

No. 04/06_e

**DVFA-Rating Standards
and
DVFA-Validation Standards**



Imprint:

DVFA
Einsteinstraße 5
DE-63303 Dreieich
Tel.: +49 (0)6103 - 58 33-0
Fax: +49 (0)6103 - 58 33-34
Mail: info@dvfa.de
Web: www.dvfa.de

ALL RIGHTS RESERVED. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher and the copyright holder.



**DVFA-Rating Standards
and
DVFA-Validation Standards**

DVFA-Committee Rating Standards

Abstract

The sustained expansion and growing internationalization of the financial markets over the past several years has given rise to a major increase in demand among market participants for company assessment benchmarks that are both comparable and concise. Additionally, a varied range of external rating services has developed since the adoption and publication of the final Basel II framework – International Convergence of Capital Measurement and Capital Standards – by the Basel Committee on Banking Supervision in June 2004, and the respective drafts published since 2001.

In response to all this, the Society of Investment Professionals in Germany (Deutsche Vereinigung für Finanzanalyse und Asset Management – DVFA) set out to develop standards that would allow an evaluation of the various rating models. Ensuring information transparency is the centerpiece of this undertaking, in order to fulfill the increased informational needs of market participants with regard to company ratings. In 2000, the DVFA established the Rating Standards Committee under the chairmanship of Prof. Dr. Jens Leker (University of Münster).

For the development of the Rating Standards, the committee created four expert groups, led respectively by Prof. Dr. Dr. h.c. Jörg Baetge (University of Münster), Prof. Dr. Harald Krehl (DATEV eG), Dieter Pape WP/StB (URA Unternehmens Ratingagentur AG) and Prof. Dr. Heinrich Rommelfanger (University of Frankfurt). The committee also includes representatives of various rating agencies, banks, investment firms, audit companies, consulting firms and universities. In 2001, the committee published its initial results in the form of the DVFA Rating Standards (FinanzBetrieb, 2001, Issue 4).

In reaction to growing international capital market trends and the final version of Basel II, the Rating Standards Committee has now “internationalized” the DVFA Rating Standards, and adapted them to the new Basel II framework.

The new Rating Standards include general standards for company ratings and special standards for the validation of rating methods. The general standards address fundamental questions relating to company ratings and the phases of the rating process. They contain definitions of company ratings as well as requirements with respect to general information and the basis for and processing of information within the rating process. Furthermore special criteria for the evaluation of mathematical-statistical rating methods have been developed.

The Validation Standards developed by expert group 2, chaired by Professor Dr. Dr. h.c. Jörg Baetge, deal with selected questions of securing validity in the development of rating methods. The subject of these standards is the conceptual basis of validation, the special aspects of mathematical-statistical validation techniques and the structure of the validation process. The Validation Standards apply to all rating methods, such as the IRB approach, as well as to rating checks, rating software and rating systems. In particular, they also apply to all externally provided ratings.

The standards are presented as a catalog of questions and information, allowing an evaluation of the method and the results generated by the rating method under consideration. Their acceptance is assured through the consideration of expertise from different financial market participants and institutions.

The committee is not aiming at evaluating external or internal (bank) rating methods offered on the market. It understands itself rather as a communication platform for all financial market participants involved in the rating process and interested experts. Compliance with the Rating Standards will strengthen transparency and the informational basis in the context of a critical evaluation of rating methods by financial market participants and supervisory authorities



Part I:

DVFA-Rating Standards Transparency for company ratings

A Definitions

1 Company rating

The term “rating” refers generally to methods for assessing certain criteria based on specific categories and assigning them a ranking. The rating models currently available on the market cover various types of ratings, such as: the issuer credit rating, the credit rating of individual financial instruments, client and supplier credit ratings or the equity rating.

The committee has limited the scope of the Rating Standards to the credit rating, as an indicator of the creditworthiness or earnings power of a company. The credit rating represents a comprehensive analysis of a company, incorporating all available relevant information, with the purpose of assessing the probability of default as defined below.

2 Default

A uniform definition of “default” is necessary as the basis for comparable default probability indicators. The new Basel Capital Accord proposes the following reference definition¹:

A default is considered to have occurred with regard to a particular obligor when either **or both** of the two following events have taken place:

- The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held).
- The obligor is past due more than 90 days on any material credit obligation to the banking group.⁸² Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings.

The elements to be taken as indications of imminent illiquidity:

- The bank puts the credit obligation on non-accrued status.
- The bank makes a charge-off or account-specific provision resulting from a significant perceived decline in credit quality subsequent to the bank taking on the exposure.
- The bank sells the credit obligation at a material credit-related economic loss.
- The bank consents to a distressed restructuring of the credit obligation where this is likely to result in a diminished financial obligation caused by the material forgiveness, or postponement, of principal, interest or (where relevant) fees.
- The bank has filed for the obligor’s bankruptcy or a similar order in respect of the obligor’s credit obligation to the banking group.
- The obligor has sought or has been placed in bankruptcy or similar protection where this would avoid or delay repayment of the credit obligation to the banking group.

¹ cf. International Convergence of Capital Measurement and Capital Standards – A Revised Framework (margin number 452-453), Basel Committee on Banking Supervision, June 2004, as translated by *Deutsche Bundesbank*.

B Questions

1 General information about the rating

1.1 Description of the rating approach

1.1.1 Purpose of the rating agency

What purpose is being pursued by the rating agency?

Is rating among the primary business areas of the agency?

A differentiation may be made between internal and external rating:

- External rating: Generally, rating is the primary business of the rating agency, which offers the service as a product to other market participants.
- Internal rating: Rating serves company management purposes, and is not the primary business of the agency. Rating is not offered to external market participants, but rather serves to increase the transparency of internal rating assessments.

1.1.2 Rating purpose

For what purpose was the rating prepared?

A differentiation may be made between the following rating purposes:

- Improved refinancing of listed and private companies via the capital market,
- Basis for the structuring of private contractual relationships,
- Prerequisite for public or private contracting,
- Collection of information before entering into individual business relationships, or
- Critical assessment of own business operations.

1.1.3 Rating process

What are the defined phases of the rating process?

In this context, rating processes may be broken down as follows:

- Few phases to none (data collection, data analysis) or
- Detailed architecture (preliminary planning, contract signing, data collection, data analysis, rating result, rating publication).

1.1.4 Informational basis of the rating system

On what types of information is the rating system generally based?

A differentiation can be made between the following quantitative and qualitative information:

- External accounting data concerning the rated company on the basis of the balance sheet and income statement,
- Supplementary external accounting data concerning the rated company (notes, management report, voluntary reporting),
- Reference data concerning other companies, sector data, competition data, trend and risk data for the sector, fundamental economic data,
- Qualitative data concerning primary and secondary value chains, management quality and company culture of the rated company, or
- Results of non-specific external ratings (e.g. company or sector ratings, etc.).

1.1.4.1 Sources of information

Which sources are used to gather the information on which the rating system is based?

Potential sources of information are internal and external accounting, company employees, suppliers, clients and other business partners of the rated company, as well as third-party information providers (e.g. external rating agencies, analysts, commercial credit reporting agencies).

1.1.4.2 Data time horizon

What is the time horizon of the input data?

The data may include:

- Preceding year's data,
- Historical data,
- Historical time series, or
- Projection for the near future based on forecasts by the rated company and the rating agency.

1.1.4.3 Data collection method

How was the data collected?

To what extent are the rating agency and the rated company involved in the rating process?

The data may be collected as follows:

- Primary data collection directly at the rated company by the rating agency's own analysts,
- Secondary data collection via analysis of data provided by the rated company, or
- Combination of primary and secondary data collection.

With respect to involvement in the rating process, the following scenarios are possible:

- The rated company is not involved in the rating process, or
- The rated company is actively involved in the collection of information.

How is this data validated by the rating agency?

The data may be validated by way of:

- Comparison with internal reference data (e.g. accounting data, employee surveys), or
- Comparison with external reference data (e.g. analysts' information).

1.1.5 Rating method

Which methods are used to consolidate the information about the rated company into a uniform rating result?

How are the criteria selected (selection of criteria), defined (definition of criteria) and aggregated (consolidation and weighting of the criteria)?

In general, the following scenarios are possible with respect to rating methods:

- Criteria selection, definition and aggregation are achieved using a mathematical-statistical model.
- Criteria are selected by experts; criteria are defined and aggregated using a mathematical-statistical model.
- Criteria selection and definition is the work of experts; aggregation uses a mathematical-statistical model.
- Criteria selection, definition and aggregation are performed by experts.

1.1.6 Quality assurance by the rating agency

Which structural and organizational principles are followed by the rating system?

Is the four-eyes principle upheld?

Is there a segregation of data collection and data analysis?

Are comparisons made with external rating systems?

Which other quality assurance measures are employed?

The following are possible quality assurance measures:

- All phases of the rating process, from data collection to communication of the final rating are subject to an internal control system (ICS), Particular care is taken to observe the principle of segregation of functions and the four-eyes principle,
- Feedback from the rating results end of the process to the data collection end as a way of influencing results is prevented (no adjustment in the event of unfavorable results),
- To the extent possible, rating results are externally compared to the results arrived at by other rating systems, or
- A plausibility check is in place.

1.1.7 Rating result

1.1.7.1 Rating scale

Is a scale employed to classify the rating result?

Can the scale used be mapped to a master scale?

What historical default rates are associated with the individual rating classes on the employed scale?

The following scenarios are possible for exhibiting the rating result:

- An internationally recognized rating scale is used, which can be mapped to a master scale,
- An individual rating scale is used; results are mapped to a master scale, and this is presented as well, or
- An individual rating scale is used, which cannot be mapped to a master scale.

Every rating result based on a rating scale should include:

- The name of the rating class and the associated default rates (in %), and
- The data basis and time horizon.

1.1.7.2 Publication

How is the rating result published?

Publication of the rating must be suited to the rating purpose.

Rating agencies should publish a report in accordance with generally accepted reporting principles. These reports should allow a comparison of track records between various rating agencies. Rating agencies should state whether:

- Rating results are normally published,
- Rating results are only published if desired or permitted by the rated company,
- Rating results are used only for the management purposes of the rated company, and not published.

1.2 Description of the business relationship between the rating agency and the rated company

1.2.1 Independence of the rating agency from the rated company

Is the rating agency independent of the rated company?

Ratings should be independent of the company's business interests. Independence of the rating agency means that no rating agent should simultaneously perform a consulting function for the rated company. The results by the rating agency should not be subject to political or economic pressure. The rating process should not be burdened by conflicts of interest arising from the management and/or partnership structure of the rating agency. As a rule, rating agencies should not provide buy/sell recommendations, market price assessments, or recommendations on the suitability of an investment for a specific investor.

1.2.2 Obligation of the rating entity to comply with ethical standards

To what extent is the rating agency subject to ethical standards?

Ethical standards apply to institutions, such as rating agencies, as well as natural persons, such as rating experts, credit analysts and other rating professionals. Their activities must be in line with statutory guidelines, in particular relevant European and German capital market legislation, as well as with the ethical principles applicable to academic professions in Germany and related international rules for the profession.

Adherence to ethical standards should be documented at institutional and personal level on the basis of at least the following catalogue of basic principles for ethical professional conduct:

- Proof of professional and ethical qualifications,
- Observance of the latest professional and academic standards; regular re-qualification in the latest standards,
- Establishment of organizational structures to prevent conflicts of interest; obligation to obey compliance rules,
- Disclosure of potential conflicts of interest, in particular by way of commercial ties with the rated company, as well as other potential sources of conflict, which could threaten a fair and professional assessment,
- Respecting the incompatibility of rating and consulting activities, in particular the prohibition against representation of conflicting interests,
- Respecting the autonomy, independence, neutrality and impartiality of the rating experts as the basis for qualified assessment results,
- Final rating only after professional, thorough and precise application of rating methods,

- Disclosure of information relevant for an evaluation of the rating in accordance with the Rating Standards,
- Regular updating of the data pool and timely disclosure to the market of material changes to the rating result,
- Compliance with due diligence obligations in data collection and processing, as well as for release of the information,
- Observance of the prohibition against publication of inaccurate or incomplete information, or the omission of material facts that could effect the rating result; inadvertent errors must be publicly corrected immediately after they are identified.
- Observance of the obligations to secrecy and confidentiality,
- Avoidance of illegal business practices, in particular violations of insider legislation and active corruption, and
- Observance of copyright protections and the prohibition against plagiarism.

1.2.3 Applicability of the terms and conditions of the rating agency

To what extent are the terms and conditions of the rating agency applicable?

The terms and conditions of business should be a component of the agreement between the rating agency and the rated company.

1.2.4 Designation of specific contacts

Are there persons within the rating agency designated as specific contacts in connection with the rating result?

The rating agency should designate the authors or producers of the study that will be available to the rated company in connection with rating issues.

1.2.5 Ensuring complete information and authorized sources

Are there measures in place to ensure that the management and designated sources at the rated company immediately forward all relevant information to the analyst team at the rating agency?

Have the designated sources and legal representatives been informed of the type and scope of information required by the rating agency?

The following are possible measures to ensure that all known or latent risks are recognized in the rating process:

- All relevant information is collected in the rating process; no separate research is required.
- The type and scope of the information required by the rating agency is defined in the rating agreement, and the legal representatives of the rating agency are obliged to make this information available by virtue of their signatures.
- At the end of the data collection phase, a declaration is obtained from the legal representatives that the analysts were provided with all relevant information, in order to validate the findings. This declaration is usually in the form of a checklist, contains the names of those designated as sources and is signed by the legal representatives.

Example:

Data on current or latent risks (process, liability, guarantee risks, etc.) of the rating agency have not yet been made available, or were first identified during the rating process. For instance:

- Key patent rights can be challenged by third parties or used without authorization,
- Recent studies reveal that certain substances used by the company could pose health risks, or
- Key suppliers may soon be facing delivery problems that cannot be compensated for through inventories or alternative suppliers.

1.3 Structural independence

Is the rating agent independent of the rating agency?

Is the quality assurance agent independent of the rating agent?

Is the rating result independent of downstream decisions?

see Item 1.2.2

2 Informational basis of the rating result arrived at by the rating agency

2.1 Evaluation of input data quality

2.1.1 Criteria selection

What is the basis for criteria selection?

Criteria may be selected on an individual, case-to-case, or systematic basis, or on the basis of a statistical distribution model.

2.1.2 Data aggregation

To what extent has the available input data already been aggregated by the rated company?

Has data aggregation led to the loss of valuable information, distortions or errors?

Care must be taken to ensure that material information is neither lost nor distorted in the process of data aggregation. A differentiation should be made between cases in which ratings are based on raw data aggregation by:

- The rating agent, or
- The rated company.

2.1.3 Credibility

How credible is the internal and external input data?

- Planning and forecast data from the rated company are included.
- Third-party data (public and non-public institutions) is included.
- Unaudited annual financial statements, interim or preliminary financial statements prepared by the company are included.
- Audited financial statements are included.

2.1.4 Processing

How are missing data and information treated?

How are aberrations treated?

When using mathematical-statistical models, aberrations and missing values must be treated properly.

2.1.5 Consistency

How is data consistency evaluated?

The consistency of the actual data for the rated company should be evaluated, and measures or methods used for this evaluation should be disclosed by the rating agency.

2.1.6 Scaling

How is the input data scaled?

The input data may be scaled nominally, ordinally or metrically. The respective scaling requirements of the selected methods must be satisfied.

2.1.7 Internal control

How is the data input by the rating agency?

How are input errors avoided?

A description must be provided for the data collection process, the transfer of data to processing and the evaluation of data.

2.2 Information categories

What data does the rating system use?

An evaluation of rating quality requires insight into the informational basis available to the rating agency. The following is a checklist of potential information sources that can be used to rate a company.

2.2.1 Information on the identity of the rated company:

Is information included, e.g. concerning the legal form, locations, size and ownership structure of the company?

- *Company name*,
- *Legal form*,
- *Registered office and locations* (commercial registry, nationality, registered offices, locations),
- *Establishment date*,
- *Sector classification(s)* (sector identifier, regional focus of procurement, production and sales),
- *Company size criteria* based on disclosure and co-determination legislation (total assets, revenue, employees),
- *Group membership* (as controlling group parent, as dependent subsidiary),
- *Ownership structure* (sole proprietor, family business, public-sector enterprise, venture capital companies...), and
- *Historical company performance*..

2.2.2 Accounting data from the rated company

Is information, e.g. concerning type of financial statements, legal basis for accounting or relevant projections included?

- *Scope of accounting* (single-entity, group),
- *Legal basis for accounting* (HGB, IFRS, US-GAAP, tax laws, other local standards, type and timing of accounting transition),
- *Type of financial statements* (regular annual, interim, extraordinary),
- *Auditor's opinion/certification* (auditor's opinion (*Wirtschaftsprüfer*), certification by tax advisor (*Steuerberater*) involved in the audit, change of auditor/tax advisor),
- *Formal approval of the annual financial statements* (by supervisory board, by shareholders' meeting, jointly by management and supervisory boards, approval of annual financial statements),
- *Time frame* (number of financial years, gaps in the current financial year, interim period, shortened financial years),
- *Projected financial statements* including planning horizon (complete annual financial statements, projections of individual data / figures), and

- *other financial projections* including the planning horizon (revenue projections, material expenditure projections, financial plans, investment plans, projected cost of capital).

2.2.3 Non-financial indicators of company potential

Is the development of future markets and market position taken into account?

Are management quality, location and environmental considerations and personnel resources taken into account?

- *Market and market position of the rated company* (life cycle of the sector and products, cost structures of the sector and the rated company, concentration trends in the relevant markets, sector performance, barriers to market entry, market regulation, market shares, dependence on volume clients, quality and level of innovation of the products and product range, distribution capacity, etc.),
- *Production and procurement* (extent of vertical integration, dependence on specific technologies, inputs or raw materials, dependence on high-volume suppliers, relative cost advantages, dependence on bottleneck production factors, raw material and energy supply security,...),
- *Management* (size of upper management, supervisory board/council, areas of qualification, recruiting and fluctuation of upper management, dependence of the company on singular personalities, networking of upper management, incentive systems, strategic development and strategy communication, company culture, quality management, management information systems, organizational structures, informational policy of the company, risk management system),
- *Personnel* (personnel structure, dependence on key persons, personnel planning, development and recruitment), and
- *Location/environment* (stability of the political environment and legal system, economic development promotion, restrictions, conditions for public procurement, economic development of the location, infrastructure, environmental management).

2.2.4 Risk analysis and risk management

Are potential risks distinctively defined and systematically identified?

Are all individual risks identified and correctly aggregated to an overall risk indicator.

Is detailed information recorded on the effectiveness and efficiency of risk management?

- *Definition of potential risks* (operational risks, strategic risks, delineation of individual risk types),
- *Key individual risks* (risks from bottleneck factors, default risks with respect to receivables from customers, supply chain risks, product liability risks, currency and derivative risks, risks from existing debt, etc.),
- *Aggregation of individual risks* (qualitative risk aggregation using a risk matrix, statistical risk aggregation on the basis of a Monte Carlo Simulation), and
- *Aspects of risk management* (risk planning, instruments of risk management, organization of risk management, Expert opinions on risk management, organizational integration and responsibilities, existence of early warning systems).

2.2.5 Information on reference companies for evaluation of the rated company

Has dynamic or static benchmarking taken place?

Have sector and cross-sector comparisons been conducted?

Are other mathematical-statistical comparisons conducted?

- *Chronological comparison* (single year, multiple years),
- *Individual company comparison* (name of the reference company, size data, legal form, specific sector classifications, significant structural differences in comparison to the rated company, data basis),
- *Sector comparison* (specific sector classification, number of reference companies, averages, distribution values),
- *Cross-sector comparison* (subjective evaluation, empirical evaluation, historical insolvency rate in the sector), and
- *Non sector-specific comparison* (scope of reference samples, prices of the reference samples, reference period).

3 Information processing in the rating process

3.1 General criteria for the evaluation of rating system quality

3.1.1 Rating model

What combination of rating models does the rating agency apply in its rating system?

In general, a rating system comprises several components or rating models. In practice, other information is used as a basis to evaluate the result arrived at by a rating model in all cases, except those in which there is absolute certainty about the accuracy of the rating. In many cases, material creditworthiness is assessed initially using a mathematical-statistical model, and other factors are applied later with respect to management quality, sector situation and market positioning, etc., in order to assign a final rating classification. This may be the result of another procedure and/or an expert opinion.

3.1.2 Transparency and plausibility

Is the rating result based solely on transparent and plausible criteria?

Are the structures of the rating system presented in a way that a qualified third party would understand?

The rating agency must provide quantitative and qualitative information that serves to answer questions of transparency and plausibility.

3.1.3 Validity

Can the rating system produce an issuer rating?

In order to ensure the validity of the rating system, care must be taken to ensure that the rating result is actually a measure of the intended feature – in this case, the issuer rating. For detailed information on the validation of the rating model, see DVFA Rating Standard Validation

3.1.4 Accuracy

How high is the probability of error attributed to the rating system used?

Accuracy means that the final rating result correctly represents the creditworthiness of the company under consideration. The probability of error serves as a measure of accuracy – as expressed in the alpha/beta errors. The alpha error (type I error) represents the empirical probability that an insolvent company may erroneously be deemed “solvent”. The beta error (type II error) represents the empirical probability that a solvent company may erroneously be deemed “insolvent”. As a further measure of accuracy, so-called “power curves”, or the error band below the curves (Gini Coefficient), indicate the discriminatory power of the rating model.

3.1.5 Reliability

Does the rating system consistently return the same results based on the same input?

What is the extent of corroboration between various rating analysts at the same rating agency?

Do different rating agencies arrive at the same result?

A rating model is considered reliable when equivalent input data lead to the same rating result. This can be assured by:

- Reducing the freedom of discretion with respect to the evaluation of qualitative criteria, and
- Properly training staff.

In addition to evaluations of reliability, the sensitivity of the rating system to subjective expert opinions should also be analyzed.

3.1.6 Granularity

Is there an adequate number of rating classes available to satisfactorily differentiate the results?

The number of rating classes should be suitable in relation to the volume of input data and the degree of aggregation.

3.1.7 Clarity of the rating result and integrity of the required input data

Can every rated company be assigned to exactly one rating class?

It should be possible to prepare a rating for each company that permits clear assignment to a single rating class.

3.1.8 Currentness and robustness

How often is the rating system reviewed and, if necessary, updated?

What events trigger a review of the rating system?

Who is responsible for updating the rating system?

The rating system should be regularly reviewed to assess its suitability. In particular, it must be assured that material changes in fundamental economic data are addressed through prompt adjustments to the parameters of the rating model.

3.1.9 Relevance

Does the system use all information relevant for the assessment of the rated company's creditworthiness?

Information is deemed relevant if its omission would modify the rating result. As a rule, all information relevant for the analysis of creditworthiness should be used in the rating process.

3.1.10 Influence of the rated company

Do the sector affiliation, legal form, size or other characteristics of the rated company have a significant impact on the rating system?

How detailed is sector differentiation?

It should be clear whether or not sector affiliation, legal form, size or other specific characteristics of the rated company are taken into account for the assessment.

3.2 Criteria for evaluating the quality of empirical rating models

3.2.1 Data basis

Does the data basis satisfy the requirements of the model with respect to integrity and quality?

For the sample, elements (e.g. companies) must be obtained, whose rating classification (e.g. solvent - insolvent) is known. The grouping variable is scaled nominally.

The companies selected for the sample should be representative of the population as a whole. Although no generally applicable method is available to ensure a representative sample, the probability of meaningful conclusions should increase with the size of the sample.

It should be possible to divide the data basis into discrete groups based on significant criteria. Criteria are considered distinctive and therefore “significant” if the characteristics of the groups under consideration are materially dissimilar.

3.2.2 Criteria selection

Do the criteria selected satisfy the requirements of the model?

Depending on the specifications of the model, there are different requirements with respect to the characteristics of the criteria. In particular, the requirements as to the (stochastic) independence and scaling (nominal, ordinal, interval or ratio scaling) of the data must be satisfied.

3.2.3 Performance of the model

Does performance measurement satisfy the requirements of the model?

The number of explanatory variables or total forms thereof should not exceed a reasonable upper limit based on the scope of the training sample. The selection of explanatory variables should cover the most important aspects of the criteria area, and the correlation of the criteria variables should be held to a minimum.

In the case of key annual financial figures: in order to fully represent the information potential of the financial statements, the criteria (in this case, external accounting figures) should reflect several areas (net assets, financial position, operating results, etc.).

3.2.4 Estimation of model parameters

The estimation of model parameters should be based on a portion of the overall sample, known as the training sample. The training sample must represent the best possible cross-sample of current and future obligors.

The training sample must be large enough so that the classification of the training set provides a useful indication as to the classifications for the entire sample.

Care must be taken to ensure that the estimated parameters lend themselves to sound economic interpretation.

3.2.5 Performance of the model

In order to determine the classification performance of the rating model, the overall sample should be divided into at least one training and one holdout sample. The data from the companies included in the training sample serve to calibrate the model parameters. The holdout sample(s) serve(s) to validate the model (backtesting), and should include only those companies that have not already been included for model calibration. Statistical testing should be used to ensure that all samples represent discrete random sets from the overall sample, and the quality measure should be provided that permits evaluation of the rating model's discriminatory power.

3.3 Additional criteria for evaluation of individual rating models

3.3.1 Criteria for mathematical-statistical models

Does the estimation of model parameters satisfy the requirements of the mathematical-statistical model?

The estimation of parameters is often based on specific assumptions, which should be documented and binding. This is relevant, for instance, to assumptions about the distribution of explanatory variables, or to the mathematical requirements for the estimation procedure.

3.3.2 Criteria for scoring models

3.3.2.1 Defining the target system

Is information provided on the means of selecting individual factors?

Information should be provided on the selection of individual factors (subjective, statistical analysis).

3.3.2.2 Ranking of alternatives

Is information included on the methods used to determine the ranking of alternatives?

The range of the evaluation scale must suit the criterion being evaluated. However, the distance between the highest and lowest values on the scale does not have to be uniform for all targets. Points on the scale should be defined and depicted in such a way as to make the evaluation scale roughly metric. Information should be provided on the method used to select the criteria and to define their weighting (subjective, statistical analysis).

3.3.3 Criteria for expert systems

3.3.3.1 Knowledge base

What is the composition of the knowledge base?

About which categories can the knowledge base provide results?

From which sources was the available data collected?

The knowledge base may comprise:

- internal/external databases,
- (specific) expert observations, e.g. individual expert opinions concerning a certain segment of the company, or
- (standardized) expert experience that can be applied to the selection of criteria and definition of rating classes, or to the relevant aggregation operators or processing rules, etc..

Information should be made available on the qualifications (education, professional experience, references, obligation to comply with professional standards, etc.) of the experts involved in building the system. The expert knowledge applied must be documented in detail.

3.3.3.2 Oversight system

Which scaling requirements apply to the input data as a result of the aggregation method selected?

Are these requirements satisfied?

In the case of operation-based aggregation, metrically scaled input data should be available; for rule-based aggregation, ordinally scaled input data is sufficient.

Are linguistic evaluations of the criteria described adequately and clearly?

In the case of fuzzy logic-based systems, linguistic variables are described using fuzzy intervals represented by associative functions. The description should elaborate on the definition of the evaluation intervals, and indicate whether data information was used in addition to expert knowledge.

Is the evaluation process transparent?

The aggregation structure should be understandable to a knowledgeable third party. The aggregation operators and/or rule sets must be documented.

How often is the expert system reviewed and, if necessary, modified?

The events that trigger a review of the knowledge base and the experts involved in any modifications should be identified.

How long has the expert system been in use and how accurate have results been in the past?

The age of the expert system should be provided. For information on accuracy, see Item 3.1.4

3.3.4 Criteria for neural networks

3.3.4.1 Structure

What is the number of neurons in the input, hidden and output layers?

The number of neurons, hidden layers and connections should be indicated.

3.3.4.2 Training

Is information provided on training of the neural network?

Is information provided on the use of holdout samples?

Information should be provided on the computational algorithm used (supervised or unsupervised), as well as the pruning methods and the number of training runs. In addition, an indication should be made as to whether sensitivity analysis was used to monitor the influence of the input parameters.

The forecast results of the holdout sample should be used as a benchmark for the performance of the neural network. Validation should be based on a third, independent sample.

3.4 Criteria for evaluating qualitative rating models (classification by rating analysts based on a unspecified model)

3.4.1 Qualifications of the rating analysts

What information is provided in the qualifications and independence of the rating experts?

The rating agency should disclose its general (minimum) requirements and work standards for its analysts (education, training, experience, obligations to comply with professional standards, etc.).

3.4.2 Rating process

To what extent is the rating process standardized?

How detailed are the guidelines followed by the experts?

The rating process and the number of participating experts must be documented. Moreover, the requirements provided to the experts with respect to criteria selection and data currency should be stated, along with information on the extent to which benchmarks and knockout criteria are applied.

3.4.3 Rating quality assurance

How is the quality of the rating assured?

The conclusion arrived at by the expert(s) should be assessed by further persons. Compliance with the four-eyes principle and the existence of ongoing quality control measures should be documented.

4 Information on the rating result arrived at by the rating agency

4.1 General information on evaluation of the rating result

4.1.1 Comparability

Is mapping to a master scale possible?

If various rating scales from different rating agencies are used, it should be possible to represent each of these on a “generally accepted scale” (master scale), in the interest of comparability. This can only be achieved, however, on the basis of a uniform definition of “default”.

4.1.2 Transparency

The rating result must be transparent. The linkages between the selected criteria and the final result must be disclosed. This involves both public and individual disclosure, in order to allow proper interpretation of the rating result by the rated company and other financial market participants.

The rating must be transparently presented to the rated company individually, so that it can clearly understand the rating result. The rating result may not be amended after the fact, unless overt errors have been made in the rating process. Minutes should be kept of informational exchanges between the rating agency and the rated company.

External disclosure of the rating results depends on the rating purpose, especially as defined by the initiator of the rating, or its stated objective.

4.1.3 Monitoring

Does the rating agency regularly publish information concerning the quality of its rating results?

Monitoring and scrutiny of the rating results should primarily be a function of the market. To this end, the rating agency should regularly publish historical default rates for each rating class (track records). If mathematical-statistical models are used, information should be disclosed with respect to model significance, classification quality and monotonicity of the historical default data.

4.2 Information of the interpretability of the rating result

4.2.1 Individual information about the rating result

4.2.1.1 Company identity

What information is provided on the identity of the company?

see Item 2.2.1

4.2.1.2 Currency and date-marking of the rating result

Can the currency and date of the rating results be readily determined?

The date of the rating result and the time horizon of the input data should be indicated. Since a rating remains in effect until it is amended or withdrawn, the assessment of the rating agency should be subject to ongoing review, and immediately modified in accordance with any new information that impacts creditworthiness. All modifications of the rating result should be documented, and a history of the rating should be maintained.

The ratings should be amended with due promptness, immediately after the rating agency receives new information (quarterly or annual financial statements, ad hoc disclosures). Material changes to the creditworthiness of the rated company should result in prompt amendment of the rating.

4.2.1.3 Consideration of country risks

Information on sector or country-specific risks may be provided separately, or as an implicit component of the rating result.

4.2.1.4 Probability of default

Information should be provided as to the average probability of default for the rating class to which the company is assigned.

4.2.2 Rating scale

4.2.2.1 Description of the rating scale

Information should be provided about range and scaling, as well as mapping to a master scale.

4.2.2.2 Definition of default

“Default” should be defined in accordance with the Basel Capital Framework by the Basel Committee on Banking Supervision. Individual definitions that deviate from this must be explained in detail within their specific context, and should take account of partial default, loan restructuring, delay of payment and security pledged.

4.2.2.3 Probability of default in the individual rating classes

The mean probability of default and, to the extent possible, the intervals should be provided for each rating class.

4.2.3 Scope of the rating

4.2.3.1 Forecast horizon in years

A differentiation should be made between short-term (< 1 year) and long term (> 1 year) rating forecasts.

4.2.3.2 Sector

The relevant sector classification should be indicated, e.g. manufacturing, other industrials, banking/insurance, services, consumer goods, etc.).

4.2.3.3 Company size

Information about company size should be standardized in accordance with sections 267 and 293 of the German Commercial Code (HGB) and sections 1 and 11 of the Germany Disclosure Act (PublG).

4.2.3.4 Regional scope of application

To the extent possible, the rating should apply globally without limitations, but may be subject to local, national or international restrictions if there is good cause.

C Special Criteria for evaluating individual mathematical-statistical models

1 Criteria for linear discriminant analysis

1.1 Data basis (sample requirements)

Does the data basis satisfy the requirements of linear discriminant analysis?

Unmodified, this model is only suitable for use with metrically scaled input data. Nominally and ordinally scaled data should only be used if they can be mapped to a metric scale with sufficient accuracy.

1.2 Criteria selection and construction of the discriminant function

Do the criteria selected satisfy the requirements of the linear discriminant function?

In order to satisfy the theoretical requirements of linear discriminant analysis, the criteria selected for grouping should be normally distributed as far as possible, with identical variance-covariance matrices.

Is information provided on the form of the discriminant function?

If linear discriminant function analysis is selected as the methodological basis, the discriminant value Z is derived from the linear discriminant function to be determined with the discriminant function coefficient a_i and the explanatory variable KZ_i , in the general form: $Z = a_0 + a_1 \cdot KZ_1 + a_2 \cdot KZ_2 + a_3 \cdot KZ_3 + \dots + a_n \cdot KZ_n$

1.3 Estimation of the discriminant function coefficients

Does the estimation of the model parameters satisfy the requirements of the linear discriminant function?

The estimation of the discriminant function coefficients should be based on a portion of the overall sample, known as the training sample. The training sample must represent the best possible cross-sample of current and future obligors. It must be of sufficient size that the classification of the training set provides a useful indication as to the classifications for the group containing the company under consideration. Care must be taken to ensure that the estimated parameters lend themselves to sound economic interpretation.

When evaluating the discriminant function coefficients, it must also be ensured that the associated coefficients, and thus the indicator combinations in a classification function, do not allow conflicting economic interpretation.

Freedom from conflicting economic interpretation means that criteria deemed positive from an economic standpoint are included in the computed classification function with a different mathematical sign than those deemed negative. A function that meets these requirements is considered consistent from an economic standpoint.

1.4 Performance measurement of the estimated discriminant functions

Does performance measurement satisfy the requirements of the linear discriminant function?

The quality measure used to evaluate the discriminant function should be indicated. A description of the correlation between the discriminant value Z and the rating class should also be provided.

2 Criteria for logistic regression

2.1 Criteria selection and construction of the classification function

Do the criteria selected satisfy the requirements of logistic regression?

All soft factors should be integrated into the model in standardized form as independent variables, and should at least approximate metric scaling.

Is information provided on the form of the logistic regression function?

Estimated *a priori* probability is expressed as a logistic sigmoid curve.

2.2 Estimation of logistic regression coefficients

Does the estimation of the model parameters satisfy the requirements of logistic regression?

The estimation of model parameters should be based on the training sample.

The estimated classification function of the logistic regression should be consistent from an economic standpoint. For more information on consistency, see Item 1.3

3 Criteria for neural networks

3.1 Criteria selection, number of layers and neurons

Do the criteria cover all relevant information categories?

Do the criteria satisfy the scaling requirements for neural networks?

In the context of neural networks, direct input of qualitative criteria is technically unproblematic. For direct processing, each qualitative criteria characteristic requires an input neuron. For instance, the qualitative criterion "accounting behavior" may have the following characteristics: "conservative", "neutral" or "progressive"; these three input neurons are thus integrated into the network. For even more detailed characteristics, additional corresponding input neurons are required. Qualitative criteria quickly lead to more complex networks and very involved computation procedures.

Does the model use Back-propagated Delta Rule Networks (multilayer perceptrons) or Radial Basis Function Networks?

How many layers are included in the neural network and how many neurons does each layer comprise?

What activation functions are employed? Which sigmoid functions are applied, or does the model use Gaussian distribution?

3.2 Training of the neural network

Which computational algorithm (supervised or unsupervised) is used?

Which pruning methods were applied and how many training runs are conducted?

Are sensitivity analyses used to monitor the impact of the input parameters?

Is the training sample sufficiently large to provide a useful indication as to the classifications of the sample to be analyzed?

Does it represent a good cross-sample of current and future obligors?

3.3 Performance measurement of the neural network

Has the overall sample been separated randomly into the three discrete sets: training, holdout and validation?

The data from the training sample serves to calibrate the model parameters. The holdout samples are applied for the selection of the best model alternatives, the results of which are used as a benchmark for the performance of the neural network. Validation (backtesting) is accomplished with the help of the third, independent validation sample, which should only contain “as yet unaffected” companies.

4 Criteria for data clustering models (e.g. support vector machines)

4.1 Data basis (sample requirements)

Does the data basis satisfy the requirements of cluster models?

Unmodified, this model is only suitable for use with metrically scaled input data. Nominally and ordinally scaled data must be mapped to a cardinal scale.

For support vector machines (SVM):

The model implements a Bayesian estimator for classification. This allows optimal *a priori* classification of imprecise input vectors.

Missing data must be accordingly represented in the input vectors. Care must be taken to ensure that the missing data is taken into account as such by the model.

4.2 Estimation of an approximation or classification function

Does the estimation of the model parameters satisfy the requirements of the clustering model?

If the data comprises numerous, variously scaled values, common, multi-dimensional scaling of the input data can be useful.

For support vector machines:

An SVM is basically a mathematical-statistical model. Certain free modeling parameters, however, such as the form of the risk function or modifications for missing variables, must be determined by an expert.

The modeling estimate results in the support vectors and the Lagrange Multipliers, which represent class prototypes in the classification process. A review should be conducted with respect to the economic relevance of the input criteria for these prototypes.

5 Criteria for decision tree models (e.g. CART)

5.1 Construction of the classification function

Is information provided on the construction of the classification function in the decision tree model?

The definition of the potential (rating) classifications of the subjects (obligors) under consideration must be provided, and the selection of potential discriminant variables explained.

The impurity function that determines the characteristics of the discriminant criteria must also be defined, and the selection of the criterion explained.

The application of substitution splits in the case of missing values for the subjects to be classified, i.e. the use of other criteria values with similar classification effects, is permitted, but must be sufficiently documented.

5.2 Estimation of the default classification rate and determination of end nodes

Is information provided on estimation of the default classification rate for the model?

The resubstitution estimator or other estimator of the default classification rate used to determine the optimal size of the decision tree, i.e. the ideal number of steps needed to reach a final grouping of the set under consideration (end nodes of the decision tree) should be defined. The selection of this estimator should also be explained.

Is information provided on determination of the end nodes?

The method used to assign the end nodes of a decision tree to their specific classifications must be described. The selected assignment must result in an optimal default classification rate of the training set.

The method for final determination of tree size based on the default classification rate should be presented and the default classification rate of the final decision tree tested and documented using a holdout sample. The method used for testing with the holdout sample should be explained, with particular elaboration on how the original training sample was separated into a new training sample and accompanying holdout sample.



Part II:

DVFA-Validation Standards

A Definition

1. Validity of rating models

The term “validity” refers to the following characteristics of a rating model: that the model delivers accurate results, as compared to the measurement targets of the model as defined by its developer and the objectives as defined by the user (“Validity of Results”), and that the model is based on sound assumptions with respect to the relationship between the characteristics of the observed cases and the probability of default in each case (“Validity of the Model”).

Accurate results can be, e.g. well-grounded statements regarding the ability of a debtor to satisfy future obligations, or correct information relating to the probability of default or migration to a different rating class.

2. Quality of rating models

While validity is the most important mark of rating quality, robustness, stability and plausibility all represent further measures for the quality of rating models.

Robustness is the ability of the model to produce results in practical application, without allowing erroneous, distorted or improper inputs. Stability of the model means that it arrives at the same result for a given company over an extended period. Plausibility refers to the requirement that the result should be reproducible based on information available outside of the model.

As a further aspect of quality, the model should be generally compliant with the DVFA Rating Standards.

3. Validation of rating models

The validation of a rating model determines the degree to which the operational assumptions and rules used to arrive at the results (“operational/methodological validation”) are valid, as well as the degree of validity of the results arrived at by the model (“mathematical validation”).

Operationally relevant assumptions and rules are those concerning the effects of individual risk parameters and company-specific circumstances on the level of default probability, and the representation of the validated empirical findings. They allow the model to reflect the relationships identified in practice between information about the rated company and the probability of default.

The mathematical validation of a rating model refers to the mathematical-statistical determination as to the validity of the results produced by the rating model. This involves the testing of the default probabilities estimated using a representative training sample through application of an equally representative holdout sample and the appropriate mathematical-statistical model.

4. Rating class

A “rating class” is defined as an indicator of counterparty risk on the basis of various rating criteria from which the probability of default (PD) can be estimated. These rating classes are used by banks, for instance, to create portfolios with optimum risk-return profiles through the suitable selection of interest rates and lending volumes for each class.

In order to fulfill this optimization function, the rating system must have the ability to allocate a representative sample of loan applications in such a way that the defined probabilities of default can be achieved in all of the rating classes. Representative means, e.g. that the samples reflect the structures of the rated client segments of the bank, in terms of both the different levels of client creditworthiness and the amounts and terms of the loans.

B Questions

1. Methodological validation of rating models

1.1 Plausibility of rating results

- 1 Does the recorded data unequivocally produce the rating result?
- 2 Is the rating result reproducible and thus objective in the sense of intersubjective verifiability?
- 3 Was equal treatment of comparable data assured, independent of user or duration, in the sense of reliability of results?

1.2 Transparency of rating models

- 4 Is comprehensive, detailed documentation of the rating model available, which would be understood by a knowledgeable third party?
- 5 Do those performing the validation have unrestricted access to the “black box” of the model?
- 6 Are their guidelines in place with respect to the integration of the model into the “lending decision” process, and is there measurement of the extent to which the model is applied in practice?

1.3 Economic causality (assumptions)

- 7 Is there a link between statistical correlations and expert knowledge-based causality factors?
- 8 Are there hypothetical solutions for every rating criterion with respect to the translation of characteristics into assessment results or quantitative effects?
- 9 Is application of the rating model subjected to comparisons with theoretical approaches to estimating and/or explaining default?

2. Mathematical-statistical validation of rating models

- 10 How many rating classes have been established for non-defaulting or defaulting borrowers?
- 11 Are measures of relative concentration determined for the validation of the rating model?
- 12 Are indicators defined based on the comparison of default rates?
- 13 Have migration matrices been created?

3. Validation process

3.1 Validation process: methodological validation

- 14 Are the content and structure of the rating model plausible from an operational point of view?
- 15 Is the model based on robust assumptions?
- 16 Are input, throughput an output data available and of adequate quality?

3.2 Validation process: mathematical-statistical validation

- 17 Is an adequate number of observations available?
- 18 Is the number of critical cases (defaults) adequate? If not, is there at least an external benchmark available?

3.3 Consequences of methodological and mathematical-statistical validation

- 19 Are the results of the rating model validation interpreted and used as a basis recommending actions?

3.4 Organization of the validation process

- 20 Who performs validation of the rating model?
- 21 With what frequency is validation of the rating models performed?
- 22 What sort of documentation is kept on validation activities and results?

C Examples and explanations

1. Methodological validation of rating models

1.1 Plausibility of rating results

Methodological validation is especially relevant in relation to mathematical-statistical rating models and methods, since these do not automatically imply congruence between operational practice and model theory.

Validation is required, in particular, for determinants such as:

- the scope of application for the rating model and the rating process,
- the criteria and circumstances relevant to the rating (input),
- the aggregation mechanisms and their methodological bases (throughput), and
- the possibilities for manual modification and representation of results (output).

Validation, however, should also take account of supporting factors, such as:

- data storage, and
- other external and internal sources that can be utilized to assess the consistency of the information used in the rating process.

A key prerequisite for plausibility is the quality of the collected data, which is determined largely by the quality of the data collection process.

1.2 Transparency of rating models

Mere "coding" in the "hardware" of the model is not sufficient to ensure transparency. Rather, detailed documentation of the rating model and its aggregation mechanisms must be available, which would allow a neutral third party to form an understanding of all key aspects of the model within a reasonable period of time.

1.3 Economic causality (assumptions)

The information used in the model can only have the desired effect on the overall rating from an operational point of view if the following requirements are cumulatively met:

- a link between statistical correlations and expert knowledge-based causality factors,
- hypothetical solutions for every rating criterion with respect to the translation of characteristics into assessment results or quantitative effects (e.g. alphabetical grades, supplements, numerical scores, weightings, amplification of other factors, etc.),
- subjecting application of the rating model to comparisons with theoretical approaches to estimating and/or explaining default.

2. Mathematical-statistical validation of rating models

For the use of the advanced IRB approach as the basis for internal rating systems, Basel II prescribes a "meaningful distribution of exposures across grades with no excessive concentrations [within a single rating class], on both its borrower-rating and its facility-rating. To meet this objective, a bank must have a minimum of seven borrower grades for non-defaulted borrowers and one for those that have defaulted."

After the PD for each rating class has been estimated using a representative training sample, the quality of the rating model can be assessed using an equally representative holdout sample.

The validation data is comprised of the number and volume of defaulted and non-defaulted exposures in the individual rating classes. The application of the rating model to all of the elements in the holdout sample produces a data set in accordance with the following:

a_k	Volume of the defaulted exposures in rating class k , $k = 1, \dots, K$
s_k	Volume of the non-defaulted exposures in rating class k , $k = 1, \dots, K$
$x_k = a_k + s_k$	Total volume of the exposures in rating class k , $k = 1, \dots, K$
$M^a = \sum_{k=1}^K a_k$	Total volume of the defaulted exposures in the holdout sample
$M^s = \sum_{k=1}^K s_k$	Total volume of the non-defaulted exposures in the holdout sample
$M = \sum_{k=1}^K x_k$	Total volume of all exposures in the holdout sample

Rating class K is the “class for defaulted borrowers”, i.e. the rating system allocates to this class those clients of the bank who are highly likely to default and should therefore be denied credit. The remaining classes 1 to $K-1$ represent significantly lower probabilities of default, in order from lowest to highest. The extent to which loan applications are approved in the individual classes from 1 to $K-1$ is determined by the structure of the optimum portfolio.

The following example, with nine rating classes, serves as an example:

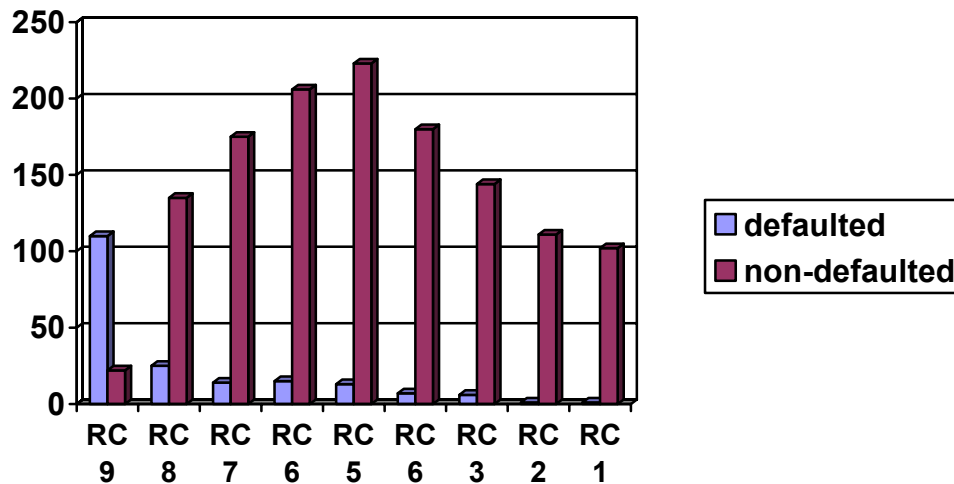


Fig. 1: Histogram of defaulted and non-defaulted exposures in 9 rating classes

For the sake of simplicity, the example assumes that all borrowers have been lent the same amount, so that only the number of exposures is relevant.

In the future, when banks have adequately comprehensive information available on individual exposures, validation should be performed annually on the basis of default experience in previous years.

Most literature on the subject recommends the use of concentration ratios based on the Lorenz Curve (power curve, cumulative accuracy profile) to gain an initial impression rating model quality. Notable examples of this are the Gini Coefficient, the Lorenz-Münzer Accuracy Ratio and the AUC profile (area under ROC curve). To this end, the shares attributable to the individual rating classes must first be determined, followed by systematic aggregation of the values.

$$A_k = \frac{a_k}{M^a} \quad \text{Relative volume of the defaulted exposures in rating class k}$$

$$S_k = \frac{s_k}{M^s} \quad \text{Relative volume of the non-defaulted exposures in rating class k}$$

$$X_k = \frac{x_k}{M} \quad \text{Relative volume of all exposures in rating class k, k = 1, \dots, K}$$

$$CA_k = \frac{\sum_{j=k}^K a_j}{M^a} \quad \text{Aggregated share of defaulted exposures}$$

$$CS_k = \frac{\sum_{j=k}^K s_j}{M^s} \quad \text{Aggregated share of non-defaulted exposures}$$

$$CX_k = \frac{\sum_{j=k}^K x_j}{M} \quad \text{Aggregated share of all exposures}$$

These values can be used, e.g. to calculate the following concentration ratios (see Fig. 2 and Fig. 3 for more information on the fields mentioned)

Concentration field = The area between the **LORENZ CURVE** (power curve, cumulative accuracy ratio) and the diagonal representing equal default frequency in all rating classes.

G = GINI COEFFICIENT = Ratio between the concentration field I and the overall field in the triangle above the diagonals.

AR = LORENZ-MÜNZER Accuracy Ratio = Ratio of the concentration field I and concentration for the ideal model (fields I + II).

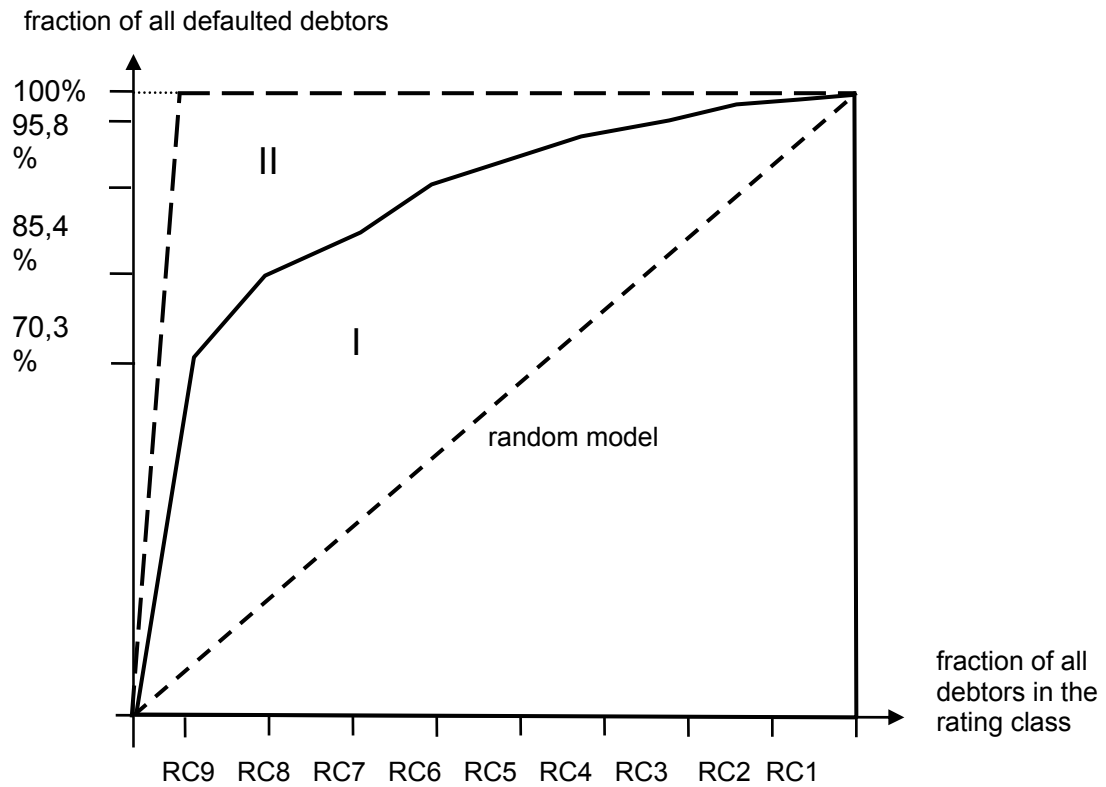


Fig. 2: Lorenz Curve

AUC = area under the ROC curve = field III + 0.5

(ROC = **Receiver Operating Characteristic**)

As defined: $0 \leq G < AR \leq 1$ and $0,5 \leq AUC \leq 1$

Hit rate
(correct „default“
Classification)

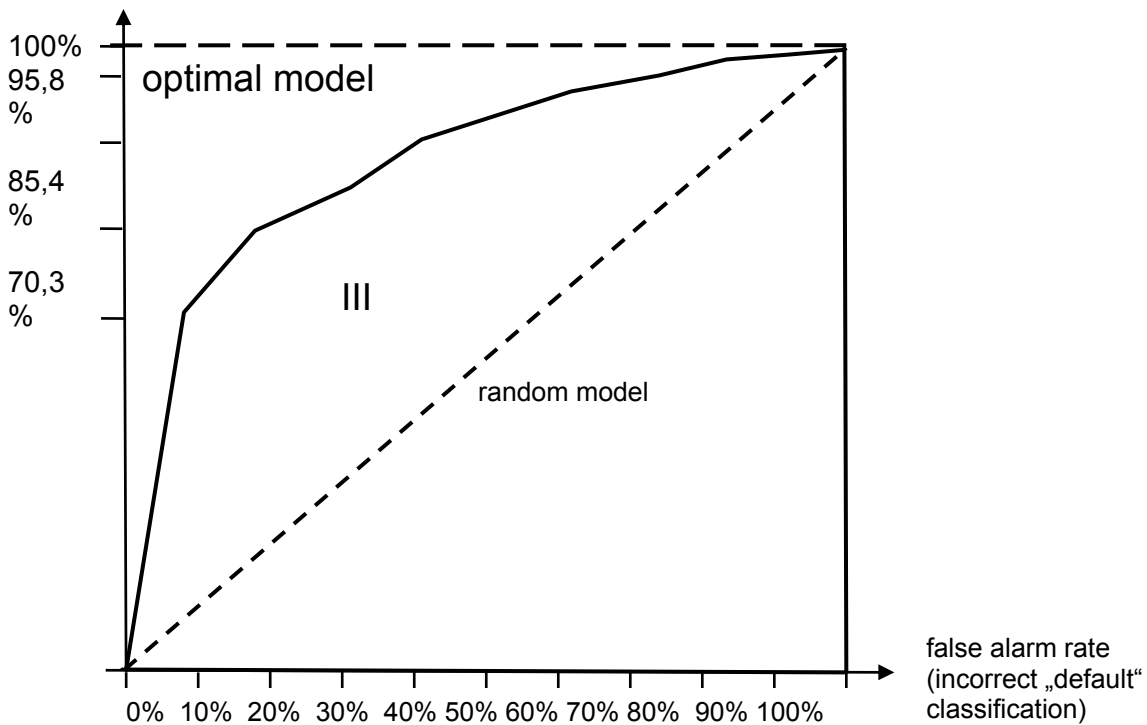


Fig. 3: ROC Curve

As a rule, the better the quality of the rating system, the higher the concentration ratio will be. But these ratios merely provide an index for adequately high-quality rating systems. Due to the broad array of applications, it is impossible to define a general benchmark. Furthermore, there is no “best” concentration ratio, since each has its own specific advantages and disadvantages. For instance, the Gini Coefficient and the Lorenz-Münzer Accuracy Ratio do not change if the number of defaults across all rating classes rises by q% and the total number of defaults in the rating classes remains constant or rises by p% distributed evenly across all classes (i.e. the number of defaults sees a relative rise to the detriment of solvent borrowers). The AUC profile does not change if, in every rating class, the number of defaults rises by q% and the solvent elements in the rating class remain constant or change evenly by p%.

Since, however, the above-mentioned concentration ratios can only provide a very rough indication as to the quality of a rating model, it is important to identify a validation method that is based on a comparison of the defined PD of the individual rating classes with the actual default rates observed in the holdout sample (hoPD). Given that lending portfolio planning is based on the defined probabilities of default, the rating model should distribute a representative sample across the rating classes in such a way that the observed default rates correlate as closely as possible to the estimated (defined) default probabilities.

An example is provided below of a validation method based on the assumption that only higher PDs are undesirable in classes 1 to K-1.¹ Since the observed cases allocated to rating class K are not to be granted loans, low PDs are disadvantageous here, and increase the β -deviation. Given that the β -deviation has a generally lower relative importance than the α -deviation, this should be accounted for via appropriate weightings g_k , which also allow varied weighting of deviations in the individual rating

¹ If positive and negative differences are equally considered, the distance $\Delta k = |hoPD_k - PD_k|$ should be used.

Additionally, a distance limit R can be defined as $\Delta k = |hoPD_k - PD_k| \leq R$ for all $k=1, \dots, K$.

classes.

If the default rates per rating class k in the holdout sample are represented by $hoPD_k$, and the defined probabilities of default with PD_k , then:

$$hoPD_k = \frac{a_k}{x_k} \quad \text{and} \quad PD_k = \frac{\bar{a}_k}{\bar{x}_k},$$

if \bar{x}_k and \bar{a}_k represent the total number of loans or lending volumes in rating class of the training sample.

Numerical example:

	RC 1	RC 2	RC 3	RC 4	RC 5	RC 6	RC 7	RC 8	RC 9
PD_k	85.3%	11.9%	8.5%	5.2%	3.8%	2.9%	1.8%	1.1%	0.5%
$hoPD_k$	83.3%	15.6%	7.4%	6.8%	5.5%	3.7%	4.2%	0.8%	1 %
$hoPD_k - PD_k$	-2%	3.7%	-1.1%	1.6%	1.7%	0.8%	2.4%	-0.3%	0.5%

Tab. 1: Probabilities of default by rating class

The following distance measure, in which higher probabilities of default in classes 1 to $K-1$ and lower probabilities in class K are deemed negative, is suited to calculating the distance Δ_k between $hoPD_k$ and PD_k .

$$\Delta_k = \text{Max}(0, hoPD_k - PD_k) \quad \text{for } k = 1, \dots, K-1 \quad (\text{in the example } k = 1, \dots, 8)$$

$$\text{and } \Delta_k = \text{Max}(0, PD_k - hoPD_k) \quad \text{for } k = K \quad (\text{here } K = 9)$$

A suitable indicator for validation of the rating model is therefore:

$$F = \sum_{k=1}^K \left(\Delta_k \cdot \frac{X_k + \bar{X}_k}{2} \cdot g_k \right)$$

with \bar{X}_k representing the relative lending volume in rating class k of the training sample.

As a benchmark, a value F_v could be calculated by assuming $\Delta_k = v\%$. The determination of a benchmark would be beneficial in that it would provide the basis for calculating a ceiling for losses stemming from negative deviations from the defined PDs in the individual rating classes.

F is not an absolute value. It can only be used in relation to the defined probabilities of default in the rating classes. In the interest of the portfolio of the lender and its own ratings, care should be taken on its part not to set the probabilities of default in the rating classes too high.

Moreover, migration matrices should be used for the validation of the rating systems when adequate data is available. The above-defined value F is suitable as a basis for the evaluation of migration between rating classes in different periods.

3. Validation process

3.1 Introductory note

On the whole, all phases of the validation process should be in compliance with the DVFA Rating Standards and the requirements of Basel II. At the same time, the validation process should be integrated into the structures and operational organization of the rating agency in accordance with the DVFA Corporate Governance Standards and the Basel II Capital Accords.

The validation process sets forth ex-ante steps with respect to the content and procedure of validation, which should be applied to every model to be validated. The necessary steps are based in each case on the characteristics of the model to be validated. The validation process results in (a) the methodological and (b) the mathematical-statistical validation of the rating model.

The possibilities for methodological and mathematical-statistical validation of rating models are determined by the characteristics of the individual models. For instance, if a rating is performed for only a small number of cases, or if the occurrence of default is very low in relation to the total number of observed cases, or if methodological validation is dependent the scope and complexity of the information about the rating model, there is only limited possibility for a mathematical-statistical validation.

3.2 Validation process: methodological validation

Methodological validation begins with a formal assessment of the content and structure of the rating model with respect to operational plausibility. This is followed by a determination as to the robustness of the applied assumptions, in the context of which, depending on the data set available (pooling data if necessary), the integrity and suitability of the input (rating criteria), the structure of the model and aggregation procedure and the type and adequacy of the output are assessed. If the input, aggregated intermediate input data (throughput) and output are not suitable from an operational point of view to achieve the rating objective, steps must be initiated to modify the rating model.

The second phase of the methodological validation is a determination as to whether or not sufficient input, throughput and output data is available, followed by an evaluation of data quality. A modification of the rating model must be initiated in the case of insufficient data or unsatisfactory data quality. Additionally, the univariate and multivariate robustness of the rating criteria and the rating model must be quantitatively and qualitatively measured.

The results of these two validation phases provide the basis for assessing the methodological validity of the model. Moreover, in the event of methodological validity problems, recommendations for improvement should be drawn up and the rating model modified.

3.3 Validation process: mathematical-statistical validation

Mathematical-statistical validation begins with the determination of the extent to which an adequate number of observations is available, pooling data if necessary. The number of observed cases is adequate if both the absolute number and the distribution between solvent and insolvent (defaulted) borrowers is suitably large. If this is not the case, despite having assessed an adequate number of cases, a modification must be made in data collection. If the actual number of cases is insufficient, a mathematical-statistical validation is not possible.

In the second phase, a determination is made as to whether an adequate number of critical cases (defaults) is available. The calculated weightings are only valid if the number of defaults is adequately large. If the number of defaults is sufficient, the standard mathematical-statistical model is applied to test the quality of the results. In the event of unsatisfactory results, further analyses must be performed. To the extent that these further analyses confirm the initial results, the cause of the inadequate forecast quality must be identified. A modification of the rating model must subsequently be initiated.

In the event that the number of critical cases is inadequate, it should be determined whether or not an external benchmark is available. This benchmark provides a basis for comparison of the cases assessed using the rating model to be validated, allowing a comparison of rating results, and thus an indirect mathematical-statistical validation. When using benchmarks, their suitability must be satisfactorily explained.

3.4 Consequences of methodological and mathematical-statistical validation

Methodological and mathematical-statistical validation is always accompanied by an interpretation of the results and recommendations for action. The recommendations may include the following alternatives: unchanged use of the rating model, recalibration of the model, modification of the model (changing the criteria or the model) or replacement of the model.

3.5 Organization of the validation process

Based on the Minimum requirements for the credit business of credit institutions (*Mindestanforderungen an das Kreditgeschäft der Kreditinstitute – MaK*), responsibility for the development, quality and monitoring of the use of risk classification procedures must not lie with the front office. The validation is to be functionally separate from the development and application of the rating model (segregation of functions). Validation results must be made available to those in charge of (further) developing the rating model. Those charged with validation may only utilize the information provided and the results of the validation as part of their ordinary professional function. Otherwise, they are obliged to hold the information confidential.

Validation activities include initiatives, triggers and bases for the (further) development of rating models. As a rule, rating models should be validated on an ongoing basis. The frequency of validation with respect to existing rating models is determined by (a) the rating intervals of the models to be validated, i.e. validation intervals should be longer than rating intervals and (b) specific validation triggers, e.g. low validity in earlier tests or new circumstances that have not been included in the process previously.

Validation activities and results must be documented in such a way that a knowledgeable third party would be able to form an understanding of all key aspects and the quality of the rating model used within a reasonable period of time.

Part III:

Members of the DVFA-Rating Standards Committee

Chair:

Professor Dr. Jens Leker

Westfälische Wilhelms-Universität Münster

Members:

Christian Albrecht, CCrA

DZ-BANK AG

Prof. Dr. Dr. h. c. Jörg Baetge

Westfälische Wilhelms-Universität Münster

Dr. Oliver Everling

Everling Advisory Services

Dr. Judith Eigermann

Deutsche Bundesbank

Martin Faßbender

Deutsche Bundesbank

Prof. Dr. Josef Fischer

Georg-Simon-Ohm-Fachhochschule Nürnberg

Christian Hesse

Creditreform Rating AG

Dr. Andreas Jerschensky

Deutsche Post AG

Christian Junk, CCrA

NRW BANK

Prof. Dr. Harald Krehl

DATEV eG

Dr. Michael Munsch

Verband der Vereine Creditreform e.V.

Klaus Ott

KPMG Deutsche Treuhand AG

Dieter Pape

Pape & Co

Inge Pawlik

Commerzbank AG

Dr. Karl Eugen Reis, CcrA

Kreissparkasse Limburg

Bodo Richardt

KPMG Deutsche Treuhand GmbH

Prof. Dr. Heinrich Rommelfanger

Johann Wolfgang Goethe-Universität

Prof. Dr. Sören Salomo

Karl-Franzens-Universität Graz

Dr. Peter Schenk

MEAG MUNICH ERGO Asset Management GmbH

Dr. Roland Spahr

BearingPoint GmbH

Dr. Hans-Ulrich Templin, CEFA

Helaba Invest Kapitalanlagegesellschaft mbH

Responsible Members of the DVFA Board of Directors:

Klaus Holschuh

DZ-BANK AG

Dr. Peter Merk

Landesbank Baden-Württemberg

DVFA

DVFA is the Society of Investment Professionals in Germany, founded in 1960. Currently, DVFA has more than 1,100 individual members representing over 400 investment firms, banks, asset managers, consultants and counselling businesses. DVFA assures the credibility of the professionals and the integrity of the market by offering internationally recognised qualification programmes, by providing platforms for professional financial communication, and by effective self-regulation.

Via EFFAS, the umbrella organisation of European Analysts Societies, DVFA offers access to a pan-European network with more than 17,000 investment professionals in 24 nations. Via ACIIA, the Association of Certified International Investment Analysts, DVFA is part of a worldwide network of more than 50,000 investment professionals.

ISBN: 3-928759-04-3
978-3-928759-04-5
(by January 2007)

DVFA
Einsteinstraße 5
63303 Dreieich
Germany

Tel.: +49 (0)6103 - 5833-0
Fax: +49 (0)6103 - 5833-34
Mail: info@dvfa.de
Web: <http://www.dvfa.de>