Strategies on the Configuration of Hyperspectral Preprocessing

> *Hyperspectral preprocessing is typically used to reduce disorder of the measurement setup (varying lighting, etc.) and, in general, aims at the transformation of spectroscopic data into a form, most suitable for further processing like information extraction.*
>
> *Therefore, information caused by the chemistry of objects is aimed to be more dominant compared to influencing disorders. This page summarizes strategies on how to find a proper preprocessing for an application.*

Note: preprocessing can generate noise which can get dominant too, and can potentially influence the preprocessed signal, and therefore can worsen the results. In general it is good practice to decide for a preprocessing mode that doesn't generate too much noise.

Comparison between Preview and Expectation

The Preview feature tends to show points of similar spectral information in similar color, while points of differing spectroscopic information are shown in distinct colors.

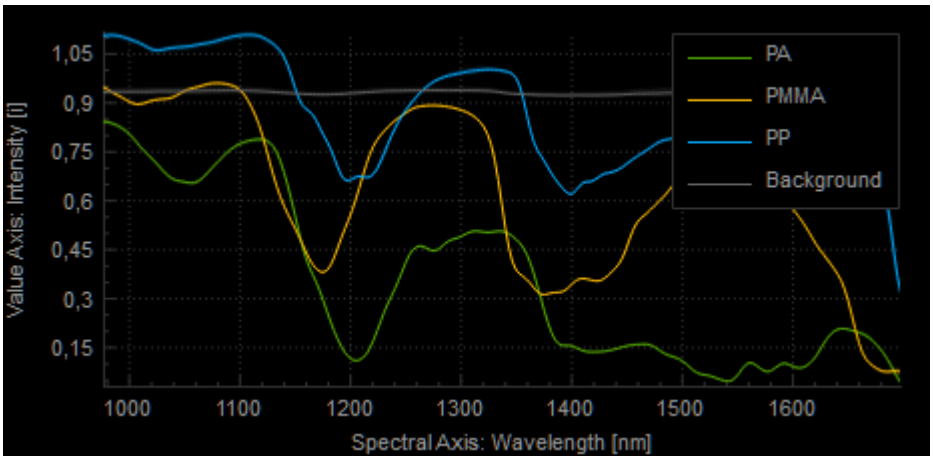
Keep in mind that

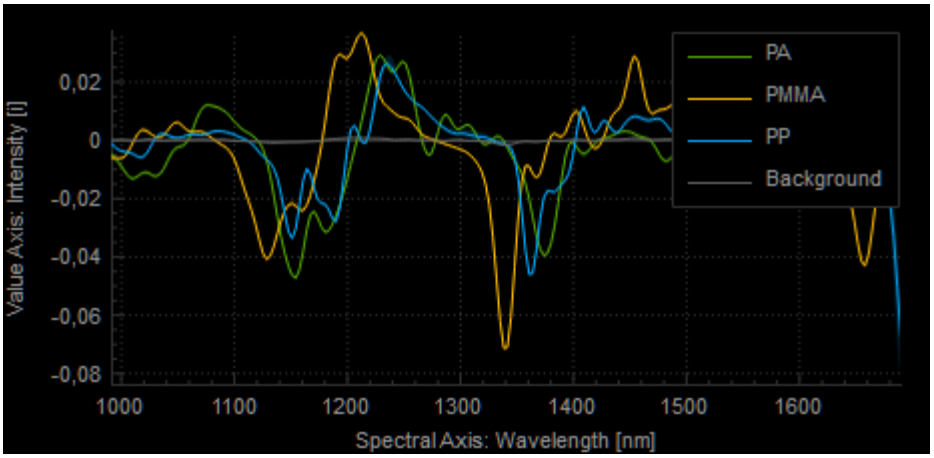
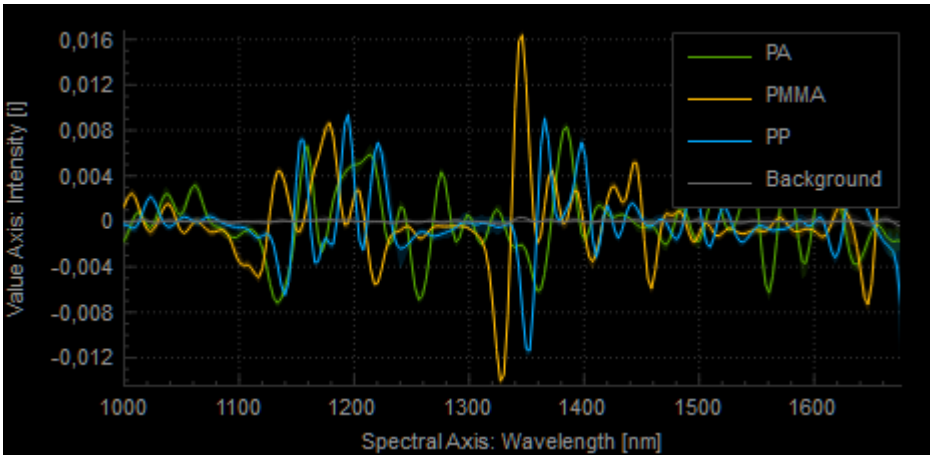
- the Preview leads to ***distinct*** color information for ***distinct*** spectroscopic information and
- proper preprocessing tends to ***reveal*** spectroscopic information caused ***by chemistry***

This allows us to ***identify*** candidates for potentially suitable preprocessing methods by ***matching*** our expectation of chemistry with gained color images.

In the example below three plastic plates of different chemistry (PA, PMMA, PP) are shown at different preprocessing modes. A color image illustrates the Preview for each preprocessing mode. Next to it, selected spectra from the plates and background are shown.
We assume the three plastic plates to be distinct in chemistry (PA, PMMA, PP) - therefore we look for the preprocessing mode which best describes the plates by distinct color information.

In the following table the influence of the preprocessing method "derivative" is summarized:

Preproc.	Preview	Spectra
Reflectance		

1st Derivative		
2nd Derivative		

The left part of the table above shows reflectance data and data obtained by applying derivation to the data. Please note: By applying the derivation to spectra, constant factors like offset is rejected. Higher order derivation tends to have noisy spectra.

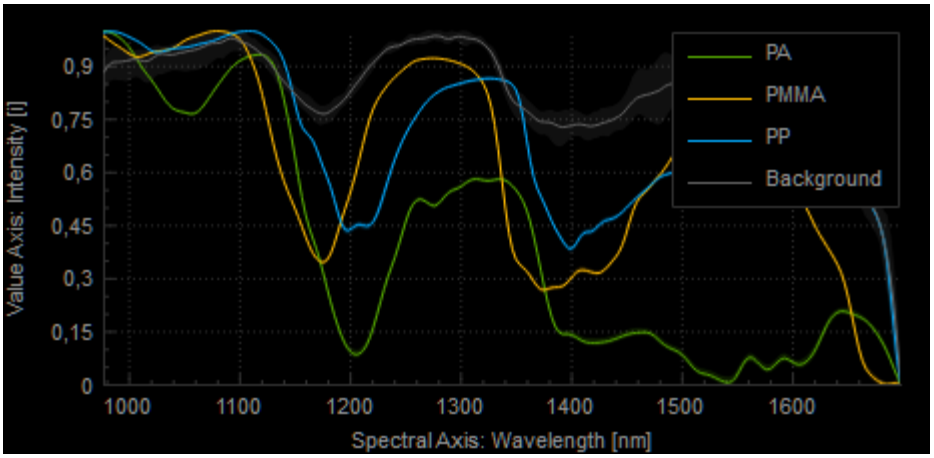
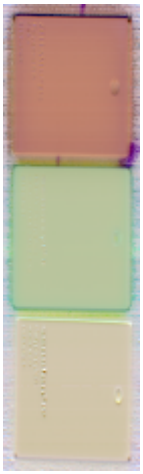
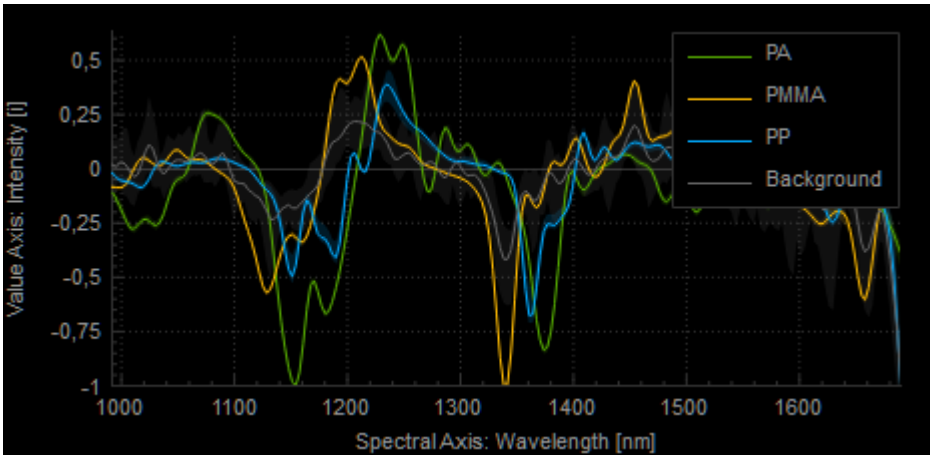
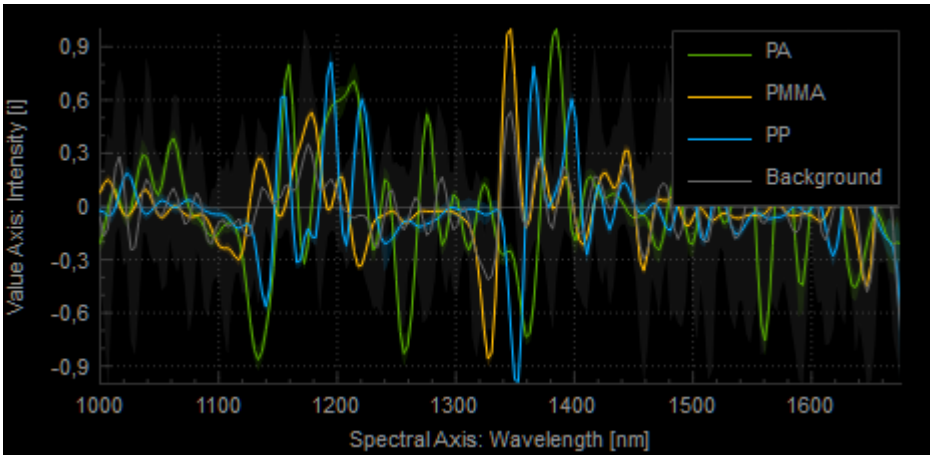
- Investigating reflectance data: The color impression reaches from bluish to bluish-gray. Therefore the plates are shown in different color but not with distinct color.
- Investigating data of 1st derivative: the upper plate is distinct (brownish) to the other plates.
- Investigating data of 2nd derivative: all plates are distinct in color. The image data shows some noise.

The 2nd derivation differentiates the chemistry best by different colors, but also shows some noise in the images.

- When good image quality is needed, the preprocessing "1st derivation" might be the best choice for further processing,
- when some degradation of image quality is tolerable, the preprocessing "2nd derivation" might be the best choice since the Preview shows the best connection to our expectations (plates are chemically different).

In the following table the influence of the preprocessing method "normalization" and "derivative" is summarized:

Preproc.	**Preview**	**Spectra**

Reflectance normalized		
1st Derivative normalized		
2nd Derivative normalized		

The table above summarizes the same data as before but additionally normalization was applied.
Please note: By applying normalization to the data the absolute values get lost - each spectra is in the same value range. Background pixels in the scene tend to cause noisy output information because normalization is done without respect of object or background.

The color impression gained is more or less comparable to those of the data without normalization.
Except:

- Object edges can get "blurred" and background pixels can tend to show similar spectroscopic information as the objects next to them. These effects are typically caused by disorders of the optical system (lens, spectrograph, etc.). Dependent on the optical system used, this disorder can be large or negligible small in value.
- Variations of the absolute value get lost - therefore, the print on the plastic plates gets more or less lost when normalization is applied. The color tone of the chemically absolutely homogeneous plates is more regular when normalization is applied.

- Small spectral influences get visible. At the lower right corner of the upper plate of preprocessing "1st Derivative normalized" a small reddish area gets visible. Later investigation has shown this information to correlate to a water film on the measurement background. It is remarkable that this information is "hidden" for us when no normalization is applied - therefore the absolute value is quite small compared to the plastic plates. When normalization is applied, spectra are transformed to be in the same value range - therefore also the tiny spectral influence of the water drop gets accessible.

When blurring effects aren't critical for the application of interest, the preprocessing "1st derivation normalized" might be suitable since the Preview corresponds to our expectations (different colors). Please note: by applying normalization the spectra are much more decoupled to external influences like light variations - so the spectra are in general much more "stable" compared to those without normalization. In case relevant information gets lost by normalization, which is often the case, this preprocessing will not be the best choice.

Preprocessing configuration for some applications

Polymer identification (e.g. sorting)	1st derivative
Food identification	Intensity, 1st derivative
Impurity detection (in e.g. food)	1st derviative
Medical tissue investigation	2nd derivative
Minerals identification	Intensity, 1st derivative, 2nd derivative
Wood inspection	Intensity, 1st derivative
Pharmaceutical investigation	Intensity, 1st derivative

*) The combination with normalization can be beneficial.