The Australian Wine Research Institute

# Development and Validation of Secondary Methods for the Measurement of Analytical Parameters in Grapes and Must

## Version Control

| Version | Changes made in this version |
|---|---|
| 1.0 | First version of this method |
| | |

## 1. Introduction/Foreword

The purpose of this procedure is to ensure that newly developed or adapted methods of analysis for key analytical parameters in grapes and must are fit for purpose. This procedure outlines the requirements for development and validation of secondary analytical methods, that rely on the use of data generated from established primary reference methods.

## 2. Scope

This procedure applies to the quantitative measurement of analytical parameters in grapes and must including, but not limited to, pH, titratable acidity (TA), sugar concentration (Brix/Baume), colour (total anthocyanins) and total tannins. The intention of this procedure is to provide guidance on the application of secondary analytical methods, specifically those reliant on spectrophotometric-based measurement of the aforementioned analytes.

## 3. Terminology

Must is defined as unfermented crushed fruit juice that contains the skins, seeds, and stems of the grapes.

A secondary analytical method is one which relies on the construction and application of calibration models, built using data produced from established primary reference methods, such as those referenced in IESP's for sugar and colour measurement.

Spectrophotometric measurement is defined as the quantitative measurement of the reflection or transmission properties of a material as a function of wavelength.

Chemometrics is the application of mathematical and statistical methods and models in order to correlate chemical parameters or physical properties with analytical instrument data.

## 4. Measurement Principle

The measurement principle is based on spectrophotometric measurement, whereby the quantification of analytical parameters is governed by the application of the Beer-Lambert Law, which describes the linear relationship between the absorbance and concentration of an absorbing species. This applies to all relevant

regions of the electromagnetic spectrum, with the most commonly utilised being the ultra-violet (UV), visible (Vis) infra-red (IR) and near infra-red (NIR) regions.

## 5. Units of Measurement

Output units are dependent on the analyte(s) being determined for the relevant secondary method.

## 6. Metrological/Technical Requirements

These will be defined by the type of instrumentation and method being employed.

## 7. Health and Safety Considerations

Refer to the key risk elements, safety considerations, control measures and personal protective equipment identified in the primary analytical method for the relevant analyte(s).

## 8. Materials/Apparatus

These will be defined by the type of instrumentation and method being employed.

## 9. Materials/Reagents

These will be defined by the type of instrumentation and methods being employed.

## 10. Verification/Calibration

In general terms, instrumentation stability and performance should be evaluated by using a suitable standard reference material against the following criteria:

- Wavelength accuracy
- Wavelength repeatability
- Response repeatability
- Photometric linearity
- Photometric noise

These performance aspects are typically evaluated using internal diagnostic checks built into the instruments' application software and should be run on a daily basis to ensure correct and accurate operation of the instrument.

Secondary analytical methods, especially those employing spectrophotometric measurement of analytes, will typically incorporate application software in order to convert absorbance responses into calculated output values for individual analytes, via the use of chemometric techniques. Where application software is used, a reference to the software, including version number used, must be recorded and monitored. The impact of any changes or updates to the application software on constructed calibration models should be assessed prior to

executing updates. The validation process must be repeated if there are any changes to the measurement accuracy or precision of the calibration models, as a direct result of software updates.

The location of operating instructions and any specific requirement of the software must be identified in individual methods, including the procedure for naming of result and/or calibration files.

## 11. Environmental Conditions

Any environmental conditions that are deemed to be critical to the quality of the secondary method, and the test result generated by that method, must be specified. As far as is practical, calibration model development and validation should be carried out using stable measurement conditions that are consistent with the intended application of the method and incorporate all relevant practical sample handling considerations.

## 12. Measurement Procedure

### Sample Handling

Refer to the individual IESP method with regard to appropriate sample handling requirements.

If grapes can be evaluated whole, these should be presented in-tact (i.e. no splitting in the skin) with analysis directed to multiple points on the berry surface. If homogenate or juice samples are required, the berries should ideally be homogenised using the settings appropriate to the model of homogenizer unit, ensuring that all seeds are thoroughly macerated. Some secondary methods can accommodate unclarified grape homogenate samples for analysis; these should be presented in duplicate, to account for heterogeneity of the samples. If clarification of the juice is required, the juice should be transferred to a 50 mL centrifuge tube and centrifuged at approximately 3500 rpm for 5 minutes, prior to sample presentation of at least two replicate samples.

### Development and Validation Procedure

The development of secondary analytical measurement methods typically involves parallel screening of samples using both the established primary reference method and the secondary measurement method, so that a direct comparison between the methods can be made and to ensure that primary method data is available for constructing an appropriate calibration model. Ideally, these measurements should be made at the same time, to reduce inherent variability, due to sample degradation or transformation.

The range and variety of samples used for the calibration sample set should be balanced and as diverse as possible, to ensure that the calibration model encompasses the expected variation in sample composition and so that samples presented for subsequent quantification fall within the ranges defined by the calibration and validation process. If relevant, this may include early and late-harvested fruit, a broad selection of grape varieties (both red and white), fruit sourced from different regions and fruit harvested across multiple vintages. Accordingly, the appropriate size of calibration sample sets is dependent on all of these factors and therefore should be as large as practically possible in order to generate a calibration model of good predictive ability.

Due to the nature of these secondary methods and their dependence on incorporating samples and reference data from multiple vintages, calibration model development and validation is expected to be an iterative

process. A typical lifecycle for this type of method would involve initial development and validation using samples from a single vintage, followed by multiple model revisions and updates, to include fruit from subsequent vintages. At each time-point, the accuracy and precision of the models should be evaluated to ascertain whether further model development and upgrades are required in subsequent years; this would be facilitated by the inclusion of additional testing via the primary reference method and comparison with output data from the secondary method on a subset of the samples being screened.

The main steps in the process of building and validating a calibration model are as follows:

- Sample scanning
- Displaying and checking spectra
- Sample selection – calibration and independent test sets
- Data pre-processing
- Building calibration models
- Validation of calibration models
- Performance verification
- Maintenance of the calibration model
- Method transfer

## 12.1 Sample scanning

Spectra should be collected using a single, optimal method of sample presentation and instrument parameters for the screening method. A sufficient number of scans should be co-averaged to obtain suitable signal-to-noise levels. Samples should be scanned at least in duplicate and selected from the sample set in random order. The secondary method output (e.g. spectral scans) may be affected by temperature, so it is important to build, validate and maintain calibration models at a consistent temperature; preferably one that represents the likely presentation of future samples. If presentation temperature cannot be controlled (e.g. in the vineyard), it may be necessary to record spectra of each sample over a range of temperatures and build this variability not the model. Consideration should be given to the potential impact of other experimental variables, possibly by use of experimental design to identify the significant factors.

## 12.2 Displaying and checking spectra

Each spectrum should be reviewed to ensure that they are complete over the entire wavelength range of interest and suitable for use in the method. Obvious outliers, such as spectra that feature excessive noise, unique and/or unexpected bands should be rejected and samples re-scanned. The spectra obtained during the development and validation phase should be comparable with those obtained during the implementation phase; this is typically evaluated using similarity indices that are generated using the instrument application software.

## 12.3 Sample selection – calibration and validation test sets

The samples collected for use in the generation and optimisation of the method should be divided into calibration and independent validation test sets. The calibration set is used to correlate the output from the secondary method with the primary reference data and should cover the full expected range and variation in grape, homogenate or juice samples. The validation (holdover) test set is used to assess the accuracy of the calibration model(s); the samples in this set should cover, but not exceed, the range of variation in the calibration set. During the implementation phase, it is acceptable to interpolate within the range of the calibration model(s), but caution should be exercised when extrapolating beyond it. As well as considering even distribution across the range of the calibration and validation test sets, the relative numbers of samples in each

set should also be carefully apportioned, with the validation test set comprising of at least 30% of the total number of samples in the entire sample set.

## 12.4 Data pre-processing

Pre-processing is a vital step in the chemometric analysis of spectroscopic data. Pre-processing can be defined as the mathematical manipulation of spectral data to enhance spectral features and/or remove or reduce unwanted sources of variation (e.g. light scattering) prior to the development of the calibration model. The selection of the optimum pre-processing tool will often depend on the nature of the samples being screened, the type of instrument being used and the relative concentration(s) of analytes present. Chemometric tools and applications are used iteratively to determine the most appropriate pre-processing method and these are commonly built into the instrument application software. In some cases, no data pre-processing will be required; this is often the case where analyte concentrations are significant and/or the incidence of interfering species is low.

## 12.5 Building calibration models

Broadly, there are two distinct approaches to the generation of calibration models: univariate and multivariate. Univariate calibration is most common, where a single response from an analytical instrument is related to the concentration of a single component. Multivariate calibration is the process of relating multiple responses from an analytical instrument to the concentration of a component. In either case, the calibration model is generated using the calibration sample set and an assessment made as to the quality of the calibration obtained using the independent validation test set.

Ideally, the accuracy and precision of the method should be comparable to that of the reference method. Consideration should be given to root mean square error of calibration (also referred to as the root mean square error of cross validation or RMSECV):

$$RMSECV = \sqrt{\sum_{i=1}^{n} \frac{(\text{Predicted} - \text{Actual})^2}{\text{n} - 1}}$$

The maximum RMSECV should be three times the standard error (StdErr) of the reference analysis method (i.e. RMSECV < (3 x StdErr).

The regression coefficient ($R^2$) for the method comparison should be calculated and considered but does not have the same relevance or importance as it does for traditional univariate methods and reliance should not be placed upon it. The slope and bias of the calibration line should be as close as possible to 1 and 0 respectively, to ensure that the calibration model is robust and accurate across the entire concentration range. When reviewing selection criteria for calibration models, consideration should be given to the spectral range(s) being utilised, to ensure that they relate directly to the component(s) in question. Other aspects that should be considered include the calibration variable factors (loadings), residuals and other species that may affect spectral covariance.

## 12.6 Validation of calibration models

The accuracy and precision of calibration models should be assessed using the independent validation test set. The performance of the calibration models is typically assessed by the root mean square error of prediction (RMSEP), which should be comparable to the RMSEC or RMSECV:

$$RMSEP = \sqrt{\sum_{i=1}^{n} \frac{(\text{Predicted} - \text{Actual})^2}{\text{n}}}$$

The maximum RMSEP should be two times the RMSECV for the calibration and no more than five times the standard error of the reference analysis method (i.e. RMSEP < (2 x RMSECV); RMSEP < (5 x StdErr)).

Again, the regression coefficient ($R^2$) can be calculated but should not be considered critical for the method. The concentration range of the validation test set should be within that defined by the calibration model. If samples exist outside of the calibration range, these should be added into the calibration model and the impact on RMSEC or RMSECV assessed, prior to validation.

### 12.7 Performance verification

Once secondary analytical methods are being used for the direct measurement of analytes in grape, homogenate or juice samples, it may be necessary to periodically analyse a sub-set of samples using the primary reference method, in order to validate the ongoing performance of the calibration models. The data for the parallel testing set should be assessed in the same way as validation data for the independent test set, (i.e. RMSEP < (2 x RMSECV); RMSEP < (5 x StdErr)).

### 12.8 Maintenance of the calibration model

If secondary analytical methods are employed for screening grapes, must or juice, there is likely to be a need to incorporate samples and reference data from multiple vintages. There may also be a need to increase the range of grape varieties included, different fruit maturity levels and fruit sourced from different regions. This would necessitate multiple model revisions and updates, with the accuracy and precision of the models re-evaluated at each time-point via the inclusion of test data generated using the primary reference method(s).

### 12.9 Method transfer

The calibration models are generally stored and applied in electronic form as part of an appropriate instrument/software package and should not be transferred to another instrument unless procedures and criteria are applied to demonstrate that the model remains valid on the second instrument. In general, calibration transfer is only recommended to another instrument of the same type and configuration.

If transition to a different instrument type or model is required, the complete calibration and validation sample sets can be used to regenerate the method on a second instrument. In this case, the same method parameters as those used in the original calibration model(s) can be used and the calibration(s) are simply regenerated using spectral scans collected on the second instrument.

## 13. Uncertainty of Measurement

The uncertainty of measurement should be estimated using the top-down approach described in ISO/IEC Guide 98-3:2008 Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement (GUM:1995) or *Assessment of uncertainty in measurement* (Cook 2002). It is noted that most of the worked examples provided in these documents refer to metrological and not chemical measurements, and therefore it remains difficult to extrapolate the principles to chemical determinations.

Therefore, the uncertainty of a method is to be estimated by performing reproducibility tests at least 7 times. The mean and standard deviation of these data are calculated and the uncertainty of measurement (UOM) calculated based on 2 x the standard deviation.

## 14. Limits of Detection

Indicative LOD and LOQ for the secondary method(s) can be determined by reviewing the prediction error, as a function of analyte concentration. The point at which errors (predicted vs actual values) start to increase, as concentration decreases, can be used to define the nominal LOQ for the method. For more significant concentration values, the LOD will likely be the same as that for the primary reference method.

## 15. Reporting Results

This is dependent on the analyte being determined and the units of measurement outlined in the individual IESP procedure.

## 16. Validation

With respect to the development and validation of secondary analytical methods, a number of considerations can be accommodated into the process of building and testing calibration models. These include establishing appropriate concentration ranges, assessing models for adherence to linearity, evaluating the potential interference (specificity) from other species and assessment of model accuracy (e.g. RMSEC, RMSEP).

The precision of secondary analytical methods can be evaluated by assessing the repeatability of multiple measurements made using the same physical sample and the reproducibility of a test result when different operators are used to perform the test. In some cases, it may also be necessary to compare the reproducibility of a test performed in a laboratory with that performed at an alternative location, e.g. in the vineyard or at a weighbridge.

The robustness of a secondary method may be evaluated by incorporating experimental factors that can influence the performance of the calibration models. These may include: the age of samples (including how they have been stored), sample preparation techniques, analysis temperature and instrument configuration.

If any of these validation characteristics are to be used to monitor and validate the performance of the secondary analytical methods, an appropriate validation plan and schedule should be developed. These characteristics should also be evaluated following any changes or updates to the individual calibration models, as a result of including additional samples.

## 17. Quality Assurance

The reliability of secondary analytical methods can be monitored using the following procedure:

1. First and every tenth determination is performed in duplicate. Duplicates should agree within the stated MU values.

2. If the duplicate analysis criteria is not met, the analysis should be repeated in another batch.

## 18. References

Cook, R. R. 2002. *Assessment of uncertainties of measurement for calibration and testing laboratories*. 2nd Ed. Australia: National Association of Testing Authorities.

ISO 98-3:2008. *Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)*. Geneva, Switzerland: International Organization for Standardization.