

New methods to optimize wine production at all stages from vineyard to bottle

Kai Velten, Jonas Müller, and Dominik Schmidt

Hochschule Geisenheim University, Modeling and Simulation Workgroup, 65366 Geisenheim, Germany

Abstract. A software platform “Gm.Linux” (“Geisenheim-Linux”) has been developed in recent years at the Center of Wine Science & Beverage Processing Technology (HS Geisenheim University), which can be used to analyze and optimize wine production processes at all stages from vineyard to bottle. Gm.Linux provides a comprehensive and up-to-date collection of free scientific open-source software with special relevance to wine production processes, and it is available at no charge at the Center’s Internet pages (modellierung.hs-geisenheim.de/). Based on free open source virtualization software, the system can be used directly after download with no installations on all common platforms such as Windows, MacOS etc. A reference card system provides step-by-step instructions for many practically relevant analysis and optimization procedures. Several examples demonstrate the effectiveness of this new approach from the practitioners, engineers and scientists point of view.

1. Introduction

At all stages from vineyard to bottle, the winemaking process relates to a great amount of electronic data, such as economical, logistical, geographical, meteorological, CAD machine data, etc. Beyond this, an increasing amount of data is generated by modern equipment involved in state-of-the-art winemaking, such as sensor data from fermentation vessels, GPS-tractors, filling machines, etc. Although these data are available in principle today, they are typically not used for the analysis and optimization of winemaking processes. In the past, the availability of suitable software for analysis and optimization was one of the main obstacles that hindered a deeper exploitation of winemaking data. Based on this analysis, a software platform “Gm.Linux” (“Geisenheim-Linux”) has been developed at the Center of Wine Science & Beverage Processing Technology (HS Geisenheim University), which can be used to analyze and optimize wine production processes at all stages from vineyard to bottle [1]. It is based on open source software that can be used free of charge and thus opens up the way for winemaking optimization even in smaller companies with low R&D budgets. Some representative example applications will be discussed in the following.

2. Gm.Linux operation modes

Gm.Linux can be operated on any existing computer hardware and operating system (Window, MacOS, Linux,...) without affecting the original system, similar to other application software. After downloading the ISO file of the system at modellierung.hs-geisenheim.de and following the “Quick Start” instructions, Gm.Linux and its software will appear and work within a window of the original operating system. Beyond this virtualized operation mode, Gm.Linux can also be used as a main operating system or

in server mode (e.g., for computationally expensive simulations, see [1]). Currently, the Gm.Linux related information on modellierung.hs-geisenheim.de is available in German only (English webpages are under construction).

3. Examples

3.1. Spreadsheet and database facilities

Winemaking data are typically available in various spreadsheet and database formats. Sensor data, for example, often use simple csv formats while the treatment of data from more complex machinery may require sophisticated database software. Gm.Linux offers comprehensive spreadsheet and database facilities as described in [1].

Particularly in the wine consulting business, sophisticated spreadsheet operations beyond the application of standard Excel-functions are often required. Gm.Linux supports computer algebra based, automatized procedures that are apt to generate any Excel function that may be required, regardless of its complexity. In [1], a simple compound interest financing example is discussed that requires a surprisingly complex Excel formula (12 book text lines). It can be automatically generated and used by a copy and paste mechanism in Gm.Linux.

3.2. Complex algebraic formulas

In modern winemaking machinery such as large fermentation vessels, GPS assisted tractors, irrigation or harvesting systems etc. optimization of relevant data is often reported in the form of x/y scatter diagrams and mathematical functions representing these data (e.g. temperature data, GPS positions, water usage etc.). These data provide important information for process optimization. To unlock this information, a mathematical analysis of the formal functions behind the measurement curves is required in many cases.

Gm.Linux supports this kind of analysis based on its computer algebra software Maxima (which performs the algebraic analysis) and based on its statistics package R (which helps to express data in algebraic format) [1–4].

A wine-related, simple example for this kind of analysis can be found in [4]: the tank-labeling problem, i.e. the question where the 1000, 2000, 3000 liter marks etc. should be placed at the front side of a lying cylindrical tank. Similar to the compound interest financing example discussed above, the result is surprisingly complex and needs computer algebra (or other numerical) treatment for a fast and efficient solution.

3.3. Image processing

Gm.Linux offers image processing tools which are particularly useful for the analysis of digital image data. Typical winemaking applications relate e.g. to the automatized assessment of plant growth and ripeness parameters or of irrigation measures, or to automatized area calculations from aerial images etc..

3.4. Automatic reporting and data analysis

After a procedure e.g. for the optimized operation of a wine fermentation vessel has been defined, it is necessary to implement this procedure in an automatized and user-friendly fashion. This can be achieved using the data analysis and automatic reporting facilities of Gm.Linux. In fact, Gm.Linux incorporates a concept called Gm.HYDRA which performs data analysis, optimization and automatic reporting in one single step [1]. In the case of the fermentation vessel, for example, Gm.HYDRA would use all available fermentation data, apply the fixed optimization scheme and then generate a standardized report for human evaluation (and correction, if required).

3.5. Systems analysis

Before optimization procedures can be applied, it is often necessary to identify the most important influencing factors in a first step. If, for example, an undesired aromatic compound arises in wine fermentation, extensive data of a great number of possible influencing factors from all stages of winemaking are often available, but the question is how to identify the relevant influencing factors from this database. Gm.Linux provides a variety of procedures based on recently developed automatized statistical procedures (such as random forest methods, decision tree methods, classification methods) for the identification of relevant influencing factors [1,5–7].

3.6. Software for computers and mobile devices

Particularly in the wine consulting business, today's customers expect more than just an uninspiring Excel based analysis e.g. of wine economical data. What they expect is appealing, easy-to-use and attractive software that works on modern platforms including mobile devices. Gm.Linux offers a powerful template based on R's shiny package that can be used to hide all technical aspects of an analysis

behind an attractive, browser based software that works on all modern devices [2].

3.7. Simulation

Sophisticated 3D-computer simulations of many wine-related processes can be performed in Gm.Linux. Similar to previous work in the modeling workgroup at Hochschule Geisenheim, heat, water and nutrient transport in soils can be predicted using OpenFOAM, a 3D simulation software available in Gm.Linux [8,9]. This (or similar) software has also been used in the workgroup to perform 3D computer simulations of filtration or bottle related processes [10–12]. Gm.Linux offers sophisticated software both for the processing of CAD data in the preprocessing step of the simulations and for the postprocessing step including high-quality video generation (Salome, Paraview) [1,13,14].

3.8. Optimization

Several Gm.Linux tools can be used for the automatized optimization of wine related processes. If e.g. a great number of parameters of a wine fermentation process needs to be adjusted towards optimal process performance, the "R Commander" software can be used to define a minimal set of trial runs based on modern DOE (design of experiment) procedures [15,16]. Simulation results or online sensor or machine data can be optimized based on the Dakota package [17]. Even modern six sigma based techniques for the optimization of large process chains can be used based on an R package [18].

4. Conclusion

Winemaking is a complex process involving a great number of technological parameters. Today, most of these parameters are still set mainly based on empirical and heuristic considerations. To ensure process stability and final product quality, winemakers tend to keep or only slightly modify working process configurations. Thus, it is likely that considerable improvements of process performance and product quality can be achieved based on a deeper understanding of the subprocesses of winemaking, and on the application of model and simulation based optimization procedures.

Even small wine companies with little or no R&D resources can use the software tools developed in this project for the optimization of winemaking parameters beyond traditional empirical procedures. This is achieved by 1) using free, scientific open source software, avoiding expensive commercial CFD, modeling and statistical software that would sum up to more than 20.000€ for the software mentioned in this paper, and 2) by using Gm.Linux based software configurations that allow for the virtualized operation of the whole software bundle on simple desktop computers without any tedious installation procedures.

The future competitiveness of winemakers around the world depends on their ability to exploit existing optimization potentials based on these modern model and simulation-assisted tools.

References

- [1] M. Günther, K. Velten. *Mathematische Modellbildung und Simulation* (Wiley-VCH, Berlin, 2014)
- [2] R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria (<http://www.R-project.org/>)) (2014)
- [3] W. Schelter. *Maxima Manual Version 5.30*, (<http://www.maxima.sf.net/>) (2013)
- [4] K. Velten. *Mathematical Modeling and Simulation: Introduction for Scientists and Engineers* (Wiley-VCH, Berlin, 2009)
- [5] M. Forina, C. Armanino, M. Castino, M. Ubigli. *Vitis* **25**, 189–201 (1986)
- [6] R. Wehrens. *Chemometrics with R: multivariate data analysis in the natural sciences and life sciences* (Springer-Verlag, Berlin Heidelberg, 2011)
- [7] G. Williams. *Data Mining with Rattle and R: The art of excavating data for knowledge discovery* (Springer Science & Business Media, New York, 2011)
- [8] H. Jasak. *Int. J. Nav. Arch. Ocean Eng.* **1**, 89–94 (2009)
- [9] K. Velten, H. Neunzert, H. Zemitis, O. Iliev. Analysis of transport processes for layered porous materials used in industrial applications. In: *Mathematics - key technology for the future* (Springer-Verlag, Berlin Heidelberg, 2003)
- [10] F. Geiger, K. Velten, F. Methner. *J. Food Eng.* **109**, 609–618 (2012)
- [11] C. Meister, K. Velten, F.-J. Methner. *Int. J. Food Sci. Tech.* **47**, 1468–1478 (2012)
- [12] K. Velten, M. Günther, B. Lindemann, W. Loser. *Filtration* **4**(4), 276–280 (2004)
- [13] U. Ayachit. *The ParaView Guide: A Parallel Visualization Application* (Kitware Inc., New York, 2015)
- [14] V. Bergeaud, V. Lefebvre. SALOME: a software integration platform for multi-physics, pre-processing and visualization. *Proceedings of SNA + MC2010* (2010)
- [15] J. Fox. *J. Stat. Softw.* **14**, 1–42 (2005)
- [16] U. Groemping. *RcmdrPlugin.DoE: R Commander Plugin for (industrial) Design of Experiments* (2013)
- [17] M. S. Eldred, A. A. Giunta, B. G. van Bloemen Waanders, S. F. Wojtkiewicz, W. E. Hart, M. P. Alleva. *DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 4.1 reference manual* (Sandia National Laboratories Albuquerque, New Mexico, 2007)
- [18] E. Cano, J. Mogueza, A. Redchuk. *Six Sigma with R* (Springer, New York, 2012)