

NIR Spectroscopy Glossary

NIR Fundamentals Course

Instructor: Vitaly Kirkpatrick

A comprehensive glossary of Near-Infrared Spectroscopy terms and concepts covered in this course.

A

Absorbance (A): The logarithmic measure of light absorbed by a sample, calculated as $A = -\log_{10}(T)$, where T is transmittance. Directly proportional to concentration according to Beer-Lambert Law.

Absorptivity (a): A measure of how strongly a chemical species absorbs light at a given wavelength. Also called absorption coefficient.

Accuracy: The degree of closeness of measurements to the true or accepted reference value.

Algorithm: A step-by-step computational procedure used in chemometrics for calibration, preprocessing, or data analysis.

B

Baseline Correction: A preprocessing technique that removes systematic variations in the baseline of a spectrum, often caused by instrument drift or sample presentation.

Beer-Lambert Law: A fundamental principle stating that absorbance is directly proportional to both concentration and path length: $A = \epsilon bc$, where ϵ is molar absorptivity, b is path length, and c is concentration.

Bias: Systematic error in predictions, calculated as the average difference between predicted and reference values.

C

Calibration: The process of developing a mathematical model that relates NIR spectral data to a property of interest (e.g., moisture content, protein level).

Calibration Set: A collection of samples with known reference values used to build a predictive model.

Chemometrics: The application of mathematical and statistical methods to chemical data, essential for NIR spectroscopy analysis.

Combination Band: An NIR absorption band resulting from the simultaneous excitation of two or more fundamental vibrations.

Cross-Validation: A validation technique where the calibration set is divided into subsets, with each subset used for validation while the others are used for calibration.

D

Derivative Spectroscopy: A preprocessing technique that calculates the first or second derivative of a spectrum to enhance spectral features and reduce baseline effects.

Detector: A device that converts light energy into an electrical signal. Common NIR detectors include InGaAs, PbS, and Silicon.

Diffuse Reflectance: Light that is reflected from a sample in all directions due to scattering within the sample matrix.

Dispersive Instrument: A spectrometer that uses a diffraction grating or prism to separate light into its component wavelengths.

E

Electromagnetic Spectrum: The range of all possible frequencies of electromagnetic radiation, including visible, NIR, and infrared regions.

External Validation: Testing a calibration model using an independent set of samples not used during model development.

F

Fiber Optics: Flexible glass or plastic fibers that transmit light, used to deliver NIR light to samples and collect reflected or transmitted light.

FT-NIR: Fourier Transform Near-Infrared spectroscopy, which uses an interferometer to collect spectral data simultaneously across all wavelengths.

Functional Group: A specific group of atoms within a molecule that determines its chemical properties and NIR absorption characteristics (e.g., O-H, N-H, C-H).

Fundamental Vibration: The lowest energy vibrational transition of a molecule, typically occurring in the mid-infrared region (2500-25000 nm).

G

Grating: An optical component with a series of parallel grooves that disperses light into its component wavelengths through diffraction.

H

Hyperspectral Imaging: A technique that combines NIR spectroscopy with spatial imaging to create a data cube containing both spectral and spatial information.

I

InGaAs Detector: Indium Gallium Arsenide detector, commonly used for NIR detection in the 900-1700 nm range.

Interferometer: An optical device used in FT-NIR instruments that creates an interference pattern by splitting and recombining a light beam.

L

Latent Variable: A hidden or unobserved variable in multivariate analysis that captures patterns in the data. PLS regression uses latent variables to relate spectra to properties.

Loading: In PCA or PLS, the contribution of each wavelength to a principal component or latent variable.

M

Molar Absorptivity (ϵ): A measure of how strongly a chemical species absorbs light at a given wavelength per unit concentration and path length. Units: $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$.

Molecular Vibration: The periodic motion of atoms within a molecule, which absorbs specific wavelengths of infrared light.

MSC (Multiplicative Scatter Correction): A preprocessing technique that corrects for light scattering effects in diffuse reflectance spectra.

N

Nanometer (nm): A unit of wavelength equal to 10^{-9} meters. NIR region spans approximately 800-2500 nm.

Near-Infrared (NIR): The region of the electromagnetic spectrum between visible light and mid-infrared, typically defined as 800-2500 nm.

Noise: Random variations in spectral data caused by detector electronics, environmental factors, or sample heterogeneity.

Normalization: A preprocessing technique that scales spectra to a common reference, often by dividing by the total area or maximum value.

O

Outlier: A sample whose spectral or reference value is significantly different from the rest of the dataset, potentially indicating measurement error or unusual sample properties.

Overtone: An NIR absorption band corresponding to a multiple of a fundamental vibrational frequency (e.g., first overtone = $2 \times$ fundamental frequency).

P

Path Length: The distance light travels through a sample, typically measured in millimeters or centimeters.

PbS Detector: Lead Sulfide detector, used for extended NIR range detection (1000-2500 nm), often requires cooling.

PCA (Principal Component Analysis): A dimensionality reduction technique that transforms correlated variables into uncorrelated principal components.

PLS (Partial Least Squares): A regression method that relates spectral data to a property of interest by finding latent variables that maximize covariance between spectra and reference values.

Precision: The degree of reproducibility of measurements, often expressed as standard deviation or coefficient of variation.

Preprocessing: Mathematical transformations applied to spectra before calibration to improve model performance (e.g., derivatives, SNV, MSC).

Principal Component: An orthogonal axis in multivariate space that captures maximum variance in the data.

Q

Qualitative Analysis: Identification or classification of samples based on their spectral characteristics (e.g., authentication, variety identification).

Quantitative Analysis: Determination of the concentration or amount of a constituent in a sample using a calibration model.

R

R² (Coefficient of Determination): A statistical measure of how well a calibration model fits the data, ranging from 0 to 1 (1 = perfect fit).

Reflectance: The fraction of incident light reflected by a sample, often expressed as a percentage.

Reference Method: A standard analytical technique (e.g., wet chemistry, chromatography) used to determine true values for calibration samples.

Regression: A statistical method for modeling the relationship between predictor variables (spectra) and a response variable (property of interest).

Residual: The difference between a predicted value and the reference value for a sample.

RMSEC (Root Mean Square Error of Calibration): A measure of calibration model fit, calculated as the square root of the average squared residuals for calibration samples.

RMSECV (Root Mean Square Error of Cross-Validation): A measure of model prediction error estimated through cross-validation.

RMSEP (Root Mean Square Error of Prediction): A measure of model prediction error for an independent validation set.

RPD (Ratio of Performance to Deviation): A metric for model performance, calculated as the ratio of standard deviation of reference values to RMSEP. RPD > 3 indicates good predictive ability.

S

Sampling: The process of selecting and presenting samples for NIR measurement.

Scatter: The redirection of light in multiple directions due to interaction with particles or inhomogeneities in a sample.

Score: In PCA or PLS, the projection of a sample onto a principal component or latent variable.

SEP (Standard Error of Prediction): A measure of prediction error for an independent validation set, similar to RMSEP.

Signal-to-Noise Ratio (SNR): The ratio of the desired signal to background noise, a measure of data quality.

Silicon Detector: A detector sensitive to visible and short-wave NIR light (400-1100 nm).

Smoothing: A preprocessing technique that reduces high-frequency noise in spectra, often using moving average or Savitzky-Golay filters.

SNV (Standard Normal Variate): A preprocessing technique that centers and scales each spectrum to zero mean and unit variance, correcting for scatter effects.

Spectral Range: The wavelength interval covered by a spectrometer (e.g., 400-2500 nm).

Spectral Resolution: The smallest wavelength difference that can be distinguished by a spectrometer, typically expressed in nanometers.

Spectrometer: An instrument that measures the intensity of light as a function of wavelength.

Spectrum (plural: Spectra): A plot of light intensity or absorbance versus wavelength.

T

Transmittance (T): The fraction of incident light that passes through a sample, expressed as $T = I/I_0$, where I is transmitted intensity and I_0 is incident intensity.

Transflectance: A measurement mode where light passes through a sample, reflects off a backing material, and passes through the sample again before reaching the detector.

V

Validation: The process of testing a calibration model using samples not used during model development to assess predictive performance.

Validation Set: An independent set of samples used to test model performance.

Variance: A statistical measure of data spread, calculated as the average squared deviation from the mean.

Vibrational Spectroscopy: Spectroscopic techniques (including NIR, mid-IR, and Raman) that measure molecular vibrations.

W

Wavenumber: The reciprocal of wavelength, expressed in cm^{-1} . Common unit in mid-infrared spectroscopy, less common in NIR.

Wavelength: The distance between successive peaks of an electromagnetic wave, typically measured in nanometers (nm) for NIR.

Acronyms & Abbreviations

- **ANN:** Artificial Neural Network
- **CNIRS:** Council for Near-Infrared Spectroscopy
- **CV:** Cross-Validation
- **FT-NIR:** Fourier Transform Near-Infrared
- **InGaAs:** Indium Gallium Arsenide (detector material)
- **IR:** Infrared
- **MLR:** Multiple Linear Regression
- **MSC:** Multiplicative Scatter Correction
- **NIR:** Near-Infrared
- **NIRS:** Near-Infrared Spectroscopy
- **PbS:** Lead Sulfide (detector material)
- **PCA:** Principal Component Analysis
- **PCR:** Principal Component Regression
- **PLS:** Partial Least Squares
- **R²:** Coefficient of Determination
- **RMSEC:** Root Mean Square Error of Calibration
- **RMSECV:** Root Mean Square Error of Cross-Validation
- **RMSEP:** Root Mean Square Error of Prediction
- **RPD:** Ratio of Performance to Deviation
- **SEP:** Standard Error of Prediction
- **Si:** Silicon (detector material)
- **SNR:** Signal-to-Noise Ratio
- **SNV:** Standard Normal Variate

- **SWIR:** Short-Wave Infrared
-

Common Units

- **nm (nanometer):** 10^{-9} meters, standard wavelength unit in NIR
 - **cm^{-1} (wavenumber):** Reciprocal centimeters, common in mid-IR
 - **% (percent):** Common unit for moisture, protein, fat content
 - **g/L, mg/L:** Concentration units
 - **mm, cm:** Path length units
-

Last Updated: January 19, 2026

Course: NIR Spectroscopy Fundamentals

Instructor: Vitaly Kirkpatrick

For additional resources, see the Bibliography and Quick Reference Guide.