

Principles of MIR, multivariate image regression I: Regression typology and representative application studies

Thorbjørn T. Lied^{*}, Kim H. Esbensen¹

Department of Technology (TF), Applied Chemometrics Research Group (ACRG), Telemark University College (HIT), Kjølnes Ring 56,
N-3914 Porsgrunn, Norway

Abstract

We present an introduction to Multivariate Image Regression (MIR) with a selection of illustrative application studies. Generalisation from two-way multivariate calibration to the three-way regimen leads to—at least—three alternative *image regression cases* depending on the nature of the available Y-data: IPLS- Y_{discrim} ; IPLS- Y_{grid} ; IPLS- Y_{total} . A systematic *image regression typology* is briefly introduced.

We here present the core of the principles of *applied* MIR. Two major MIR application studies are worked through, a food mass product industrial inspection study (IPLS- Y_{discrim}) and a food product (fruit) storage stability image analytical monitoring (IPLS- Y_{grid}). These exemplifications are presented as *archetypes*, representing a much wider range of potential industrial/technological application areas. Based on simple three-channel imagery (in order to *simulate* many industrial systems), they nevertheless represent all higher-dimensional multivariate image cases as well, since the pertinent MIR principles and software are invariant w.r.t. any number of channels/variables employed.

The present paper represents one major element of our work towards establishing a complete, stand-alone facility for Multivariate Image Regression (MIR); the second paper in this series deals with the development, implementation and extensive exemplifications of a complementary cross-validation facility. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Multivariate image regression; MIR; Multivariate image analysis; MIA; Multivariate image texture analysis; MIX; 2-D images; 3-D image arrays; *Image regression cases*; Applications

1. Introduction

The introduction of the Multivariate Image Analysis (MIA) concept in chemometrics was not longer ago than Esbensen and Geladi [1]. In the intervening 10 + years, the development of MIA has been rela-

tively slow, but would appear to begin to take to speed more recently—the entire field was summarised in the comprehensive textbook by Geladi and Grahn [2]. Much of the theoretical background for Multivariate Image Regression (MIR) necessitates a thorough understanding of the principles and methods in MIA, which we shall here assume known. It is especially important that the concept of the *multivariate image* is well understood.

Because regression calculations on the extensive amount of data in multivariate images can be easily a technical challenge (growing quadratically with the

^{*} Corresponding author. Tel.: +47-901-85-661.

E-mail addresses: thorbjorn@lied.no (T.T. Lied), kes@aue.auc.dk (K.H. Esbensen).

¹ Aalborg University Esbjerg, Niels Bohrs Vej 8, DK-6700 Esbjerg, Denmark. Tel.: +47-79-12-76-88; fax: +47-75-45-36-43.

number of variables, or channels), some important recent evolutions have made it more feasible in the last few years. The continuously developing technical improvement in computer hardware efficiency is of course a major contributor, but especially the KERNEL PLS algorithm introduced by Lindgren et al. in 1993 [3] has dramatically speeded up the present type of calculations, as was outlined in detail in Ref. [5].

Despite of this, few studies have yet shown the true potential of MIR as a tool for predicting quantitative features in multivariate image data. Hopefully, this condition will be improved by the current paper. Below, only simple three-channel imageries are used (the rationale is in order to *simulate* many industrial systems), but these nevertheless represent all higher-dimensional multivariate image cases as well, since the pertinent MIR principles and software is invariant w.r.t. any number of channels/variables employed. This should not, however, lead any reader to conclude that the present exposition is aiming at contesting conventional three-channel image analysis, based on e.g. R/G/B relations, such as thresholding, etc.—quite the contrary is the case (sic). The examples chosen serve as master examples for *any* dimensional multivariate imagery (nos. channels 3 →).

1.1. Concepts

Several concepts are used in this paper, some of which may be relatively unfamiliar. A brief introduction of these is given to help the reading of the article.

1.1.1. Multivariate image (MI)

The MI is a digital image of *one* scene, consisting of *many* variables (channels), e.g. colour bands, channels. At the outset, the simplest situation is the one in which each image *pixel* is treated as an *object*, which requires rigid consistency in scene layout for all variables. An object in a given scene position in one variable must be found in the same scene position in all other channels; for regression cases also in the Y-image. MIs are usually presented as a 3-D matrix, but because the two object-ways can be treated as one way, the MI may also be reorganized into a 2-D-matrix prior to modelling, and two-way

methods can be applied [2] by way of the so-called *unfolding* operator.

1.1.2. MIR

Multivariate Image Regression [2,4] builds regression models between the multivariate X-image and the (uni-/multi-)variate Y-image. MIR is here performed using KERNEL-PLS [3,5] on reorganized Multivariate Images, i.e. each variable is reorganised into a (very long) object vector. In this basic unfolded form, MIR uses only the variable-signatures, i.e. the spectral information in the analysis and only indirectly makes use of the spatial information analogous to MIA [1,2]. But even though MIR technically uses two-way analytical methods, there is an enormous visualisation potential in image data that is also used fully in MIR. Displaying results not only in score space, but also in the so-called *backfolded* image space, enhances the insight in the data structure and the models developed. Applying colour coding to score plots, MIA, or by combining three-score images in one composite (“R/G/B”) colour image, it is often possible to capture comprehensive model presentations of great interpretation value, etc.

1.1.3. MIX

Multivariate Image texture analysis. MIX is an extended MIA–MIR approach which includes spatial, especially textural, information in the analysis. In cases where spatial information is important, this can be included in the MIR model by, e.g. adding derived textural variables calculated from the original variables [6–8]. Sometimes, enhancing details using, e.g. edge-detectors is favourable, in other cases the opposite (smoothing details) might be required, etc. In addition, *combining* textural filters might often give very useful results. MIX has the potential for explosive data growth, thus powerful means for variable selection are required. We do not treat the MIX aspects in any depth in the present work however.

1.1.4. Regression typology

Perhaps surprisingly, going from the 2-way realm in which the concepts of multivariate calibration is well-known—and need no further presentation here—there is a corresponding three-fold multitude of analogous but in principle different image *regression modes* (Fig. 1).

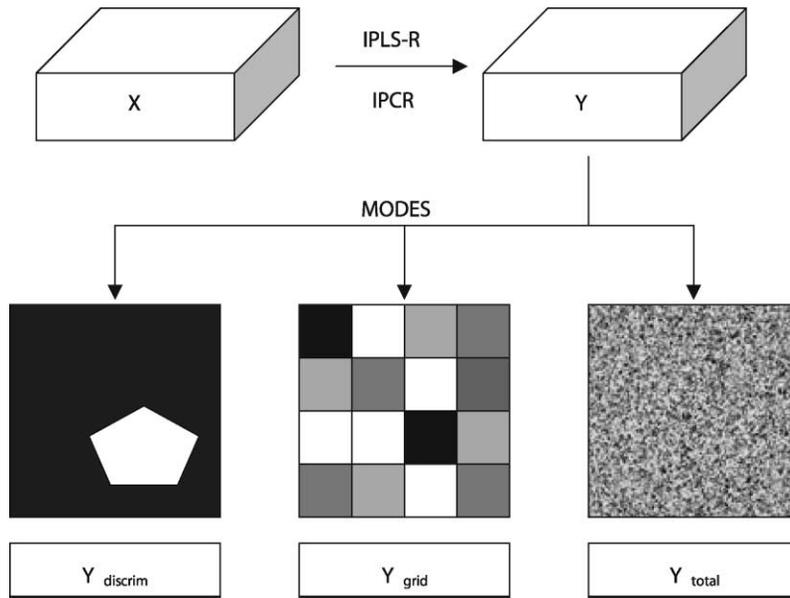


Fig. 1. The three different MIR modes, Y_{discrim} , Y_{grid} and Y_{total} .

1.1.4.1. $IPLS-Y_{\text{discrim}}$. The Yes/No classifier/discriminator. In every position in the Y-image, a pixel is either 1 (one) if it is part of a current class, otherwise 0 (zero). The approach is suitable for classifying one class among (many) others. Used as a pre-processor, this method can easily be taught how

to pick out desired classes. This case is also easily extended to cover several classes, by using several one-class Y-discrim masks (Fig. 1).

1.1.4.2. $IPLS-Y_{\text{grid}}$. If every Y-condition is not available in one image, several images can be juxtaposed

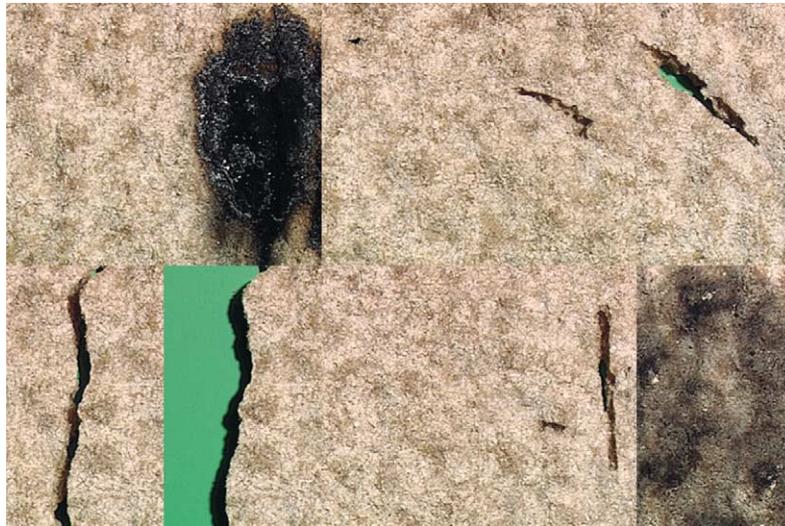


Fig. 2. “Normal” and “flawed” Scandinavian crispbread (“knekkebrød”). Three representative types of flaws are displayed; broken, perforated and burnt cases.

in a compound, so-called *gridded* image. This way the total experimental design can be represented in one image, i.e. one model. Extensive illustration of IPLS- Y_{grid} is given in this work. In some cases, especially when predicting an overall value for each sub-image in the grid, the corresponding Y-image will have a constant value within each sub-image. When this appears, some kind of smoothing of each sub-image in X will usually be useful, i.e. reducing non-classification related variations in X.

1.1.4.3. IPLS- Y_{total} . When the entire experimental design is covered in one frame, merging images together, as in the IPLS- Y_{grid} is not required. In these cases, each pixel in X also has a separate, unique value in the Y-image. Typical examples come from, e.g. remote sensing. Because most of the still limited MIR-literature explicitly discusses this kind of data, and because it is merely a special, extreme case of the Y_{grid} , it will not be treated further in this paper

1.2. Software

All calculations in this paper are performed using a self-developed program, described in Lied et al. [5].

Table 1
Technical imagery details of the crispbread case

| Image capture | Camera | Lens | Focal length |
|--------------------|--------------------|------------------------|--------------|
| | JVC 3CCD KYF-50 | Micro Nikkor AF | 105 mm |
| Measures | With (pixels) | Height (pixels) | #Variables |
| Total Image | 1000 | 666 | 4 |
| Sub Images | 200 | 333 | 4 |
| Spectral Variables | Colour | Wavelength | Bandwidth |
| 1 | Red | N/A | N/A |
| 2 | Green | N/A | N/A |
| 3 | Blue | N/A | N/A |
| Textural variables | Filter | Window size and passes | Applied to |
| 1 | Variance | | |

Table 2
Technical imagery details of banana image analysis set-up

| Image capture | Camera | Lens | Focal length |
|--------------------|---------------|------------------------|--------------|
| | SILVACAM | Fujinon | 120 mm |
| Measures | With (pixels) | Height (pixels) | #Variables |
| Total Image | 800 | 600 | 3 |
| Sub Images | 200 | 200 | 3 |
| Spectral variables | Colour | Wavelength (nm) | Bandwidth |
| 1 | NIR | 760–900 | |
| 2 | Red | 580–680 | |
| 3 | Green | 490–580 | |
| Textural variables | Filter | Window size and passes | Applied to |
| 1 | N/A | | |
| 2 | N/A | | |
| 3 | N/A | | |

The software is available for Microsoft Windows® (9X NT 4 + 5) and is written in National Instruments' LabVIEW v. 5.1. Both MIA and MIR is implemented; for MIR regression calculations, KERNEL-PLS [3] is used exclusively.²

2. Applications

Below, the terminology IPLS (Image PLS) is used throughout, but it is evident that PCR *may* also be used alternatively should one so desire, albeit with the well-known distinctions regarding PCR vs. PLS [9,10], etc. Here, PLS is employed exclusively because of its well-known chemometric advantages [9–14]. In both major examples below, data are mean centred and scaled to uniform standard deviation. All

² When developing this prototype, serious efforts were made to enhance the flexibility and user interaction facilities. For large datasets, 10 M pixels or above, calculations start to become slow however. Development of a professional system is now under way. Contact the corresponding author for instructions on how to download the freeware *prototype*.

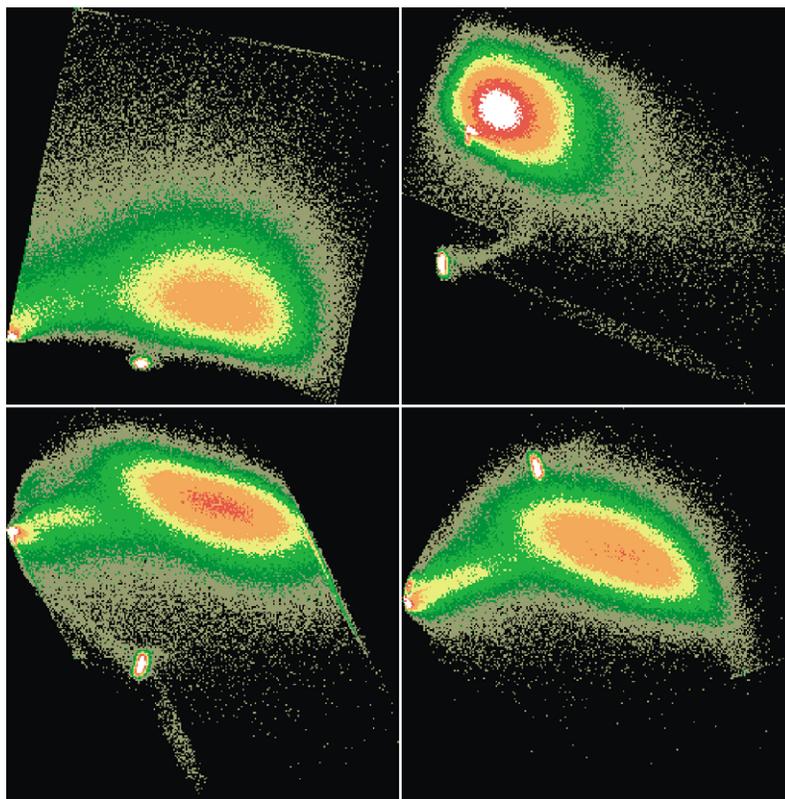


Fig. 3. MIA standard score-plot for the crispbread case in Fig. 2. Top row, left: t_1 vs. t_2 , right: t_2 vs. t_3 , bottom row, left: t_1 vs. t_3 , right: t_1 vs. t_4 . Clusters in these projected plots may overlap in some dimensions, while not in others. In addition, there is no guarantee that the relevant problem-dependent information is best described in a specific plot. It is thus useful to view several dimensions simultaneously, comp. [1–4].



Fig. 4. Y-image mask of (0/1) discrimination areas. Note that by using a relevant background discrimination feature, it is possible to zoom in only on the true flaws present in the gridded calibration imagery, which have been designated white here.

variables thus have equal variance weights, making the Kernel-PLS decompositions pertain to *correlations*.

2.1. $IPLS-Y_{discrim}$: discrimination prediction

Motivation: A *pilot study* of image analytical *industrial inspection* of a mass production food article, Swedish crispbread (“knäckebröd”) is presented. This item, by nature of its mass consumption status, is produced in very large quantities in industrial bakeries in many countries. Output from the industrial

ovens is necessarily way outside *complete* human inspection capabilities, for which reason an automated, industrial image analytical system would be of considerable interest. This in turn could form the basis for a truly 100% inspection system.

In our restricted pilot study involving some 10 pieces of crispbread, parallel representative oven outputs are available, 5 with an “accepted” status and 5 with three types of representative faults, typically encountered in the industrial production situation. Fig. 2 shows these three faults together with examples of the directly acceptable product (“normal”). Technical details regarding this image is found in Table 1.

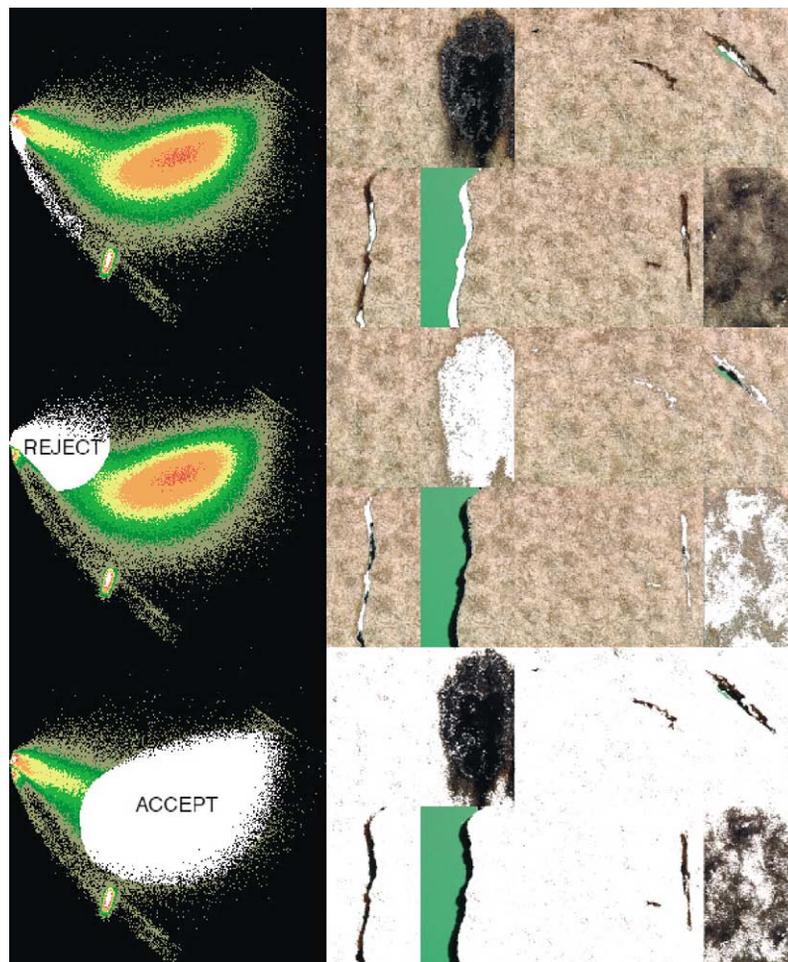


Fig. 5. $IPLS-Y_{discrim}$ t_1-t_4 score plots showing all three resolved classes in the crispbread case: broken/perforated (top panel); burnt (middle) and “accepted” (lower). Note complete discrimination. Also comp. with similar t_1-t_2 score plot from the MIA-solution in Fig. 3. X-variance: t_1 : 73.43%, t_4 : 1.18%, Corresponding Y-variance: 60.18% and 0.19%.

Rationale for choosing three-channel imagery for illustration purposes: All exemplifications used here employ a simple three-channel camera technology

(Tables 1 and 2), in order to *simulate* typical industrial image analytical systems (which certainly does not want to carry more channels than absolutely nec-

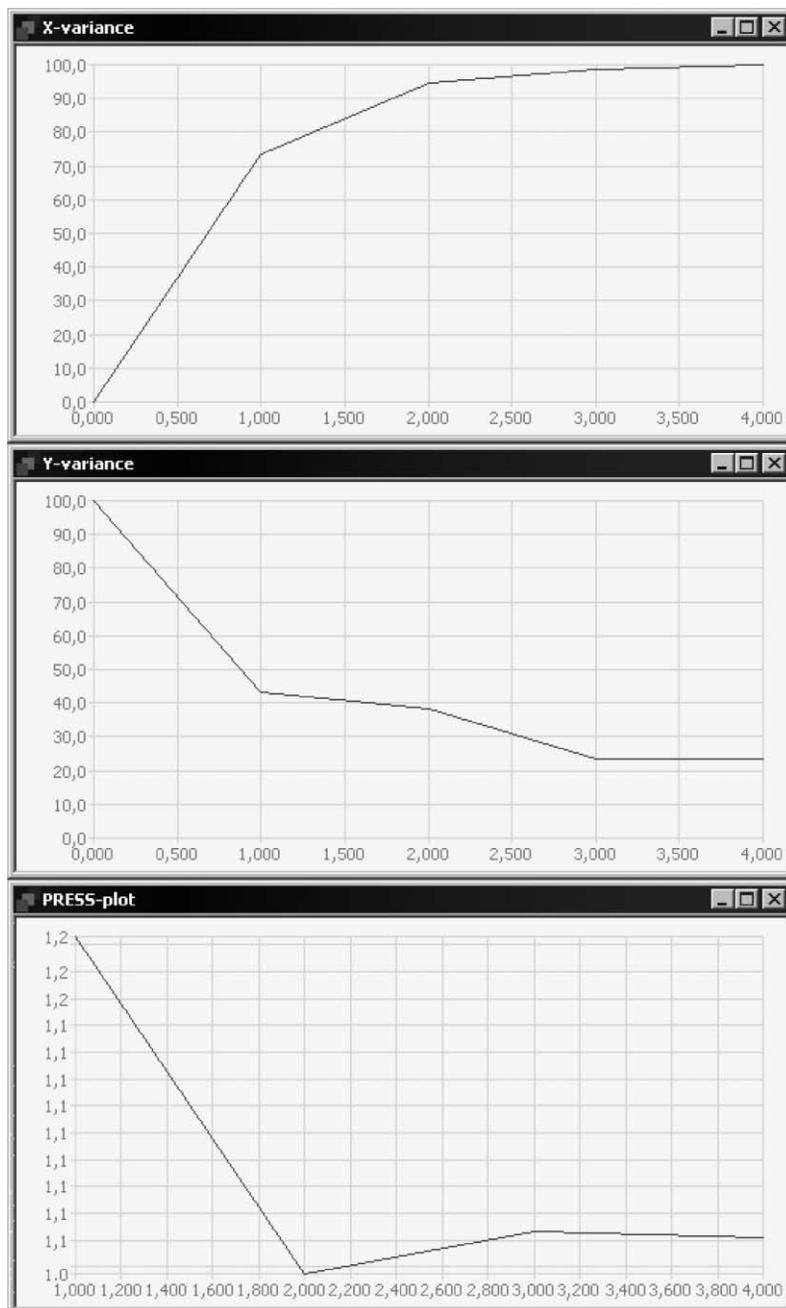


Fig. 6. Calibrated X-variance (top) and Y-variance (middle) and validated PRESS (low) for the crispbread case in Fig. 2.

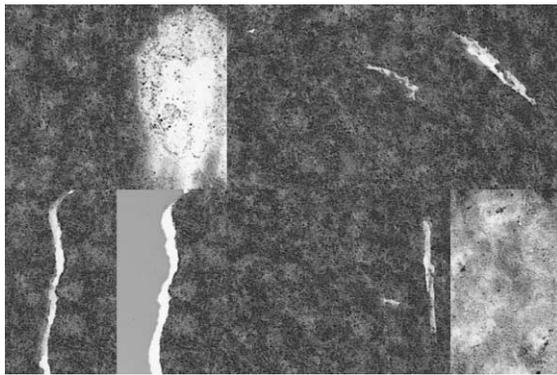


Fig. 7. Predicted \hat{Y} -image using two components. Note how the model distinguishes clearly between faults/non-faults.

essary). Whereas three channels constitute the absolute minimum number of variables in order to qualify for multivariate data analysis, it is of paramount importance to observe that this number is fully able to emulate all the essential features of, e.g. PCA, MIA, PLS-R and MIR also when using higher-dimensional data and image types, as we have shown extensively in our other studies [1,4,5,17,19]. The MIR software and principles illustrated and displayed here are completely invariant w.r.t. any number of channels, *ibid.*, and must thus be viewed as bona fide exemplars of any dimensionality.

Fig. 3 shows standard MIA score plot set-ups for the crispbread case (Fig. 2) [1,2].

In the interest of the wider application horizon for this specific IPLS- Y_{discrim} regression example, a more general IPLS- Y_{discrim} regression case will be set up. This is done by using a *Y-image mask* of the areas of interest in the image which are recognised as “rejects”, i.e. areas which are underlain by those parts of the *Y-image* which depicts flaws of the various type. Fig. 4 shows this “flaw”-mask.

Fig. 5 has been designed to bring forth the full potential of the IPLS- Y_{discrim} -case, showing (in standard MIA-style) the corresponding t_1 – t_4 score plots vs. the original (raw) image domain layouts of three fault classes present in the fully background-discriminated crispbread case: shadow (top); burnt (middle) and “accepted” (bottom). It is gratifying to observe *complete discrimination* between all relevant classes, i.e. all three types of rejects/accepted and the back-

ground as well. This successful discrimination points directly to the desired use of image-based *prediction* of all these types of crispbread. This pilot study, while extremely simple, allows full conceptual delineation of a complete automated image analysis system, by way of the relevant PLS-prediction facility (Fig. 7). The fact that the *problem-dependent information* in the current example is best described in the t_1 – t_4 score plot, illustrates that there is no knowledge prior to calibration or analysis of *where* or *how* this information will show up. In the present case, component directions t_2 and t_3 model communalities *not* relevant to the discrimination optimisation.

Incidentally, observe that in this particular case, there would appear to be very little “tilting” of the IPLS-solutions relative to the corresponding simpler PCA-solutions (MIA), contrary to many other two-way experiences [4,5,15,16]. In the present case, this reflects a rather direct correspondence of the X-block data structure with the Y-structure(s), i.e. the information gathered in the image analytical X-description “happens” to be directly correlated to the guiding Y-discrimination dummy variable; see also

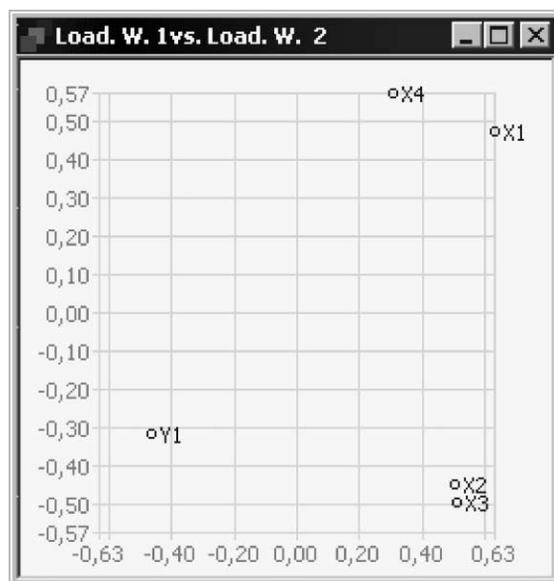


Fig. 8. Loading weights 1 vs. loading weights 2 for the crispbread case. Note X-variables 2 (green) and 3 (blue) which seem redundant.

below however. Regarding Fig. 6 (also relevant for Fig. 13), cross validation was used in order to calculate prediction statistics (PRESS). *How* and *why* this was applied, will not be discussed here, this is a topic for the second paper in this series [19]. Figs. 6 and 13 here serve to illustrate the use (fullness) of MIR in practise.

While Figs. 4–6 give the statistical facts in this case, Fig. 7 shows the actual predicted image in scene space using one component. This figure demonstrates that the model is excellent for predicting all the relevant types of faults. This result is, needless to say, of a much larger generalisation potential than the specific crispbread example chosen. The illustration in fact has merit as an *archetype* for IPLS- Y_{discrim} multivariate image regression.

MIX-aside: In this example, a variance filter (Table 1) has in fact also been applied to extract local textural variations in the X-image. While the background is flat, the crispbread has a very distinct, regular texture. The variance filter, which returns the local variance in a small window in every position in the image, will thus greatly help to *distinguish* crispbread from background as well as textureless burnt parts, assisting the spectral information in the classi-

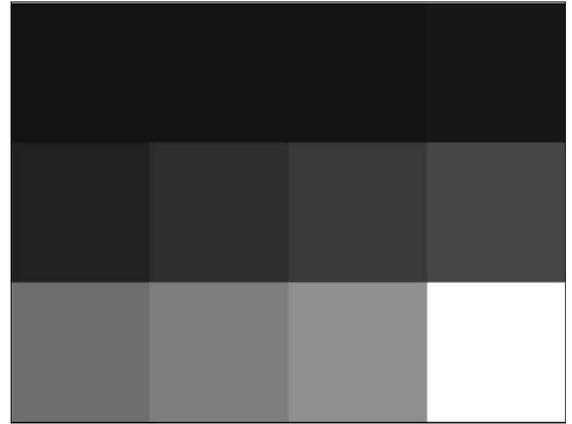


Fig. 10. IPLS- Y_{grid} Y-image, delineating the problem-dependent Y-levels for the banana deteriorating process. This figure (Y-image) corresponds to Fig. 9 (X-block).

fication. The effect of this filter addition is visualized in Fig. 8, which shows the *loading weights* (w) for this example. From this figure, it can also be seen that X2 and X3 (green and blue) contains mostly the same information, indicating that one of them could be left out in later calibrations, etc.

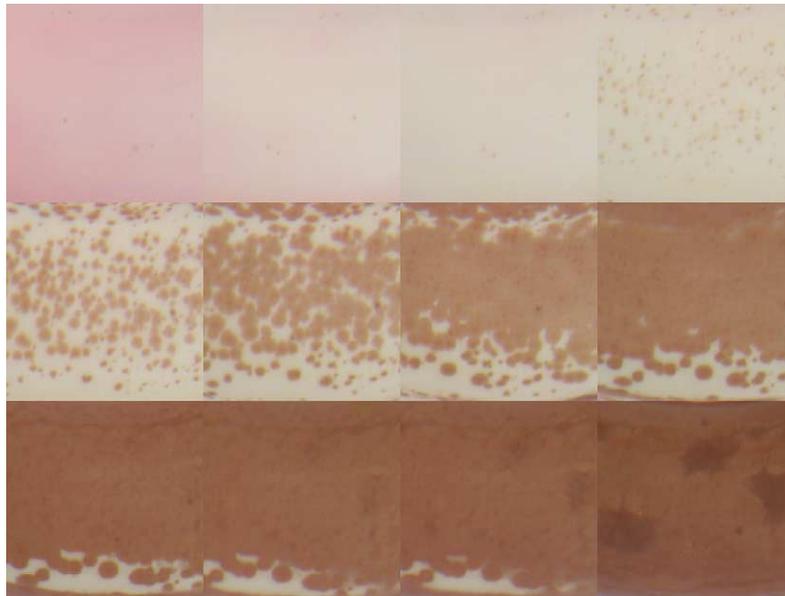


Fig. 9. Storage time *aging* of representative fruit (banana) in the interval 1–20 days.

Note how judicious use of a *relevant background* (colour, texture) is essential to bring about the successful discriminations in this case. As in all image analysis applications, *illumination* and *colouring* (not treated here) is often of equal importance compared to the data analysis proper, etc.

2.2. $IPLS-Y_{grid}$: monitoring and estimating storage time for fruit (bananas)

Motivation: The objective of this application example is to monitor storage stability by a series of (multi-temporal) images of the same fruit(s), with great efforts to keep all storage and imaging parameters *constant*, the only variable being *time elapsed* since storage start. Successful monitoring will allow for quantitative storage deterioration prediction directly from the captured multivariate X-images [5]. In the current example, the task is to get a quantitative measurement of the age of a banana, based on the colour given specific storing conditions. These quan-

titative measurements are based on subtle spectral variations in colour over the entire grid image fields, while the initially green banana is turning yellow and later brown. In this process, as can be seen in Fig. 9, the intensity of a colour is central for the calibration. The number of brown pixels increase, but individually they also turn darker in the process. Counting the number of, say “brown pixels”, would therefore not give the quantitative measurement requested, because the degree of brownness is not accounted for. Fruit deterioration is both a spectrally, as well as a spatially (contextually) related development process. A more subtle technique is thus required.

In this context, the calibration-parameter “storage time” shall be represented by juxtaposed part-images, making up a complete, so-called *gridded*, multivariate image, hence, the suggested name for this second image regression mode: $IPLS-Y_{grid}$. This example also serves as an *archetype* not only of multi-temporal studies but also of analogous objectives, conf. below.

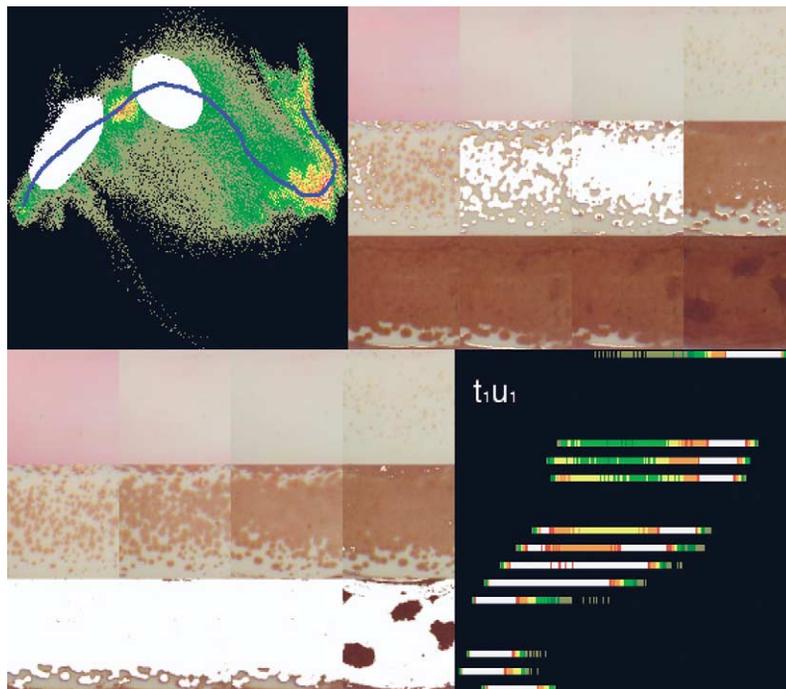


Fig. 11. IPLS of the fruit aging process, in the $IPLS-Y_{grid}$ regression case. Upper left: t_1-t_2 score plot, with two sets of corresponding scene-space (upper right and lower left panels). Lower right: corresponding t_1-u_1 score plot. Note discretisation along the u_1 -axis, corresponding to the Y-levels presented in Fig. 9.

Fig. 9 shows the compound, gridded multivariate X-image of a deteriorating banana, for which the storage times involved are (from upper left to lower right): 1/2/3/6; 7/8/9/10; 13/14/15/20 days, respectively. This gridded layout is necessary in order that all storage times can be *analysed together* by MIA or MIR. It is emphasized that it is the *objective* of the image analysis (in this case: storage stability monitoring) which *dictates* that the individual grids represent a succession of different storage times. For other image regression cases, these individual grid cells will often represent different, typical “object-like” categorical entities to be similarly *compared*, e.g. a series of different meats to be characterised, as was the case with Geladi and Esbensen [15] (in fact also predicting a storage-related parameter, “harshness”) and Geladi and Grahn [2] a.o.

In Fig. 10, which shows the particular *Y-image*, the *array* of grid cells forms the basis for an IPLS1. Observe how the deterioration process interval of 1–20 days has been *mapped* into an image analytical appropriate grey-level interval, spanning 0–255. Again, it is the *Y-image* mask that makes the regression problem immediately appreciable. In one sense, as soon as the multivariate X-image has been defined, it is the *Y-image* which sets up the entire MIR. In this example, the objective is to try to estimate the age of the banana, by way of a common *Y-value* for each sub-image. In the final prediction, the goal is to get an estimate of the *overall age* of the banana, and *not* the local variation at pixel-level. The *Y-grid* model will undoubtedly become more overall than one where local variations at pixel level is reflected in the *Y-image*, but this is not of significance for the

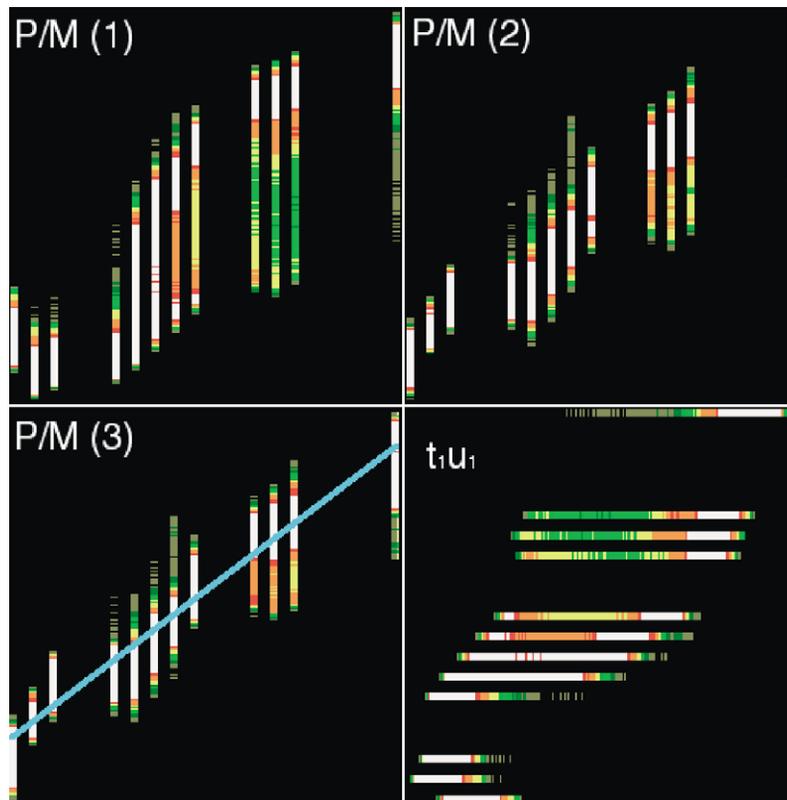


Fig. 12. Complete predicted vs. measured (P/M) layout of the banana aging process IPLS- Y_{grid} -analysis. The standard P/M assessment plot is shown for one, two and three IPLS-components, while only the t_1-u_1 score plot is shown (lower right). Significant improvement of prediction precision using three components.

ability to predict the age of an entire banana, however. The objective of an individual fruit inspection system of course has the *individual fruit* as its smallest “object-of-interest”.

Fig. 11 performs an identical role as Fig. 5, encompassing the essentials of the IPLS-analysis. In the t_1-t_2 score plot (upper left panel), one may appreciate, in full detail, the *trace* of the fruit deterioration

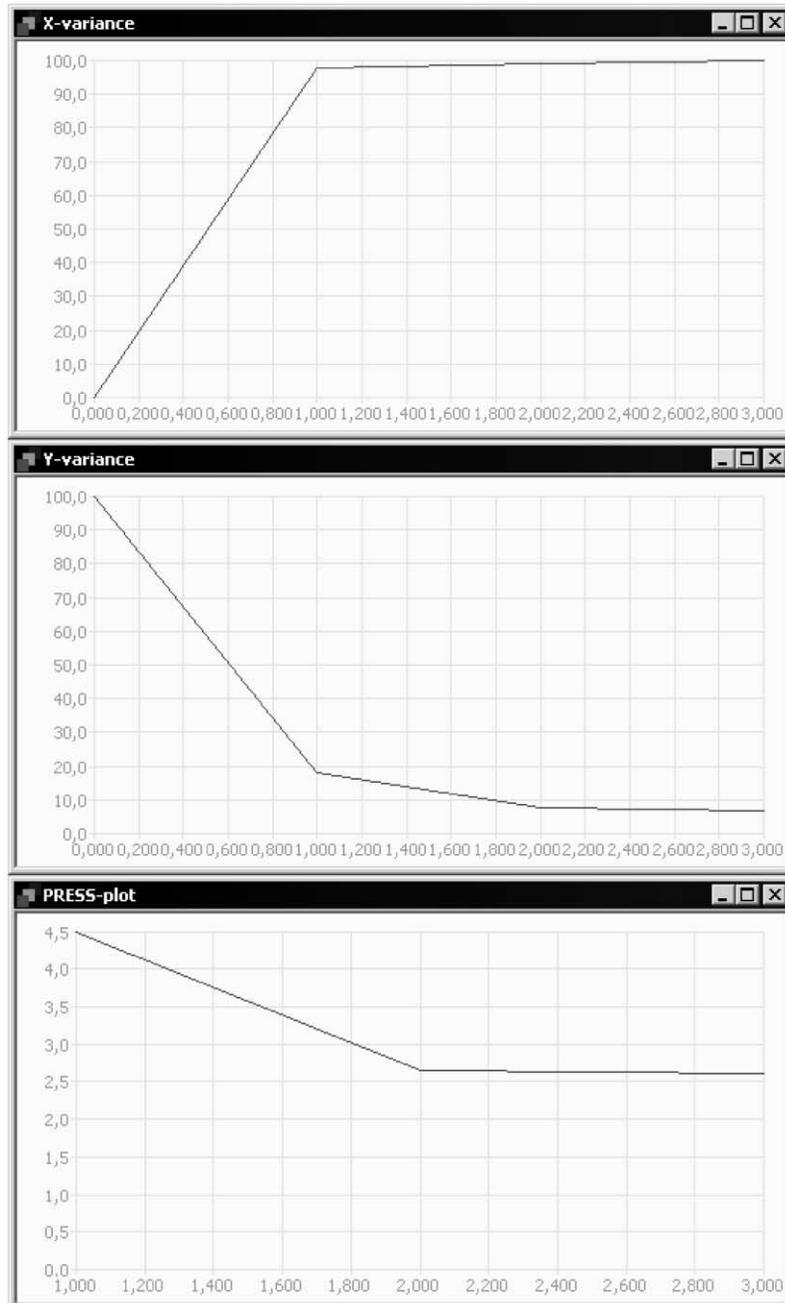


Fig. 13. Calibrated X-variance (top) and Y-variance (middle) and validated PRESS (bottom) for banana aging case.

process.³ We have illustrated two representative *process stages* along this trace, an intermediate stage and the penultimate sad, almost totally rotten end of the banana development (upper right and lower left panel, respectively). With reference to MIA [1,2,17], the scene-space back-projections of these two classes are self-explanatory in Fig. 11, especially when compared with Fig. 9. Fig. 11 shows how it is possible to delineate the *entire* deteriorating process in the X-space because the *entire* storage time *calibration span* has been compounded in the one (X,Y)-image.

For IPLS solutions, the *t–u score plot* allows valuable, indeed critical insight into the *effective* regression relationships between the X- and the Y-space [18]. For example if the t_1-u_1 relationships is (close to) already linear, this is a certain reflection that a strong prediction model will be achieved; likewise, smaller non-linearities in the t_1-u_1 score plots are usually “ironed out” by inclusion of one or a few, *additional* PLS-components t_2-u_2 , t_3-u_3 , etc.

For the present first presentation of the most used features in multivariate image regression, these few aspects of the general use of the *t–u*-plot will be enough to allow appreciation of the way the IPLS- Y_{grid} -modelling works. Figs. 11–13 represent salient central aspects of our work leading up to a complete *MIR strategy* [5,19].

For an assessment of the *modelling strength* of the IPLS-analysis, Figs. 11–13 will also suffice. From these X–Y relationships, it is evident that a satisfactory model can be achieved using three IPLS-components. Observe, e.g. how the P/M (*predicted vs. measured*) relationship improves quite considerably when adding the second- and third IPLS-compo-

nents. From the t_1-u_1 relationships *alone*, it was however already clear that this would *per force* result. We are also able to follow how one would go about identifying *outliers*, etc., by using the appropriate *t–u* score plots, following Ref. [18]. In the specific present plots in Fig. 11, we did not actually have reason to perform any outlier deletion, since none were found.

3. Discussion and conclusions

The examples above represents our first presentation of a *simulation* of an *automated image-analysis monitoring system* in the guise of IPLS: MIR (and MIX), in which we focused on the general aspects of the IPLS- Y_{discrim} and the IPLS $_{\text{grid}}$ cases.

The specific choices of illustrative food article systems are not in any way an absolute indication of the general applicability of this approach. What have been shown feasible for perishable fruit articles, and for on-line food product characterisation, is of course equally applicable to, say, cereals, bread, meat, fish—indeed the food and feed areas at large are potentially opened up for a similar approach, e.g. the berries—, beverages,—dairy sectors, etc.

Continuing outside the human, and animal, food and feed areas, an *analogous* automated image monitoring approach can of course equally well be envisaged for quite different application areas, at first primarily within the general technological and industrial sectors—but perhaps even further removed. One common denominator could be any *multi-temporal* aspect, which would lend itself to an appropriately modified image recording and—analysis approach, *similar* to the one *illustrated* here, e.g. degradation studies: paints, coatings, corrosion inhabitation—industrial inspection in general i.a. Within the field of *remote sensing*, there is also a plethora of similar multi-temporal objectives.

The on-line image monitoring example, while relatively simple in the crispbread case, also has many, much broader application potentials within much of the *industrial inspection* realm, in which there is often a distinct need for automated image analytical monitoring.

For the present feasibility studies, we are satisfied with the above results for both the *IPLS- Y_{discrim}* and

³ We have elsewhere worked out a complete image analysis strategy, which—while originally presented as related to MIA—also applies to the *analogous* *t–t* score plots derived by an IPLS-solution [17]. As but an example we there followed another biological process, albeit of considerable greater complexity, i.e. a forest clearing *regrowth* process, using geomorphological analogies in order to characterise MIA score plots. From this review [17], a range of interpretation guidelines for *t–t* score plots were developed, all of which may also be applied to the present PLS-solutions. Observe that a slightly different *modus operandi* applies to the *t–u* plots [18]. These subtle differences will be addressed in several sequel papers on a comprehensive *MIR strategy*, which are in the works.

the $IPLS-Y_{grid}$ approaches. We have shown that the multivariate image regression approach (MIR) is now fully established. It bears in mind though, that there is always a series of critically important *specific* associated image-analytical problems, e.g. problem-specific illumination, shadows, reflections, non-constant object sizes a.o.—much interesting work remains.

The present first foray into the possibilities of multivariate image regression has focussed on the ways-and-means of *modelling* (using bilinear IPLS) and *prediction*. What remains is the equally important aspect of multivariate calibration, *validation* (in the form of image-regression validation), which forms the subject-matter of the second paper in this series, in which we also will make use of the third IPLS-regression case only identified here: Y_{total} [19].

References

- [1] K. Esbensen, P. Geladi, Chemom. Intell. Lab. Syst. 7 (1989) 67–86.
- [2] P. Geladi, H. Grahn, Multivariate Image Analysis, Wiley, Chichester, UK, 1996, p. 316.
- [3] F. Lindgren, P. Geladi, S. Wold, J. Chemom. 7 (1993) 45–59.
- [4] K. Esbensen, P. Geladi, H. Grahn, Chemom. Intell. Lab. Syst. 14 (1992) 67–86.
- [5] T.T. Lied, P. Geladi, K. Esbensen, J. Chemom. 14 (2000) 585–598.
- [6] N. Lamei, K.D. Hutchison, M.M. Crawford, N. Khazenie, Opt. Eng. 33 (1994) 1303–1313.
- [7] T. Yamazaki, D. Gingras, IEEE Trans. Image Process. 4 (1995) 1333–1339.
- [8] J.R. Carr, Comput. Geosci. 22 (1996) 849–865.
- [9] H. Martens, T. Naes, Multivariate Calibration, Wiley, Chichester, UK, 1989, p. 419.
- [10] K. Esbensen, S. Wold, P. Geladi, J. Chemom. 3 (1988) 33–48.
- [11] A.K. Smilde, J.A. Westerhuis, R. Boque, J. Chemom. 14 (2000) 301–331.
- [12] R. Bro, Doctoral thesis, Multiway analysis in the food industry, 1998.
- [13] C. Anderson, R. Bro (Eds.), J. Chemom. vol. 14, 2000, pp. 103–334.
- [14] A.K. Smilde, H.A.L. Kiers, J. Chemom. 13 (1999) 31–48.
- [15] P. Geladi, K. Esbensen, J. Chemom. 5 (1991) 97–111.
- [16] J.P. Wold, K. Kvaal, Appl. Spectrosc. 54 (2000) 900–909.
- [17] K. Esbensen, T.T. Lied, K. Lowell, G. Edwards, Principles of Multivariate Image Analysis (MIA) in remote sensing, technology and industry, submitted for publication.
- [18] K. Esbensen, Multivariate Analysis in Practice, 5th edition, CAMO ASA, Oslo, Norway, 2001.
- [19] T.T. Lied, K. Esbensen, Principles of MIR, Multivariate Image Regression: II, Cross validation—what you see is what you get, submitted for publication.