TECHNICAL NOTE TN2022_5 – INSIGHT PRE-PROCESSORS

Introduction

Specim offers Insight and IQ studio to process hyperspectral data. In those software, pre-processors are included in order to improve the predictive performance of the created models. The aim of this Technical Note is to go through those pre-processors, explaining their effect on the spectra.

SG = SAVITZKY – GOLAY

SNV = STANDARD NORMAL VARIATE

MROI = MULTIPLE REGIONS OF INTEREST

Article

Pre-processors are widely used in spectroscopy when one wants to improve the predictive performances of models. It helps making models more robust and accurate by highlighting relevant spectral features in spectra. Specim Insight (*) and IQ Studio (¤) includes the following pre-processors:

- Savitzky – Golay filtering (* ¤)
- Derivative (*)
- Max-norm (*)
- P-norm (*)
- Mean-centering (*)
- Standard Normal Variate (*)
- Pseudo-absorbance (*)
- minmax-Norm (*)

Pre-processors indicated with a * are included in Specim Insight, whereas those indicated with a ¤ are included in IQ Studio.

- **Savitzky - Golay (SG) filtering**

  This is probably one of the most widely used pre-processor in spectroscopy**. The SG filtering reduces the noise in the data**, by fitting a polynomial curve along with the spectra. This fitting is estimated over successive subsets of the data (i.e. successive spectral windows with a fixed width). The user can input 2 parameters for the curve fitting: the degrees of the polynomial function, and the filter width.

  Notice that a SG filtering of the 0 order corresponds to a running average.

  In Fig.1, we can observe that:

  - A large filter width reduces significantly the noise, but may smooth too much spectral features. If sharp spectral features are expected, a large smoothing window would tend to filter those out.
  - A polynomial function with a too high order may tend to build overcomplicated smoothers, whereas a too low order would create artifacts in the spectra (like steps).

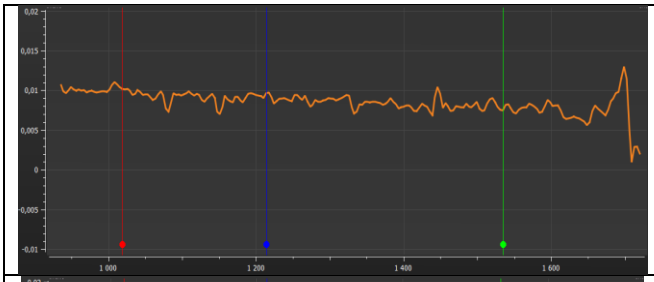
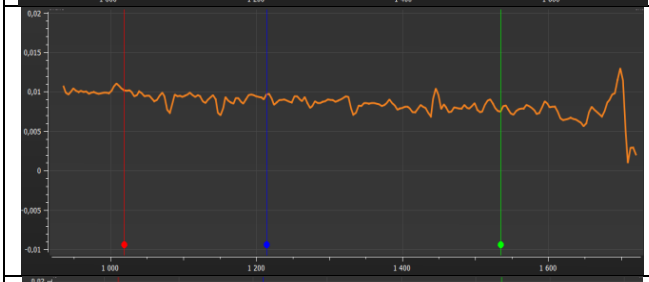
	Original spectra
	Pre-processed spectrum Spectral windows: 5 bands Polynomial order: 2
	Pre-processed spectrum Spectral windows: 10 bands Polynomial order: 2
	Pre-processed spectrum Spectral windows: 10 bands Polynomial order: 1
	Pre-processed spectrum Spectral windows: 10 bands Polynomial order: 0 (= running average)

Figure 1: Effect of SG filtering on spectra with different fitting parameters.

Notice that a derivative computation is very often bound with a SG filtering. For more information about derivative as pre-processor, please consult the following section “Derivative”.

- **Derivative**

  Derivating a continuous function (like a continuous spectrum) would transform it and highlight several aspects of its shape:

  - It will **remove any intensity offset** (the derivative of a constant is zero). In the situation where 2 samples are made of the same material, but only their reflective intensity is different, after a derivative, both samples would look much more similar, and the intensity difference would not play any role anymore in the modelling (Fig.2). Also, if an original spectrum is biased not only by an offset, but by something similar to a first order polynomial function (like a slope as a baseline), the second derivative would eliminate it.

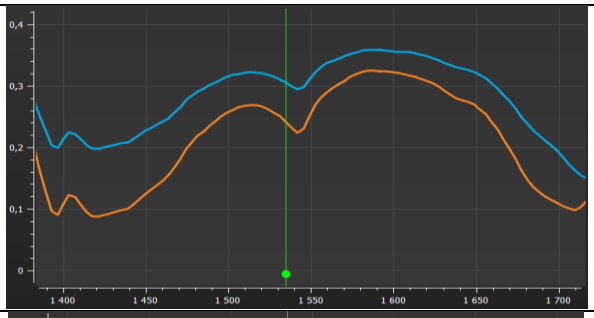
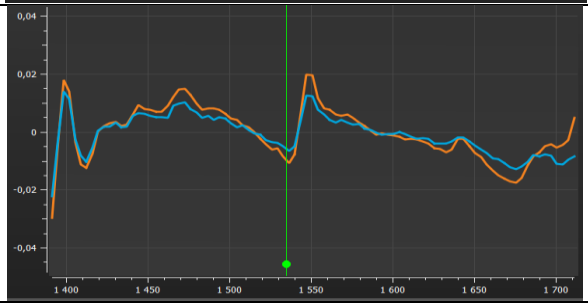
	Original spectra from 2 samples with the same material, but with a reflective offset as difference.
	Processed spectra after 1st order derivative

Figure 2: effect of the first derivative on intensity offset between spectra.

  - A first order derivative will quantify the slope of a spectral curve at each band. **It means that when the derivative is equal to 0, the spectrum can have a local maximum or minimum**. Therefore the derivative helps in identifying their position. It will also highlight if spectra are increasing or decreasing faster than others (see Fig.3).
  - A second derivative is also widely used. **When it equals to 0, it highlights inflection points in the original spectra**. Notice that when the first derivative is equal to 0 and the second derivative is positive, it means that original spectra contained at that band an absorption peak, whereas on the opposite, it would contain a reflection peak (see Fig.3).

  It is important to notice that the derivative of a noisy spectrum is also noisy. It is always recommended to use the derivative function after a smoother, like a SG for instance (hence they often go in pair).

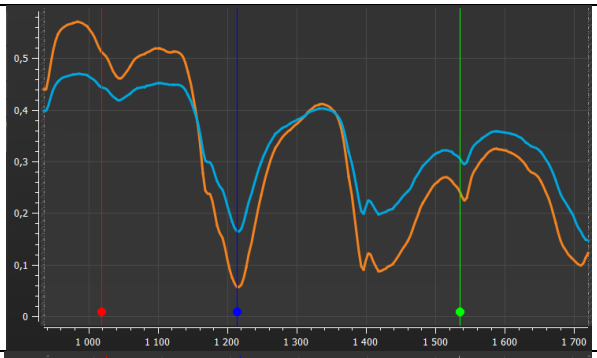
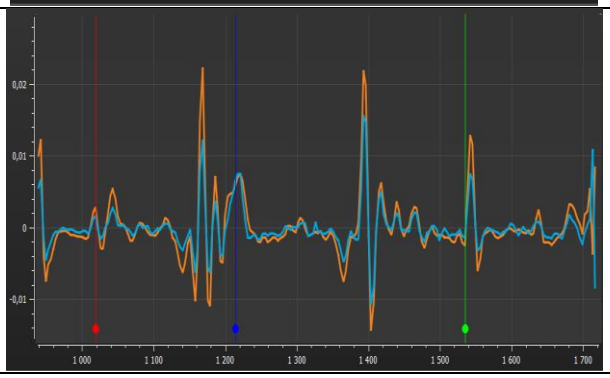
	Original spectra (both from the same material)
	Pre-processed after SG and derivative: Spectral windows: 5 bands Polynomial order: 2 **Derivative order: 1**
	Pre-processed after SG and derivative: Spectral windows: 5 bands Polynomial order: 2 **Derivative order: 2**

Figure 3: effect of the first and second derivatives on spectra.

Notice that the derivative is always computed locally, taking into account for each band the next consecutive one only. For instance, the derivative at the band λ is computed only by looking at the spectral values at bands λ and λ+1. Notice that even in the case of a spectral gap between (due to MROI for instance), the derivative will still be computed).

- **Max-norm (or Maximum normalization)**

  Each spectrum, for each individual pixel is normalized by its maximum value. In practice, it means that for each spectrum the maximum value is found and is used as a divider for each value for each band.

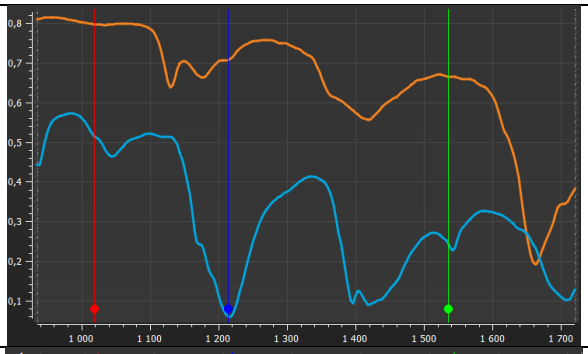
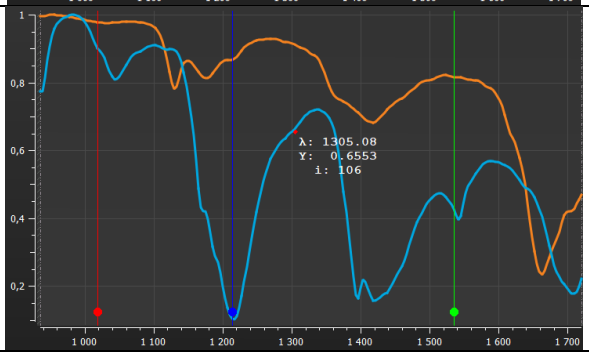
	Original spectra of 2 different materials
	Spectral after maximum normalization

Figure 4: effect of the maximum normalization on spectra

A maximum normalization would scale the spectrum of each pixel to a somewhat comparable level and would enhance comparison for in turn building a more accurate model.

- **p-norm**

  In statistics, a p-norm is a statistical distance and **it helps in comparing vectors (here spectra) between each other**. It is defined as in equation 1:

$$\|spectrum\|_p = \left(\sum_{i=1}^{n} |\lambda_i|^p\right)^{\frac{1}{p}} \quad (1)$$

  - p is the power of the norm
    - if p = 1, the p-norm is also called the “taxicab” norm.
    - If p = 2, the p-norm is also called the “Euclidian” norm.
    - If p tends to ∞, the p-norm becomes the max norm.
  - λ_i is the spectral value at the band “i”. Notice that only bands selected for the models are taken into the equation. See for example:
    - in a case where the full spectrum of a FX17 is needed, without binning, i will range from 1 to 224
    - in a case where the full spectrum is needed of a FX17, with a x2 spectral binning, i will range from 1 to 112
    - in a case where 2 spectral ROIs are included for modelling (from bands e.g. 12 to 42 and 100 to 112), i will range from 12 to 42 and from 100 to 112 only.

  In a p-norm transform, each band of the original spectrum is divided by the p-norm value of the spectrum. In Insight, the user can specify the order of the norm.

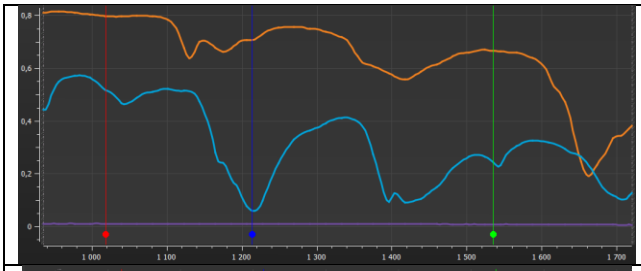
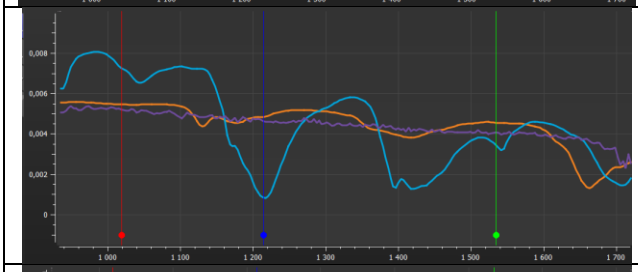
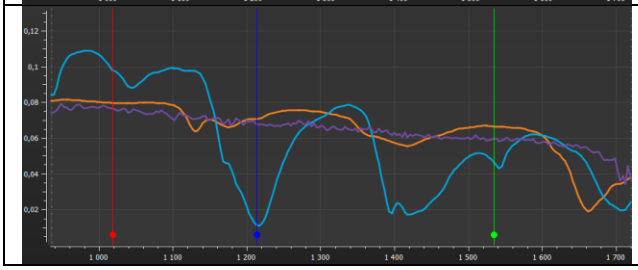
	Original spectra of 3 different materials
	Spectra after 1-norm transform
	Spectra after 2-norm transform

Figure 5: effect of the p-norm transform on spectra

The p-norm enhances the spectral dynamic on low amplitude spectra, whereas it reduces the dynamic on spectra which have a large one. **It is scaling the spectra between each other** (Fig.5). Notice that the p-norm should be used on denoised or smoothed data.

- **Mean-centering**

  A Mean-centering transformation would apply a corrective offset for each spectrum related to each pixel so that its average is zero.

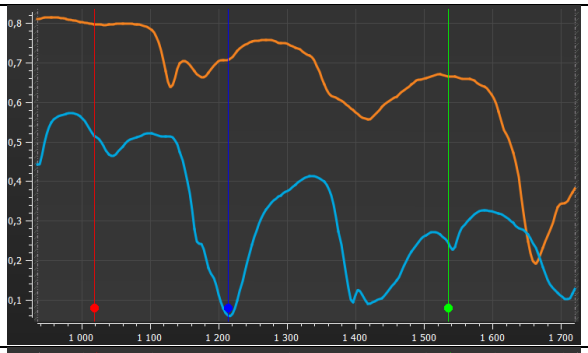
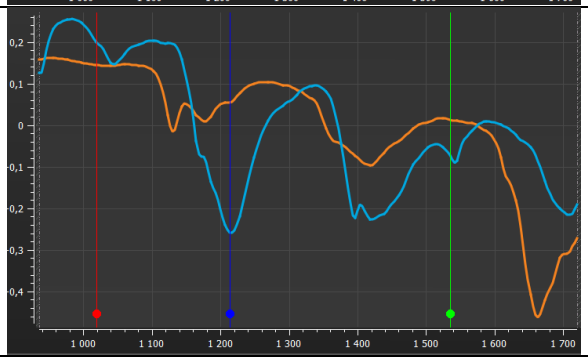
	Original spectra of 2 different materials
	Spectral after Mean centering transform

Figure 6: effect of the Mean-centering on spectra

Notice that in Insight, there are 2 Mean-centering transform available:

- Mean-centering, which is described as previously, per pixel
- Mean-centering (whole training set): in that case, only pixels selected in training selections are taken into account, and a global averaged spectrum is computed. This spectrum is then subtracted over each individual spectrum of the full image.

  In this latest variant of the Mean-centering, the size of the training selection per class strongly impacts the results (Fig.7). When the training selection between classes are very unbalanced, the one dominating the others would tend to become the reference, and relative spectral differences between the classes and this reference will be enhanced.

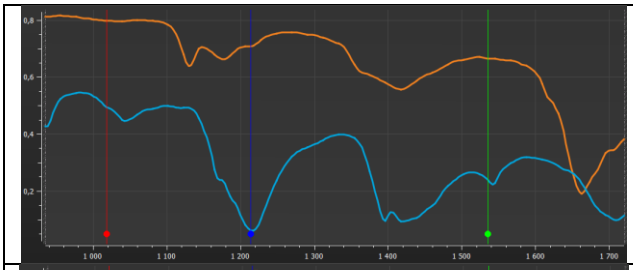
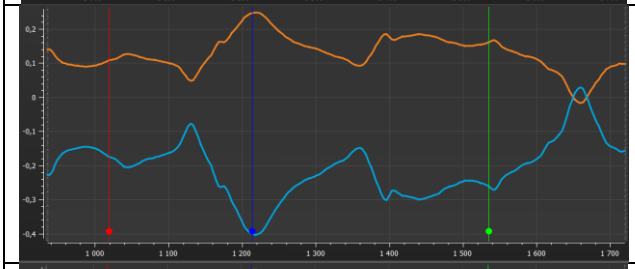
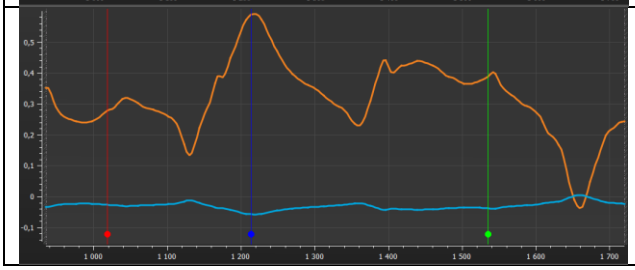
	Original spectra of 2 different materials
	Spectral after Mean-centering (Whole training set) transform, with similar amount of pixels selected to train the “blue” and “orange” classes.
	Spectral after Mean centering (Whole training set) transform, with a 7.5 time more pixels selected for the “blue” class than for the “orange” one.

Figure 7: effect of the training selections when applying the Mean centering (whole training set) on spectra

- **Standard Normal Variate (SNV)**

  In a SNV transform, the spectra are processed so that their mean is centered to 0 (like Mean centering would do) and is then divided by its standard deviation. It means that:

  - It will enhance the dynamic of spectra when there is originally a little one
  - Those spectra with already a high dynamic will be somewhat unchanged (except of their Mean centering).

  In practice, it improves the transfer of models between cameras of the same kind (with for instance sensitivity variations), and will increase the modelling performance on the edges of samples (Fig. 8).

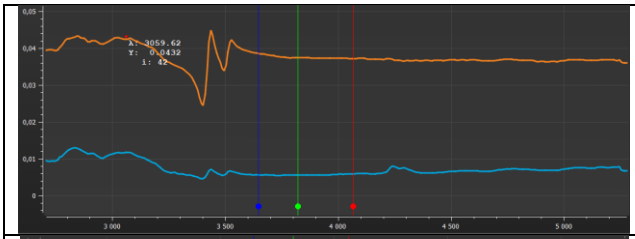
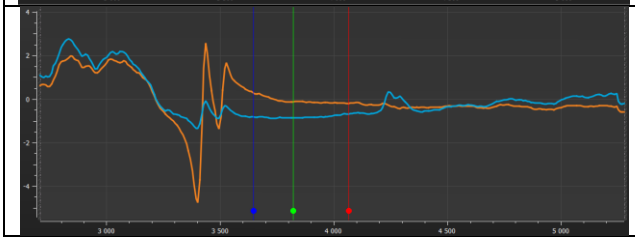
	Original spectra (both from the same material) Orange: high dynamic Blue: low dynamic
	Spectra after SNV transform

Figure 8: effect of the SNV transform on spectra

Besides, The SNV transform will also improve the performances of models when sorting samples with low reflection from some with high one. In Fig.8, the orange spectrum corresponds to a sample made of black PE with a diffuse surface, whereas the blue one refers to a similar sample but with a specular surface. Applying a SNV transform significantly improves the accuracy of the model.

- **Pseudo-absorbance**

  There are applications the reflectance is not the preferred metrics to analysis spectra, but absorbance. Absorbance can be deduced from reflectance data, by applying the formula below (Eq.2). It results in estimating what absorbance would be, as pseudo-absorbance. Assumption needs to be made that the sample is opaque (no transmission).

$$Pseudo - absorbance = -log_{10}(Reflectance) \quad (2)$$

  It is a rather simple equation, based on the logarithmic function (see Fig.9).

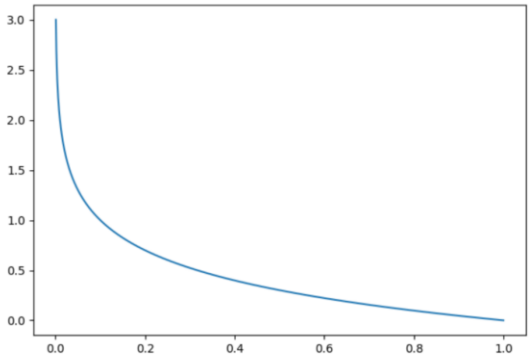


Figure 9: Pseudo-reflectance function. X-axis: reflectance between 0 and 1; Y-axis: pseudo reflectance

  Notice that:
  - For reflectance values > 1, the Pseudo-absorbance is negative.
  - For reflectance ≤ 0, the Pseudo-absorbance can not be computed (those become Not a Number, i.e. NaN). The corresponding pixels will be sorted as “unclassified”.
    Because of this, the pseudo-reflectance should not be used after all pre-processors. For instance, a derivative (including SG with derivative) or a mean-centering produce negative spectral values. Using a pseudo-absorbance after those would lead to poor modelling.

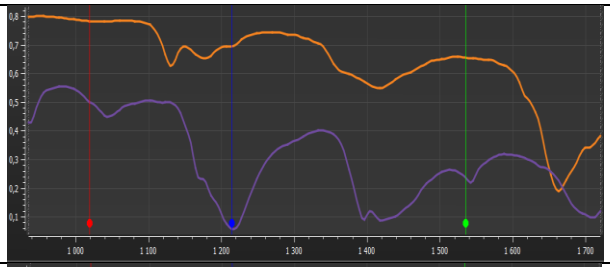
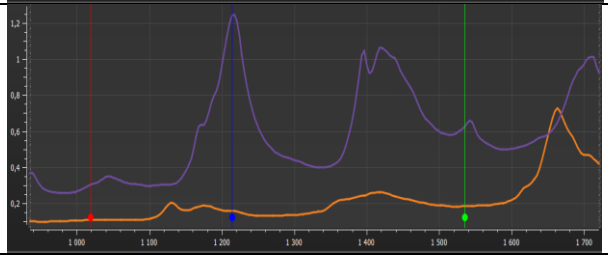
	Original spectra of 2 different materials
	Spectra after pseudo-absorbance transform

Figure 10: effect of the p-norm transform on spectra

Notice that the pseudo-absorbance transform is based on a non-linear equation, and the spectra may look distorted (meaning it is not 1:1 mirrored with the reflectance). This can be seen in Fig.10. **This non-linearity will enhance deep reflectance absorption peaks in respect to those with small depth. This may improve the dynamic of some absorption peaks, which in turn would improve the accuracy of quantitative models (hence its wide used in pharmaceutics)**.

- **minmax-Norm (i.e. Min-Max normalization or Min-Max scalling)**

  The minmax-Norm preprocessor will scale the spectrum of each pixel from 0 to 1.

  - It will enhance the dynamic of spectra when there is originally a little one
  - This will enhance spectral features

  In practice, it might make models more robust to shadow effect. Also, the identification between more or less diffuse samples of the same kind will be more reliable (Fig. 9).

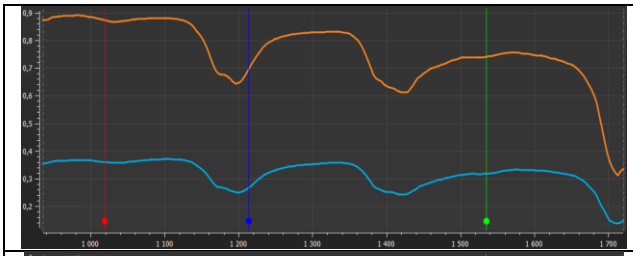
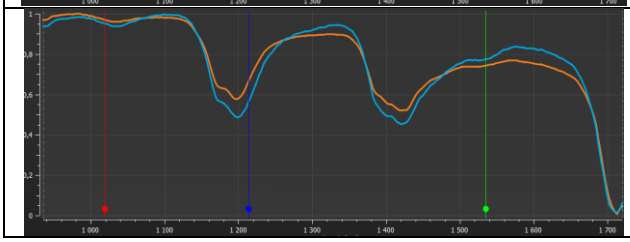
	Original spectra (both from the same material) Orange: well exposed area Blue: with shadow
	Spectra after minmax-Norm transform

Figure 8: effect of the minmax-Norm transform on spectra

Notice that this preprocessor may enhance the noise or artifacts in the data. We recommend to use it after a smoother, like Savitzky – Golay filtering.

Finally, we would like to highlight to the users that some of the pre-processors may generate NaN data, which are sorted as “unclassified” by the software Insight. Pseudo-absorbance and Max-norm can generate such data and they should be interpreted with care.

In Insight, several pre-processors can be used, and their order can also be adjusted. We recommend to test different combinations in order to make models most robust and accurate.

Disclaimer

Version history

Version	Date	Author	Comments
1.0	Sept. 28th 2022	MMA, HHA, JJA, RBA	The pre-processors mentioned in this TN version are included in Insight 1.1.0.34. All the mentioned pre-processors are not necessarily all present in previous version of the software.
2.0	Oct. 28th 2022	MMA, JJA	The pre-processors mentioned in this TN version are included in Insight 1.2.1.47. All the mentioned pre-processors are not necessarily all present in previous version of the software.
3.0	June 12th 2023	MMA	The pre-processors mentioned in this TN version are included in Insight 1.3.0.63. All the mentioned pre-processors are not necessarily all present in previous version of the software.