes bayesian networks," in *Twenty-Fourth International Florida Artificial Intelligence Research Society Conference (FLAIRS'2011)*. Florida, USA: Association for the Advancement of Artificial Intelligence, 18-20 MAY 2011, pp. 608–613.

[25] K. Fukunaga and D. R. Olsen, "An algorithm for finding intrinsic dimensionality of data," *IEEE Transactions on Computers*, vol. 20, no. 2, pp. 176–183, 1971.

[26] L. Ljung, *System Identification: Theory for the User*. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[27] P. Spirtes, C. Glymour, and R. Sheines, *Causation, Prediction and Search*. Cambridge, Massachusetts. U.S.A.: MIT Press, 2000.

[28] E. A. Osman, O. A. Abdel-Wahhab, and M. A. Al-Marhoun, "Prediction of oil pvt properties using neural networks," in *SPE Middle East Oil Show*. Bahrain: Society of Petroleum Engineers (SPE), 17-20 MAR 2001, p. 14 pp.

[29] S. K. Andersen, K. G. Olesen, F. V. Jensen, and F. Jensen, "Hugin: a shell for building bayesian belief universes for expert systems," in *Proc. Eleventh Joint Conference on Artificial Intelligence, IJCAI*, Detroit, Michigan, U.S.A., 20-25 August 1989, pp. 1080–1085.

 ISBN: 978-1-61208-246-2