Time Series Analysis and Manipulation in Industrial Energy Systems

Philip Voll[1*], Stefan Kirschbaum[2] and Gregor Wrobel[2]

[1] Chair of Technical Thermodynamics, RWTH Aachen University, Germany
voll@ltt.rwth-aachen.de

[2] Society for the Promotion of Applied Computer Science, Germany
kirschbaum@gfai.de, wrobel@gfai.de

ABSTRACT: Industrial energy systems are manifold and often highly integrated and thus complex to analyze. For the analysis and optimization of these systems time series evaluations are vastly useful and recurrent tasks, e.g. to design peak load management systems or to evaluate the possibilities of an integration of cogeneration technologies in existing energy systems. The necessary steps an energy consultant has to perform to undertake typical time series analyses and manipulations are generally costly in terms of labor and time. This paper presents a software tool, the Engineering Time Series Analyzer, short η, which is tailored to energy consultants acting in the analysis and optimization of industrial energy systems. It covers a multitude of typical tasks for the import, the conditioning, the analysis and the manipulation of time series. Starting with a classification of different types of time series, this paper presents a work flow for time series processing in the field of industrial energy systems. As an example application a waste processing plant has been analyzed to evaluate the developed software with regard to its relevance in practical applications.

Keywords: industrial energy systems, time series classification, time series analysis and manipulation, load profile synthesis, design of peak load management systems, waste recycling

1. Introduction

Industrial energy systems are manifold and often highly integrated and thus complex to analyze. The term industrial energy system covers the energy related aspects of the production process as well as its corresponding energy supply system. For the analysis and optimization of these systems time series evaluations are vastly useful and recurrent tasks, e.g. to design peak load management systems or to evaluate the possibilities of an integration of cogeneration technologies in existing energy systems. The necessity of time series processing in the analysis and optimization of industrial energy systems becomes especially obvious if energy demand load profiles are considered: since these profiles are highly dependent on the economic growth of a production business and in particular on changes of the production process, these time series have to be recorded and reviewed for any significant change in the production process.

The necessary steps an energy consultant has to perform to undertake typical time series analyses and manipulations are generally costly in terms of labor and time. While a special-contract customer's total electricity demand profile can be obtained directly from the energy provider, the acquisition of heat, steam and cooling energy related data often has to be realized performing measurements in the process. In cases where these measurements are not possible, the energy consultant is reliant on the synthesis of load profiles from key data or the use of standard load profiles which have to be adapted for the specific object of investigation.

Existing software tools are mostly based on data warehouse solutions and generally not suitable for the particular purposes of energy consultants. This is because of their objective targets, application scopes and license fees. Hence these consultants typically use ill-suited do-it-yourself solutions based on popular spreadsheet programs.

2. Approach

This paper presents a software tool for the analysis and manipulation of time series in energy systems engineering. The **E**ngineering **T**ime Series **A**nalyzer, short **η**, is tailored to energy consultants acting in the analysis and optimization of industrial energy systems. Harkening back to available algorithms from statistics and numerics, e.g. from [1] and [2], this paper focuses on the design of a generally applicable work flow for time series processing in energy systems engineering. Therefore **η** is based on a holistic approach that ranges from the import of measured values from data loggers to the export to typical file formats used in engineering applications. The presented software tool covers a multitude of typical tasks for the import, the conditioning, the analysis and the manipulation of time series. The general approach makes it possible to use **η** in any technical field where time series analysis and manipulation are central tasks in the work flow. **η** is embedded in the TOP-Energy framework, which is described in the following part.

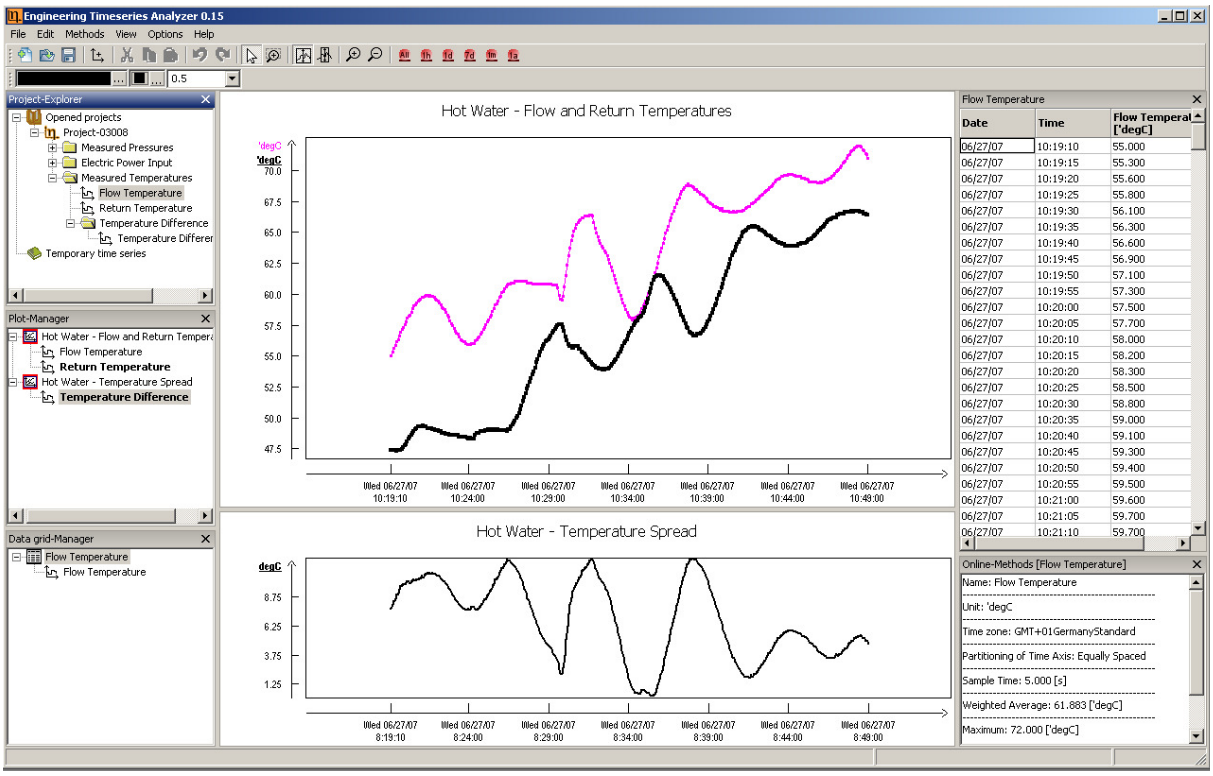

Figure 1: Engineering Time Series Analyzer GUI.

2.1 TOP-Energy

Over the last years the Toolkit for Optimization of Industrial Energy Systems (TOP-Energy) [3], [4] has been developed. This software package supports energy consultants in an efficient overall analysis and optimization of industrial energy systems. The software features modules for the typical tasks an energy consultant has to deal with, such as modules for simulation, optimization and economic feasibility studies. With its graphical user interface (GUI) TOP-Energy inherits the typical windows look & feel, which is a necessary requirement for user-friendly and effective operating software in engineering applications. Moreover, its modular design allows for arbitrary expansions in its applicability. In [5] and [6] it was shown that this concept can be applied successfully to the analysis and optimization of production processes and their related energy supply systems.

While tools like the Aspen software package [7] for process engineering applications or EBSILON [8] for the optimization of energy supply systems lack of appropriate methods to cover the typical tasks in the field of time series processing, software tools like the R project for statistical computing [9] are exclusively designed for this task. In contrast to this, **η** caters for a stand-alone use in the daily work of energy consultants as well as for the connection with other TOP-Energy modules.

2.2 Engineering Time Series Analyzer

The presented software tool is capable of illustrating time series in plot windows as well as in data grids. From there simple manipulations like the deleting or moving of single or multiple data points can be performed by simple mouse click interactions. Calculations for more complex methods are realized using algorithms that are integrated via a generic interface. This allows for arbitrary modifications and expansions of the available method repertoire. The single methods are executed by a typical windows menu structure. Figure 1 depicts the **η**–GUI. In the center two measured temperature curves are illustrated in one *plot window*, while in a second plot window the difference of these time series is plotted. On the right side of Figure 1 the corresponding data is listed in the *data grid*. The *property window* below the data grid provides the user with more information about the time series, such as analytical values like the minimum, the maximum or the integral value. The *project-explorer*, the *plot-manager* and the *data grid-manager* on the left side of the illustration manage the representation of time series data in plot windows, data grids and property windows.

3. Classification of Time Series

Time series are collections of measured values that have been recorded at certain points in time. The recording of measurement signals is either done at fixed time intervals or only at specific points in time. Due to this sampling, the generated time series data is never continuous in a mathematical sense, but always discrete with finite distances between two consecutive samples. So far, there is no unique classification of time series, compare e.g. [2], [10] and [11]. Hence, in this paper a classification is proposed that caters for time series applications in energy systems engineering. The following types of time series are distinguished: *equally spaced* and *inequally spaced*, *momentary* and *interval*, *continuous* and *discrete*, as well as *concrete* and *abstract* time series.

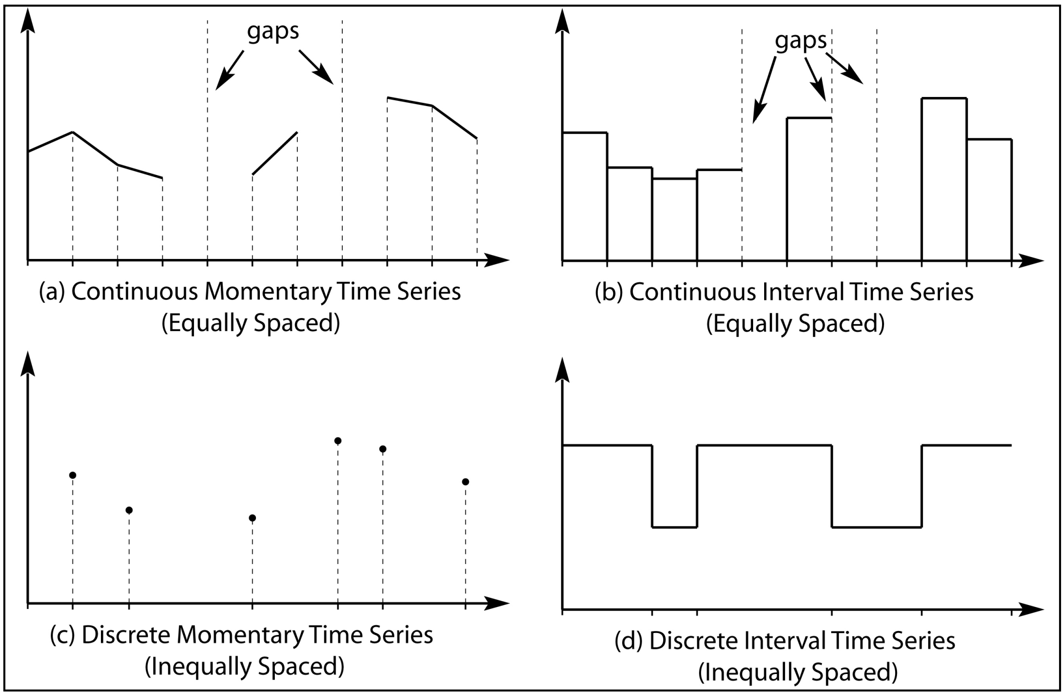

Figure 2: Schematic illustration of four different time series types.

3.1 Equally Spaced Time Series vs. Inequally Spaced Time Series

Time series are equally spaced or inequally spaced. The time axis of equally spaced time series represents a fixed time grid with constant time intervals. In contrast to this, the time intervals in inequally spaced time series have irregular ranges. Equally spaced time series are generated when a signal is recorded with a constant sampling rate. On the other hand, inequally spaced time series are generated when, e.g., the recording is performed manually only at certain points in time. If measurement errors occur while recording an equally spaced time series, the generated time series still will be equally spaced, but it will contain data points with invalid ordinate values.

3.2 Momentary Time Series vs. Interval Time Series

In momentary time series the ordinate value of each data point is valid only for the corresponding point in time. In contrast to this, an interval time series contains ordinate values, that are valid for particular time intervals, e.g. the time span between two consecutive data points.

3.3 Continuous Time Series vs. Discrete Time Series

A time series is said to be continuous if observations of the represented phenomenon can be taken continuously with regard to the time. A time series is said to be discrete if observations of the represented phenomenon can be taken only at specific times. In most cases discrete time series are inequally spaced.

Figures 2a and 2b depict equally spaced continuous time series. While Figure 2a shows a continuous momentary time series, e.g. a temperature curve, Figure 2b represents the plot of a continuous interval time series, e.g. the power input of an electric engine averaged over short time intervals. To visualize the run of the temperature curve between the available samples linear interpolation is used. Since each sample of the interval time series keeps its value up to the next sample, a bar chart is used to visualize the run of this curve. Both time series have gaps in between the available data. Gaps in continuous time series can be filled with calculated substitute values. Figures 2c and 2d represent plots of inequally spaced discrete time series. Figure 2c illustrates a discrete momentary time series, e.g. investments of a business company at specific points in time. Since the observed phenomenon can't be quantified in between the available time series data, if data points are missing, there is no way in calculating reasonable substitute values for these samples. Figure 2d shows a discrete interval time series, e.g. the switching of a digital controller. Since each sample value keeps its validity up to the next sample value, it is possible to describe the phenomenon's behavior in between the available time series data. To visualize the run of the curve a histogram is used.

The type of time series has an important influence on the application of certain analysis and manipulation methods. Consider the integral values of all four time series depicted in Figure 2. As indicated, the areas under the plotted curves are calculated in a different way for each time series. Moreover, for the discrete momentary times series it is not possible at all to calculate an integral value.

3.4 Concrete Time Series vs. Abstract Time Series

Concrete time series give information about concrete points in time, e.g. the recorded hot water temperature curve of a heating system on January 27, 2009 at 4:00 pm. In contrast to this, abstract time series only give information about the time spans in between the single data points, e.g. the sorted load duration curve of an industrial production process, or the typical load profile of an office building's heating system on a typical winter's working day.

4. Work Flow for Time Series Analysis and Manipulation

Energy consultants need to get profound knowledge of the analyzed processes. When it comes to the characterization of a machine's operating behavior, necessary information is often available in form of time series, e.g. measured temperature and volume flow curves for the characterization of an absorption chiller. Figure 3 schematically illustrates the work flow for the analysis and manipulation of time series with **η**. It depicts the use of **η** as stand-alone tool as well as its interaction with other TOP-Energy modules, .e.g. to process simulation results for an economic feasibility study. In the following the focus is on the use of **η** as stand-alone tool for time series evaluations.

When time series are measured with different measurement instruments and recorded with different data loggers again, these time series will generally be incompatible in their data formats. The presented tool features an *import method*, which provides a basis for a simultaneous processing of these time series in one tool. Therefore files with arbitrary ASCII- or Excel-formats will be converted to a unique binary data format. In general, data of these *raw time series* is characterized by deficiencies like invalid or missing data points, or the neglect of daylight saving time shifts. Most analysis and manipulation methods are applicable only to well-defined equally spaced time series. Thus, imported raw time series have to be conditioned by appropriate *conditioning methods*. On the other hand, time series that are generated using the *synthesis method* don't have the deficiencies mentioned before and can therefore be categorized instantly as *conditioned time series*. The application of further methods enables a detailed investigation of the considered time series. Finally, an *export method* makes it possible to export time series data to a couple of predefined data formats.

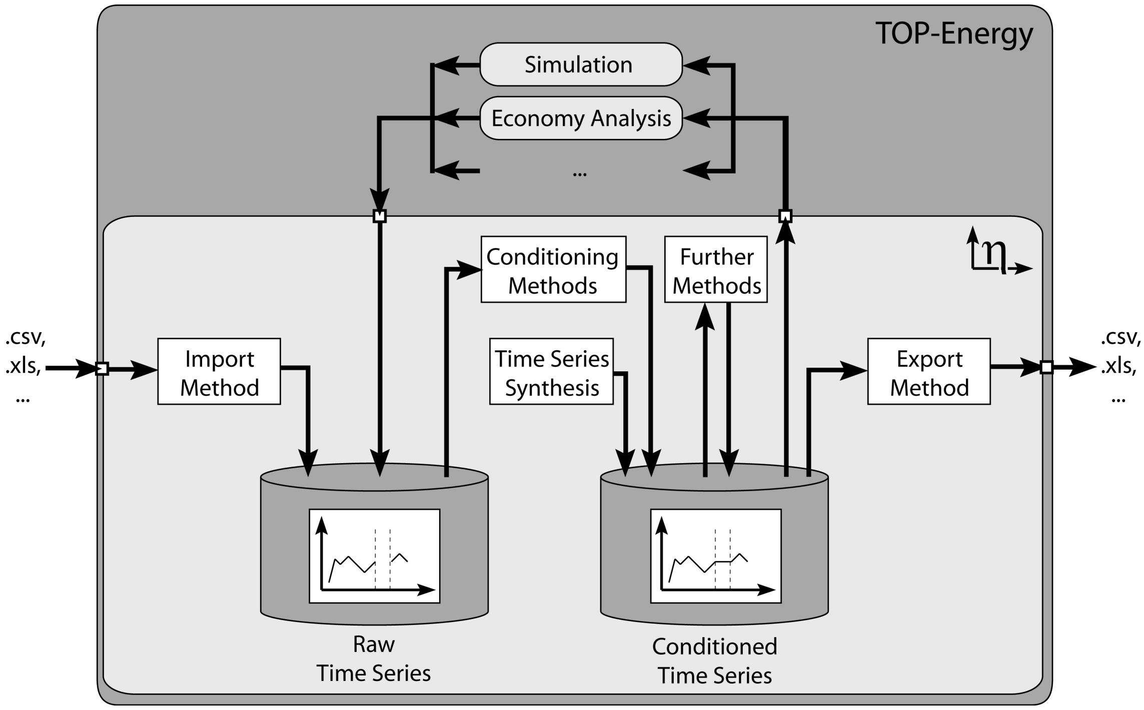

Figure 3: Schematic illustration of the work flow with **η**.

4.1 Import and Export

Time series raw data is recorded using measurement instruments together with appropriate data loggers. Data loggers typically save the time series data to ASCII- or Excel-files. Since **η** uses a binary time series data format, which enables for engineering calculations taking into account numbers with physical units, this raw time series data has to be converted to the **η**–file format. Using a chart based GUI the user can easily predefine import profiles with information about the arrangement of the recorded data in the files to be imported. For some typical formats the method offers predefined import profiles. Moreover, the export method makes it possible to save time series data to ASCII- and Excel-files. In addition to that, this method also caters for the generation of vector and pixel graphics from the plot window views.

4.2 Conditioning

Time series manipulation and analysis methods are generally applicable only to well-defined equally spaced time series. Thus, in a first step imported raw time series have to be conditioned. This process covers the detection and cleanup of invalid data points as well as the calculation of substitute values for missing data points in order to generate equally spaced time series. Depending on the type of time series, different approaches for the calculation of substitute values are applicable, compare part 3. The calculation of substitute values using interpolation or down sampling methods is referred to as *normalization*. In order to allow for a simultaneous application of processing methods to more than only one time series, these time series have to be *synchronized*. Synchronization means that all time series are normalized in order to get equally spaced time series with identical start, end and samples times. While normalization and synchronization methods are applicable only to continuous time series, there are other approaches that go with any type of time series, e.g. the generation of substitute values by adopting values of preceding data points. Others again rely on the adoption of values from time series from the previous period under consideration.

4.3 Analysis and Manipulation

The repertoire of time series methods offers support for most recurrent standard tasks in time series processing. It ranges from simple methods like scaling and moving (with regard to the time and number axis) to more complex methods like integration and differentiation in order to calculate energy valued time series from power valued time series (or the other way round), smoothing in order to remove noise from time series or sorting in order to generate sorted load duration curves. However, many tasks in the field of time series analysis and manipulation can be solved using the *time series calculator*. The calculator makes it possible to perform computations like the summation and subtraction of time series, e.g. to calculate total respectively individual energy consumption curves for single or multiple machines of an industrial energy system. Further calculation methods enable to calculate parameters such as the specific energy demand by dividing a total energy demand time series by its corresponding product quantity time series. The calculator makes it futile to implement single methods for all standard tasks, since the user can create just about every method on his own using the calculator.

5. Example: Analysis of a Waste Recycling Plant

The developed software was evaluated with regard to its relevance in practical applications. Therefore, a waste recycling plant has been analyzed. In a first step the main energy consumers were identified. Moreover, a peak load management system was designed in order to reduce the level of peak loads that occur during the recycling process.

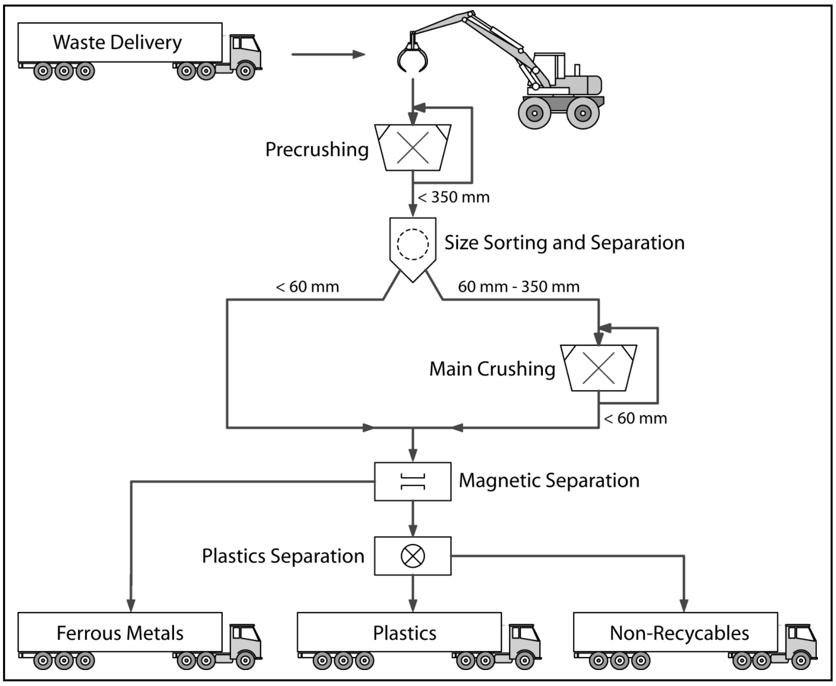

Figure 4: Simplified process flow sheet of the waste recycling plant.

Figure 4 depicts the simplified process flow sheet of the analyzed recycling process. The production proceeds non-stop from Mondays 6am to Saturdays 4am in a three shift operation. In the first process step *"Precrushing"* a waste-crusher reduces the delivered waste to pieces with sizes smaller than 350 mm. In the next step *"Size Sorting and Separation"* three mechanical separators sort and separate pieces minor than 60 mm from the main flow. The main flow is fed to the *"Main Crushing"* section, where four waster-crushers reduce the sizes of all remaining pieces to sizes of 60 mm as well. Finally, the merged waste flow is split into the components ferrous metals, plastics and non-recyclables. While the selection of ferrous metals is realized in the process step *"Magnetic Separation"* by two magnetic separators, the partitioning of plastics and non-recyclables is performed in the *"Plastics Separation"* step by two compressed air separators. All mentioned recycling machines are connected by electric conveyor belts. All of the mentioned machines are powered with electrical energy. Only a few machines use compressed air as auxiliary power. This compressed air is supplied by electronic onsite compressors again. In addition to the mentioned machines, the recycling plant also includes a dust extraction plant, and an administration building which incorporates sanitary facilities.

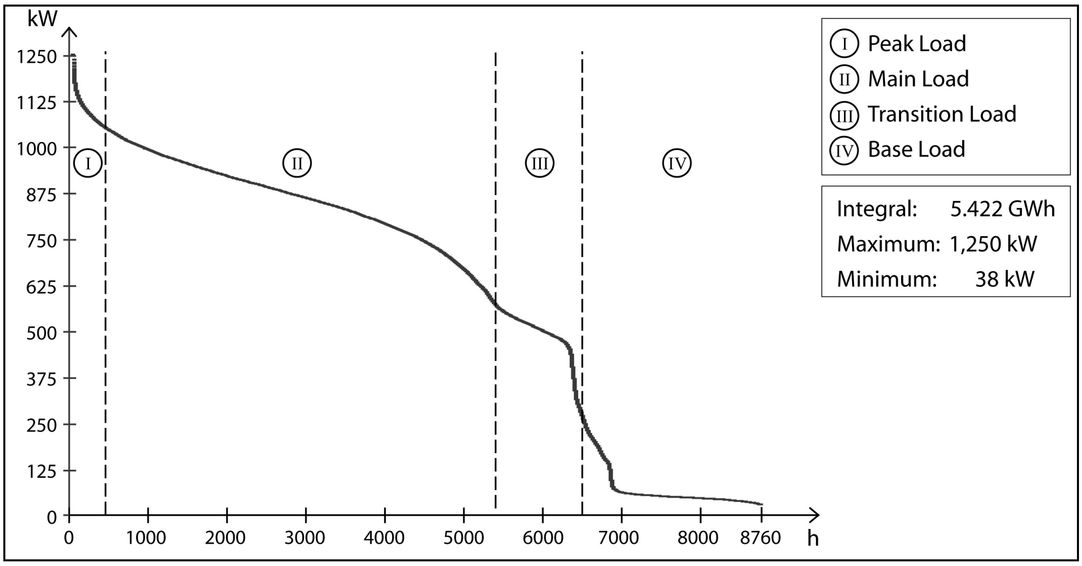

Figure 5: Annual load duration curve.

The available production data was limited to the cumulative electricity load profile of all machines averaged over 15 minute intervals, which was retrieved from the local energy supply company. The production data neither included hours of operation of the single machines nor measured values of their electrical power inputs. The load profile was imported as a continuous interval time series and conditioned for further analysis using the normalization method. Figure 5 depicts the synthesized annual load duration curve of the cumulative electricity usage, which can be divided into four domains: the peak load, the main load, the transition load and the base load domain. In the 434 hours of the peak load domain the load exceeds 1,150 kW. In the main load domain, which exists in about 4,900 hours per year, the load ranges between about 600 kW and 1,150 kW. In between the main load and the base load domain lies the transition domain. This domain represents 1,200 hours of a year's operating time, in which the load goes considerably below the power input of the main load domain, but still exceeds the base load level of about 40 kW to 60 kW. The latter only occurs on weekends and holidays. The maximum power input adds up to 1,250 kW, the minimum value amounts to 38 kW. The total usage of electrical energy sums up to 5.422 GWh.

5.1 Identification of Main Consumers

In order to identify the main energy consumers of the recycling plant, for each machine type the power input of one particular machine was measured over 8 hours of a typical working day. These measurements were used in order to synthesize a load profile for the recycling plant. Therefore, in a first step the imported momentary time series of the high-resolution measurements were down sampled to 15 minute interval time series as to adapt them to the real load profile, which was retrieved from the energy supply company. The electrical energy use of the administration building was not measured over a whole day, but by random measurements over several days. The evaluated data is listed in Table 1. For each consumer three values are given: an average value, which represents the power input averaged over the measured interval of eight hours, and the minimum and maximum values that occurred during the measurements. Furthermore Table 1 contains information about how many machines of the listed type exist in the considered plant.

Consumer (Quantity)	Average / kW	Minimum / kW	Maximum / kW
Mechanical Separator (3x)	47.0	46.4	47.3
Magnetic Separator (2x)	43.0	42.2	43.4
Plastics Separator (2x)	23.7	23.3	24.4
Conveyor Belt, Installed Load: 17 kW (Total Installed Load: 166 kW)	17.0	16.8	17.1
Dust Extraction Plant (1x)	72.0	70.7	72.7
Administration Building, by day (1x)	130.0	129.4	130.8
Administration Building, at night (1x)	20.0	19.4	20.3
Compressor (1x)	21.6	17.5	45.0
Waste-Crusher, Precrushing (1x)	56.5	20.6	141.3
Waster-Crusher, Main Crushing (4x)	45.5	28.0	104.2

Table 1: Measured power inputs (January 2009).

Comparing the listed data it turns out that the consumers can be separated into two groups: one group with energy consumers that mainly run in steady state operation (separators, conveyor belts, dust extraction plant, administration building), and another group with machines, whose power inputs vary considerably depending on the load (compressor, waster-crushers). In order to get a good idea about the distribution of electrical energy on the single machines, the occurrence of peak loads was neglected as to synthesize an averaged load profile. Besides, only production key data such as operating times and the number of installed machines was taken into account.

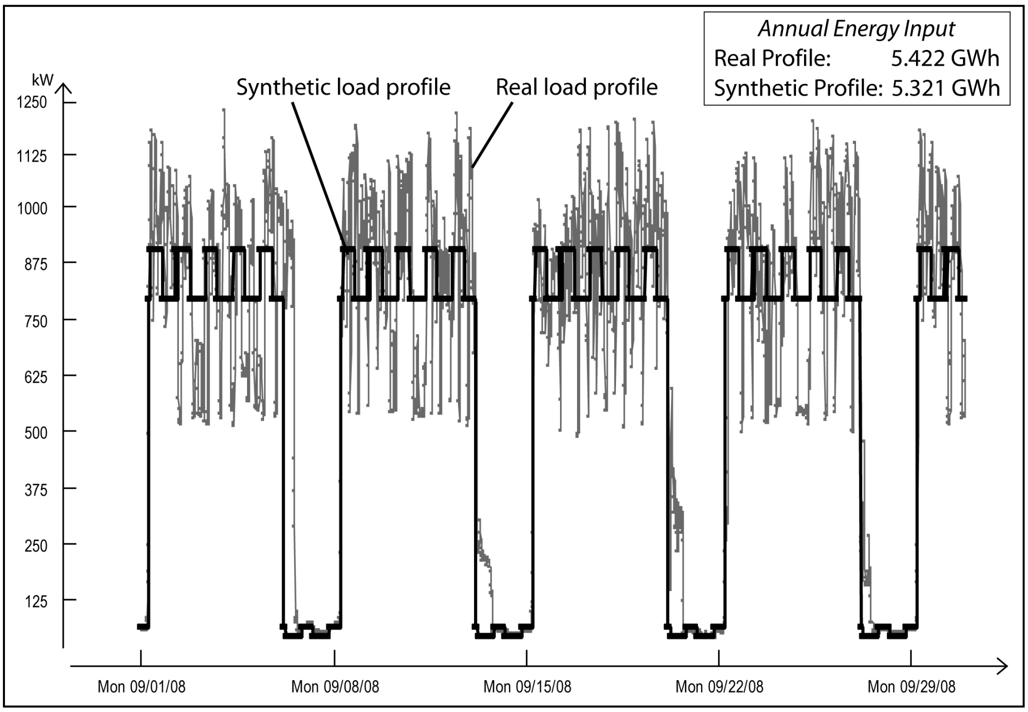

Figure 6: Comparison of real with synthetic load profile (September 2008).

In Figure 6 the real load profile of September 2008 is opposed to the averaged synthetic load profile from the same month. While the real power input values range from about 500 kW to peak values of about 1,200 kW during working days, and between 40 kW and 60 kW on weekends, the synthetic load profile only takes four states: during working days depending on the daytime either 792.5 kW or 902.5 kW, and on weekends, again depending on the daytime, either 40 kW or 60 kW. The integral value of the synthetic annual load profile adds up to 5.321 GWh. The deviation from the real annual energy demand of 5.422 GWh is less than 2%. Thus, it is concluded that the synthesized load profile represents the energy distribution in the recycling plant rather accurately.

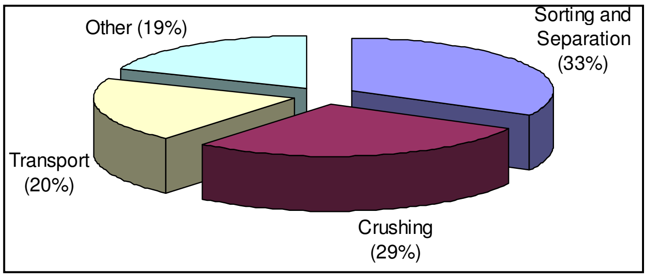

Diagram 1: Distribution of electrical energy.

Diagram 1 illustrates the distribution of electrical energy to all energy consumers. The sorting and separation machines just like the waste-crushers use about one third of the total electricity input. One fifth is expended to power the conveyor belts, while the remaining 19% are used in the dust extraction, the compressed air supply as well as in the administration and sanitary facilities section.

5.2 Design of a Peak Load Management System

The annual load duration curve in Figure 5 shows that there occur distinctive peak loads in more than 400 hours of the annual recycling process. Investigating the given load profile, it turns out that these peak loads don't occur at particular times such as start-ups of the crushers, but randomly distributed over the time. The measured power inputs listed in Table 1 reveal the simultaneous full load operation of all five waste-crushers as possible reason for the occurrence of these peaks. In order to validate this assumption, the averaged power input of all five crushers is subtracted from the synthesized load profile. In Figure 7 the result of this operation is illustrated for September 03, 2009. If the given assumption for the occurrence of peak loads is correct, the difference between this time series and the real load profile yields the cumulative peak loads of all waste-crushers. For this day the maximum value of the power input amounts to 1,207 kW. Compared with the base load of 664 kW, the peak load caused by the waste-crushers adds up to 543 kW. The maximum peak load over the whole year adds

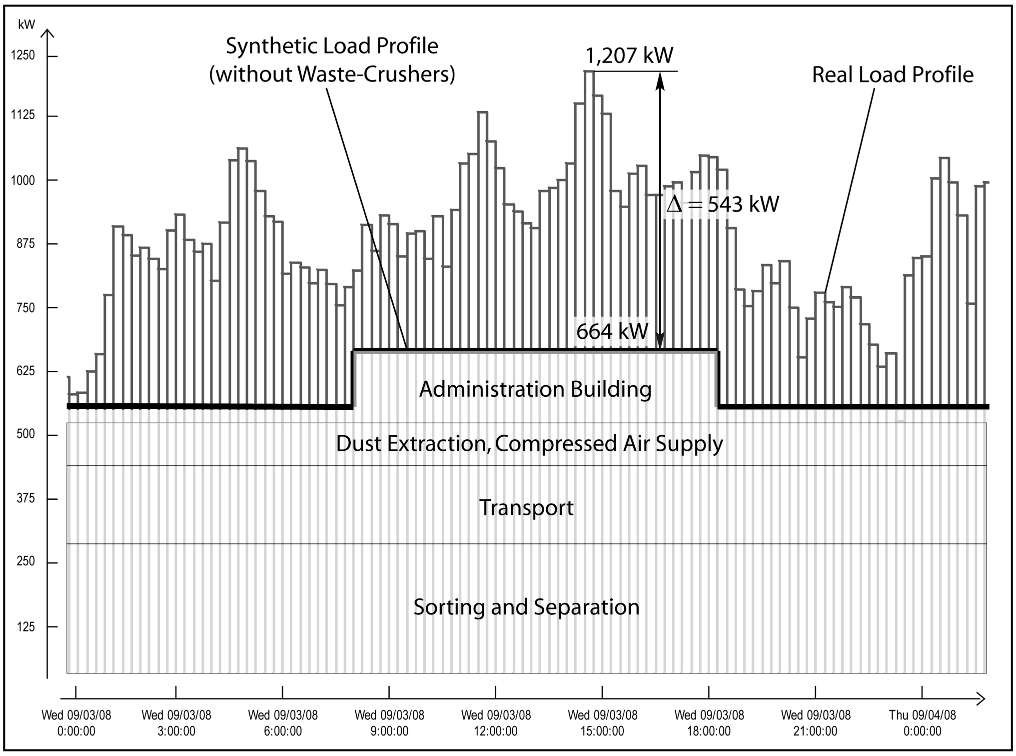

Figure 7: Comparison of real with synthetic load profile, no crushers (September 03, 2009).

up to 1,250 kW. Accordingly, the maximum electrical power demand of all crushers amounts to 586 kW. This is in good accordance with the assumed value of 558.1 kW, which is composed of 141.3 kW for the pre-crusher and another 416.8 kW for the four main-crushers, compare Table 1. Thus, the assumption that the fluctuating peaks in the use of electric energy are caused by the crushers seems to be correct.

Investigating the real load profile it turns out, that the described peak loads never occur in more than one 15 minute interval after another. In all cases there are time intervals in between those peak load times during that the power inputs take considerably lower values. With this knowledge it is possible to design a peak load management system that reduces the level of peak loads by controlling the run of the waste crushers. Therefore, the proposed peak load management system measures the power input of all five waste-crushers. Each time four of the five crushers operate in full load, the conveyor belts transporting waste to the fifth crusher stop. In that way the peak loads can be limited to values of about 1.180 kW. Taking into account the 20 €/kW monthly demand charge the considered business has to pay, this simple measure reduces the annual energy costs by approximately 17.000 €. Considering the costs for the setup and the calibration of the peak load management system, in total circa 10.000 €, the expenditures will amortize in less than a year.

6. Summary

The use of **η** with its standardized methods speeds up and enhances the quality of time series evaluation in energy systems analysis and optimization. In particular, its holistic approach enables the user to perform all necessary steps from the import to the conditioning up to the final export of time series in one single tool. The software is tailored to energy consultants acting in the analysis and optimization of industrial energy systems. However, the generic interface allows for arbitrary modifications and expansions of the available method repertoire. Therefore the general approach makes it possible to use **η** in any technical field where time series analysis and manipulation are central tasks in the workflow. Taking the considered waste recycling plant as an example it could be shown that practically relevant time series manipulation and analysis applications can be performed using **η** to help identifying main energy consumers as well as to find measures in order to improve the rational use of energy and to reduce energy costs.

References

[1] Shasha D., Zhu Y.: *High Performance Discovery in Time Series - Techniques and Case Studies.* Springer, 2004.

[2] Chatfield C.: *The Analysis of Time Series: An Introduction.* Taylor & Francis Ltd., 6th Edition, 2003.

[3] Augenstein E., Wrobel G., Kuperjans I., Pleßow M.: *TOP-Energy - Computational Support For Energy Sytem Engineering Processes.* 1st International Conference "From Scientific Computing to Computational Engineering", Vol. III, pp. 1284 – 1291, 2004.

[4] TOP-Energy Project: *TOP-Energy - Toolkit for Optimization of Industrial Energy Systems*. Chair of Technical Thermodynamics, RWTH Aachen University. Available at: www.top-energy.de

[5] Jung J., Kirschbaum S., Wrobel G.: *Multi-objective Approach to Analysis of Industrial Energy Systems.* 21st International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems, 2008.

[6] Augenstein E.: *Rechnergestützte Analyse und Konzeption industrielle Energiesysteme*. PhD Thesis. Chair of Technical Thermodynamics, RWTH Aachen University, 2009.

[7] Aspen Technology, Inc.: *Aspen Plus – Conceptual design of chemical processes*. Available at: www.aspentech.com

[8] Evonik Energy Services GmbH: *EBSILON Professional – Simulation of thermodynamic processes with emphasis on power plant technology.* Available at: www.steag-ketek.de

[9] The R Foundation for Statistical Computing: *The R Project for Statistical Computing*. Department of Statistics and Mathematics, WU Wien. Available at: http://www.r-project.org

[10] Zander F.: *Verwaltung von Zeitreihen in datenbank-basierten Informationssystemen am Beispiel von RBIS.* Diploma Thesis. Institute of Geography, Friedrich Schiller University of Jena, 2007.

[11] aqua_plan GmbH: *Time series and their Applications – Analysis and Design of a Data Model.* Available at: http://www.aquaplan.de/public_papers/time_series/