Database-supported thermal analysis involving automatic evaluation, identification and classification of measurement curves

Alexander Schindler¹ · Claire Strasser¹ · Stefan Schmölzer¹ · Mihal Bodek² · Rafał Seniuta² · Xiayun Wang³

Received: 15 July 2015 / Accepted: 1 September 2015 / Published online: 14 September 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The philosophy of the recently presented computer-based curve recognition and database system for differential scanning calorimetry (DSC) measurements, Identify, is highlighted. This involves autonomous evaluation of measurements by the novel AutoEvaluation software function. A substantial expansion of Identify is furthermore introduced, including data not only from DSC but also from thermogravimetry (TG), dilatometry (DIL) and thermomechanical analysis (TMA) as well as data on the specific heat capacity, \( c_p \), within the same database system. Libraries with more than 1000 measurements and literature data from the fields of ceramics and inorganics, metals and alloys, polymers, organics, food and pharmaceuticals as well as chemical elements are included in the database, which can also be enlarged by user data. Despite of general limitations of database search routines discussed, both AutoEvaluation and Identify considerably simplify the evaluation and interpretation of thermoanalytical measurement curves. Demonstrations are given for the automatic evaluation, identification and classification of several polymers from their DSC and TG curves; this can be used for quality control and failure analysis. Classification of DIL measurements of the linear thermal expansion of alumina samples is also depicted.

Keywords Database · Curve identification · Classification · Quality control · Automatic evaluation

Introduction

In the history of thermal analysis, the consideration of measurement signals originating, for example, from the methods differential scanning calorimetry (DSC), thermogravimetry (TG), dilatometry (DIL) or thermomechanical analysis (TMA) required intensive examination by experienced users in most cases [1, 2]. Effects such as glass transitions or other caloric effects that may occur in a measurement curve first need to be identified visually before they can be evaluated. Nowadays, the evaluation of effects—once identified—is supported by software routines that work according to known DIN or ASTM standards. Moreover, thermoanalytical measurement curves usually require careful interpretation that demands both a certain level of experience and a certain expenditure of time: Measurement curves typically need to be compared with literature or textbook data, which certainly do not exist for every material or for every measurement condition applied—although the amount of information available in materials science that can be searched and found actually via the internet grows continuously [3]. Computer-assisted analysis of thermoanalytical measurements using also the possibilities of relational databases was carried out soon in order to extract, e.g., kinetic parameters [4]. However, the possibility of a computer-based identification of a measurement by means of a comparison with a database—as is well known for spectroscopic techniques such as mass spectrometry or Fourier transform infrared spectroscopy, for example—did not exist for thermoanalytical curves. In 2014, the first computer-based curve recognition and
database system for DSC curves, called Identify, was introduced [5–8]; the history of databases in thermal analysis was reviewed recently [9].

This work is intended to focus on the philosophy of Identify, which also involves a novel, autonomous and automatic evaluation of measurement curves by means of a software function called AutoEvaluation [8]. Identify and AutoEvaluation, which are both part of the NETZSCH Proteus® evaluation software, significantly simplify the evaluation and interpretation of thermoanalytical measurement curves. The scope of this work also includes the introduction of a substantial expansion of Identify implementing not only DSC curves but also data from TG, DIL, TMA and specific heat capacity, \( c_p \), within the same database system. The NETZSCH part of the database has been significantly enlarged and currently contains libraries with more than 1000 entries from the fields of ceramics and inorganics, metals and alloys, polymers, organics, food and pharmaceuticals as well as chemical elements. Furthermore addressed in this work are general limitations of database search routines. Applications of Identify presented generally concern the recognition of “unknown” measurement curves or just particular parts of a measurement curve, the classification of measurements that can be used for quality control and finally, the use of Identify as an archiving system for storing and searching of database entries. The designation “unknown” measurement used in this work means an input measurement for a database search; it is not necessarily a measurement on a completely unknown material.

Philosophy of Identify involving AutoEvaluation

In general, the Identify database system always contains all of the NETZSCH database entries and offers the possibility of overlaying an “unknown” measurement with any database curve—including those of other supported signal types. Users can create libraries containing their own data that can also be shared with several other users at the same time in the computer network.

1. Effect-based approach:
   For DSC and TG signals, Identify uses an effect-based approach [7] and effect-based algorithms for identification of a measurement curve similar to those employed, for example, in modern image recognition software for identifying persons or objects [10]. This approach can be divided into three main tasks:
   (a) Segmentation of a measurement curve: Significant effects such as glass transitions, endo- or exothermic effects in the case of DSC signals or mass changes in the case of TG signals must be identified and distinguished from irrelevant parts of the measurement curve. This task can either be done visually by the user—as was always the case in the past—or by applying AutoEvaluation, which can autonomously find effects in a measurement. One frequent case is the automatic evaluation of the melting effect of metal samples. In the case of polymers, AutoEvaluation searches systematically for any effects present, such as melting of the sample or melting of components such as plasticizers or additives, as well as for glass transitions, crystallization, curing or evaporation. Results from AutoEvaluation, which are not only fast but also reproducible and objective, are clearly helpful for non-experienced users as well as for experienced scientists who might gain a “second opinion.”
   (b) Extraction of the properties of the effects found: Under consideration of known DIN or ASTM standards, properties such as the extrapolated onset temperature or a peak area are evaluated automatically using standard features of the evaluation software.
   (c) Recognition of a measurement curve: All properties of the effects found are taken into account in a calculation of the “similarity value” between the “unknown” measurement and database measurements or literature data (see below) using advanced mathematical algorithms. In general, properties of an effect may concern its temperature position, its magnitude and also its shape.

This effect-based approach has several advantages. It provides results very fast (in situ), its algorithms incorporate intelligence and user experience regarding the significance of effects, and it can be adapted by users to the application: Algorithm types for single- or multi-component samples as well as different parameter setups such as “amorphous” or “crystalline” can be selected that would take into consideration any additional information on the sample that the user may have. Identify is also able to automatically select suitable algorithm settings with respect to the “unknown” measurement.

2. Datapoint-based approach:
   For DIL, TMA and \( c_p \) signals, Identify uses a datapoint-based approach because of fundamental differences compared to DSC or TG signals: In contrast to typical DSC or TG measurements, neither \( \Delta L/L_0 \) signals from DIL and TMA nor \( c_p \) measurement curves exhibit any effects in many cases, but they do
show, for instance, just an increase in the signal as a function of temperature. Effects occurring in DIL, TMA and \( c_p \) curves sometimes do not have a standardized shape and the evaluation of such effects may be difficult. In contrast to DSC signals, the values of \( \Delta L/L_0 \) and \( c_p \) signals are clearly in the scope of interest where values of \( \Delta L/L_0 \) curves can differ by more than four orders of magnitude for different materials. The datapoint-based approach uses the signal data as a function of temperature to calculate the “difference” between two curves. Identify can use three different datapoint-based algorithm types regarding the absolute difference (default for \( c_p \) curves), regarding the slope (default for DIL and TMA curves) or regarding the shape of the curves. The latter allows for recognition of selected parts of a curve exhibiting an effect such as a peak or a step due to phase transformations. The “curve difference” is in case of the algorithm “shape” in dimensionless units between 0 and 100 % as is the case for effect-based algorithms; in case of the algorithm “absolute,” it is in absolute units such as in J g\(^{-1}\)K\(^{-1}\) for \( c_p \) curves; the algorithm “slope” applied for \( \Delta L/L_0 \) curves reveals the unit K\(^{-1}\). All cases yield a hit list as the search result, in which the database entries of selected libraries are sorted according to their similarity or curve difference as compared to the “unknown” measurement. Identify contains not only measurements but also literature data entries without a measurement curve. Such literature data entries contain properties about endo- or exothermic effects, glass transitions and mass changes, but also about the coefficient of linear thermal expansion, \( \alpha \), and the specific heat capacity, \( c_p \)—both at room temperature. Literature data entries already included originate in most cases from references [11–14]. Users can also expand the database with further literature data entries.

In summary, Identify is able to compare “unknown” measurements of any supported type with other such measurements—but also with literature data present in the database as illustrated in Fig. 1. Besides the direct one-on-one comparison between the “unknown” measurement and database entries, there is at the same time also the possibility of classification which assigns the “unknown” measurement to certain predefined classes: Such classes might be material classes (MCs) containing, for example, all measurements obtained for a certain material. Quality classes (QCs) would contain, for instance, only acceptable measurements of a material stemming from “good” parts. A single measurement or literature data entry can, in principle, belong to several classes at the same time. Some exemplary MCs are already included, but users can create classes of their own that incorporate additional user knowledge and expertise into Identify.

Use and limitations of Identify

Methods of thermal analysis such as DSC, TG or DIL are typically applied, for example, for the determination of melting temperatures, the degree of crystallinity, reaction temperatures and enthalpies, the specific heat capacity, the thermal stability of materials or their thermal expansion [1, 2]. Identification of completely unknown samples is certainly not the main purpose of traditional thermal analysis techniques such as DSC, TG or DIL—at least when not coupled to any evolved gas analysis. A more frequent application of, e.g., stand-alone DSC and TG instruments is, however, the detection or confirmation of particular phases where especially chemical compounds such as salts exhibit characteristic DSC peaks due to melting or other phase transformations [15]. Another example is the characterization of materials with regard to quality and failure analysis where DSC and TG allow for the detection, and in some cases, the identification of undesired components or impurities present in a sample [16]. In this case, additional effects are observed in the DSC or TG signals or expected effects are altered in comparison with the pure material.

As shown below, Identify can recognize measurement curves or—when the search temperature range is restricted—just selected parts of a measurement curve via comparison with database measurements. This is helpful and time-saving for interpretation of the measurement with respect to the applications mentioned above. One should, however, keep in mind that recognition of a thermoanalytical measurement curve is not automatically an identification of a material. An absolutely definite identification of materials is sometimes difficult due to the following limitations:

1. Multiple interpretations:
   At first, it is challenging that multiple interpretations of
a measurement curve are sometimes possible. This limitation, also existing in the field of MS and FT-IR spectroscopy, is typical for database search routines. As an example, the DSC curve of the melting effect of chocolate may look similar to that of the metal gallium with a melting temperature of 29.8 °C. Even the linear thermal expansion coefficient—which differs by more than four orders of magnitude for different materials—is not a unique identifier for a material, nor is the specific heat capacity. This problem can be reduced by selecting appropriate search libraries containing only entries that are really a possibility, for example, for the identification of gallium, just those from the fields of metals and alloys. It is additionally helpful that the Identify search result is not only a single suggestion, but a hit list from which several hits can be considered—also from different measurement types. In general, a material identification is the more definite the more characteristic the curves are. And this is the case if the curves have additionally to their values also characteristic effects to consider where more effects in one curve are by trend more helpful for an identification. In future work, combined signals such as STA referring to simultaneous TG-DSC or TG-c-DTA signals will be included where both caloric effects and mass changes of a sample measured are taken into account. This will certainly lead to a more definite material identification. As mentioned above, a further coupling with evolved gas analysis techniques such as MS, FT-IR or GC–MS of course enhances the possibilities regarding material identification significantly [17].

2. Dependence on the measurement conditions:

Another general obstacle for a material identification via thermal analysis is formed by the dependence of the measurement curves on the measurement conditions. For example, higher values of the heating rate and also of the sample mass would tend to shift
caloric effects observed in the DSC curve to higher temperatures. Consequently, the similarity value between DSC measurements on exactly the same material but carried out with different sample masses or heating rates would be lower than 100 %. The effect-based algorithms of Identify are, however, elaborated in such a way that variation of the heating rate or the sample mass by a factor of two in many cases still results in high similarity values between such measurements, allowing Identify to produce satisfying search results. A user can furthermore add curves measured on the same material but with different conditions to the database; those measurements will be helpful for an identification of that material in the future. The measurements included into the database search can additionally be filtered according to their measurement conditions.

3. The database does not yet contain a similar measurement or literature data:

It might furthermore occur that a similar database entry that could lead to a satisfying search result does not yet exist in the database. In such a case, users can build up their own application-specific libraries not only for identification of measurements in the future but also for qualitative and quantitative comparisons of new measurements on a material—for instance from a new batch—within the database. This leads to the application in the field of quality control and failure analysis (see below).

Recognition of measurement curves

The following examples should demonstrate the capability of Identify to recognize measurement curves. Identify was, for instance, applied to a DSC measurement of an “unknown” polymer sample (solid lines) including results of AutoEvaluation in comparison with the TG curves of PVC-P (polyvinyl chloride with plasticizer, dashed line) and CR (chloroprene rubber, dotted line). The TG curves of PVC-P and CR are selected Identify search results (see Fig. 3b). Fig. 3a

(a) Temperature-dependent mass change—or TG curve—and the corresponding rate of mass change, DTG, of an “unknown” polymer sample (solid lines) including results of AutoEvaluation in comparison with the TG curves of PVC-P (polyvinyl chloride with plasticizer, dashed line) and CR (chloroprene rubber, dotted line). The TG curves of PVC-P and CR are selected Identify search results (see Fig. 3b). Fig. 3b Results of Identify (hit list) for the “unknown” TG curve shown in Fig. 3a
It has to be emphasized that all DSC measurements shown in Fig. 2a and all database curves included in the search are so-called second heating runs at a heating rate of 10 K min\(^{-1}\) performed after a first heating and a cooling, both at 10 K min\(^{-1}\). The sample masses were in the typical range of 10 mg, aluminum crucibles with pierced lids were used, and measurements were taken in a nitrogen atmosphere.

Further examples for the recognition of DSC curves including mixtures with different ratios of PE (polyethylene) and PP (polypropylene) as well as mixtures of PC (polycarbonate) and PBT (polybutylene terephthalate) are reported in Ref. [9].

Figure 3a, b depicts the application of Identify to a thermogravimetric measurement, reflected by the TG curve of an “unknown” polymer sample. AutoEvaluation revealed two mass-loss steps of 72.2 and 20.1 % with maximum mass-loss rates at 569 K (=296°C) and 741 K (=468°C). A search library containing 66 TG measurements on typical polymers [11] was selected, and standard algorithm settings were used assuming a single-component sample. The Identify search results shown in Fig. 3b are PVC-P (polyvinyl chloride with plasticizer) and PVC-U (polyvinyl chloride without plasticizer) with similarities of 86 and 84 % as hits no. 1 and 2. Hit no. 3, CR (chloroprene rubber), has a similarity of only 52 %, which is mostly due to the fact that the first mass-loss step visible in the TG curve of CR (see Fig. 3a) occurs at about 75 K higher temperature than the curve measured on the “unknown” polymer.

These results demonstrate that Identify is well able to differentiate between various thermogravimetric measurements.
on different materials exhibiting, for example, two mass-loss steps, respectively. Measurements that—in contrast to the measurement on the “unknown” polymer of Fig. 3a—include just one or more than two mass-loss steps, are strongly discriminated against, with similarities of less than about 10% when standard algorithm settings are applied.

All thermogravimetric curves shown in Fig. 3a as well as all database curves included in the search were measured at a heating rate of 10 K min⁻¹ in a dynamic nitrogen atmosphere. The sample masses were about 10 mg, and open alumina crucibles were used for the measurements.

Quality control and failure analysis

Characterization and classification of recycled polyamides discussed earlier showed how Identify is able to distinguish polyamide samples with respect to the temperature position of the glass transition and the melting effect or the presence of additional caloric effects present in the DSC signal [7, 18]. Such a situation is shown in Fig. 4a where a sample, presumably a blend of PE (polyethylene) and PA (polyamide), was measured by DSC and analyzed by means of Identify. The blend sample with a mass of 5.19 mg was heated up at 10 K min⁻¹ in a dynamic nitrogen atmosphere; the second heating curve is shown in Fig. 4a. The algorithm settings were for multi-component samples, and furthermore, the size and the shape of the effects found were disregarded. Again, just search libraries for polymers were selected containing second heating curves measured at 10 K min⁻¹ in a dynamic nitrogen atmosphere on samples with a typical mass of 10 mg; the search temperature range was restricted to the range between 333 K (=60 °C) and 543 K (=270 °C). AutoEvaluation of the DSC measurement on the polymer blend

![Fig. 5](image-url)
revealed two endothermic effects at peak temperatures of 383 K (=110 °C) and 494 K (=221 °C) which are due to the melting of the PE and PA components. This is confirmed by the Identify search results (see Fig. 4a, b). The message “QC: FAIL!” is automatically triggered by the software in this example (see text). For the example depicted in Fig. 4, the quality control function of Identify was activated, which compares the “unknown” measurement with a selected class of database entries. In this case, the “PA6.x_semi-cryst.” class, containing measurements and literature data for pure semicrystalline PA6.x samples (no blends), was selected. The similarity between the measurement on the PE/PA blend and the “PA6.x_semi-cryst.” class is only 26 % (see Fig. 4b)—mainly due to the additional melting effect of PE at 383 K peak temperature—and thus much lower than the user-defined threshold of 95 %. This was automatically leading to the message “QC: FAIL!” (and not “QC: PASS!”; see example below) visible in Fig. 4a, which would be part of a customizable report of the Identify results. In summary, this example demonstrates that Identify is able to recognize and distinguish the common engineering thermoplastic polyamide of course from the cheaper commodity thermoplastic polyethylene and also from PE/PA blends and vice versa; a quantitative analysis of unknown blends could also be carried out, where exemplary database measurements with different mixing ratios would be required [9].

Another example concerns quality control by means of dilatometer (DIL) measurements. The linear thermal expansion of several alumina (Al₂O₃) samples with a length of 25.0 mm was measured with a dilatometer at a heating rate of 5 K min⁻¹ up to 1280 K (=1007 °C) in a dynamic helium atmosphere. The second heating curves “QC_Al₂O₃_cw01,” “QC_Al₂O₃_cw02”—that were considered to fulfill the quality criteria—were added to the database and grouped to the class “QC_Al₂O₃_passed.” Another DIL measurement “QC_Al₂O₃_cw03” on a further alumina sample was investigated by applying Identify (see Fig. 5a, b) where the standard algorithm type for DIL curves, “slope,” was applied (see datapoint-based approach described
Database-supported thermal analysis involving automatic evaluation, identification and archiving system for measurements and literature data. A measurement can be added to one or several libraries and particular measurements and literature data can also be searched for and found within the database by applying, for example, alphabetical filtering. Evaluated measurements found can then be taken into consideration; the evaluations can be edited and these edits can be applied to the database. Searching and finding of database entries are particularly useful for gathering information about what behavior and what measurement results can be expected from a sample that has not yet been measured by the user. It is also helpful prior to a measurement because suitable measurement conditions can be seen in the database entry.

Conclusions

The philosophy of the first computer-based curve recognition and database system for thermal analysis, Identify, was discussed. This included effect- and datapoint-based approaches for recognition of measurement curves of different signal types. Furthermore, a novel software function called AutoEvaluation was introduced which allows for autonomous and automatic evaluation of DSC and TG measurement curves for the first time. AutoEvaluation can be applied by Identify but can also be used independently. Both AutoEvaluation and Identify considerably simplify the evaluation and interpretation of thermoanalytical measurement curves. General limitations of database search routines were also addressed, such as the challenge of multiple interpretations of a measurement curve.

Identify—introduced in 2014 for DSC measurements on polymers—was substantially expanded upon, also implementing data from TG, DIL, TMA and the specific heat capacity, c_p, within the same database system. The latter was also significantly expanded, currently containing libraries with more than 1000 measurements and literature data from the fields of ceramics and inorganics, metals and alloys, polymers, organics, food and pharmaceuticals as well as chemical elements. Users can create libraries with data of their own that can be shared with other users at the same time in the computer network.

The use of Identify as presented in this work is, on the one hand, for the recognition of “unknown” measurement curves or just particular parts of a measurement curve in order to identify, for example, materials or particular phases of a material. It was shown how AutoEvaluation and Identify automatically evaluated the DSC and the TG measurements of two polymers, most probably identified as EVA (polyethylene-co-vinyl acetate) in case of the DSC curve and PVC (polyvinyl chloride) in case of the TG measurement. On the other hand, Identify can be used for classification and thus for quality control and failure analysis under consideration of user-defined quality criteria: One application example using DSC measurements demonstrated that PA (polyamide) can automatically be distinguished from PE (polyethylene) and from PE/PA blends and vice versa. Another example concerned the classification of measurements of the linear thermal
expansion of several alumina samples using DIL. The general use of Identify as an archiving system for storing and searching of measurements and literature data was also discussed.

One of the future projects is that also combined signals such as STA referring to simultaneous TG-DSC or TG-c-\textit{DTA}® signals will be included into the database system.

\textbf{Open Access} This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

\textbf{References}

1. Hemminger WF, Cammenga HK. Methoden der thermischen analyse. Heidelberg: Springer; 1989.
2. Ehrenstein GW, Riedel G, Trawiel P. Thermal analysis of plastics: theory and practice. New York: Hanser Gardner Publications; 2004.
3. Fiala J, Sestak J. Databases in materials science—contemporary state and future. J Therm Anal Calorim. 2000;60:1101–10.
4. Reich L, Patel SH. Use of databases in thermal analysis. Thermochim Acta. 1993;222:85–90.
5. Schindler A. Materialerkennung und Qualitätskontrolle: Auswerten, identifizieren und interpretieren. Plastverarbeiter. 2014;01:30–2.
6. Schindler A. \textit{AutoEvaluation}—automatische Auswertung von DSC-Kurven. \textit{Identify}—das neue DSC-Kurvenerkennungs- und Datenbanksystem. Presentation during annual symposium of the AK Thermophysik (part of the GEFTA). 2014; March 17–18.
7. Schindler A. Automatic evaluation and identification of DSC curves. Plast Eng. 2014; April 2. http://www.plasticsengineering.org/ProductFocus/productfocus.aspx?ItemNumber=20498.
8. Moukhina E, Schindler A. Automatic evaluation and identification of DSC curves. Presentation during international GEFTA symposium “thermal analysis and calorimetry in industry and research”. 16–19 Sept 2014.
9. Fueglein E, Kaisersberger E. About the development of databases in thermal analysis. J Therm Anal Calorim. 2015;. doi:10.1007/s10973-014-4381-3.
10. Moeslund ThB, Hilton A, Krüger V, Sigal L, editors. Visual analysis of humans—looking at people. London: Springer; 2011.
11. Thermal Properties of Elements, Polymers, Alloys, Ceramics. www.netzsch.com/TPoE, www.netzsch.com/TPoP, www.netzsch.com/TPoA, www.netzsch.com/TPoC.
12. Touloukian YS, Powell RW, Ho CY, Klemens PG, Kirby RK, Taylor RE, Lee TYR, Desai PD, Nicolaou MC, Buyco EH. Thermophysical properties of matter. New York: IFI/Plenum; 1970.
13. Liptay G, editor. Atlas of thermoanalytical curves (TG-, DTG-, DTA-curves measured simultaneously). Budapest: Akademiai Kiado; 1976.
14. D’Ans J, Lax E. Taschenbuch für Chemiker und Physiker. Berlin: Springer; 2013.
15. Klassen VK, Ermolenko EP. Problem of impurity of salts of alkali metals in cement raw materials. Middle-East J Sci Res. 2013;17(8):1130–7.
16. Schmölzer S. Temperature taken. Kunstst Int. 2009;10:55–7.
17. Schindler A, Neumann G, Rager A, Füglein E, Denner T, Blum J. A novel direct coupling of simultaneous thermal analysis (STA) and Fourier transform-infrared (FT-IR) spectroscopy. J Therm Anal Calorim. 2013;113:1091–102.
18. Fueglein E. About the use of \textit{Identify}—a thermoanalytical database—for characterization and classification of recycled polyamides. J Therm Anal Calorim. 2015;. doi:10.1007/s10973-015-4583-3.