



CSV Data Format Specification

For the SA and WA Gas Retail Markets

Version: 3.3
Last Update: 31 October 2016

Version History

Version	Date	Author(s)	Changes and Comments
1.1	26/08/02	MK	Original from Michael Kroumer distributed per Victorian Market. Distributed with REMCo Specification Pack V2.01
2.9	19/01/04	BE/DB	Updated for South Australian and Western Australian markets. For distribution with REMCo Specification Pack V2.9. GRMS application refers to Section 2. The Interface Control Document takes precedence over the remaining sections of this document for GRMS specific requirements.
3.0	11/5/04	BE	<p>REMCo front page added.</p> <p>Cross references to other elements of REMCo Specification Pack replaced with a generic cross-reference to the REMCo Specification Pack Usage Guide.</p> <p>The caveat in section 1.3 about the relationship between this document and the ICD has been removed.</p> <p>Section 2.8.5 added to document date/time formats and policies regarding time zones.</p> <p>Examples in 3.1 and 5.1 revised to be SA/WA examples.</p> <p>Section 3.1 – File name and extension formats separated to reflect differences between REMCo to market transactions and other transactions.</p> <p>Section 3.1 – Point 5 amended to redefine the date/time stamp as a 14 character unique identifier. Also text 'all upper case' added.</p> <p>Section 4.3 amended to specify that all elements within the file name and extension are upper case.</p> <p>Section 2.2 amended to specify that for B2M transactions the header row is case sensitive and all upper case.</p> <p>Caveat added to example of a CSV transaction with 'all' as the 'to' participant.</p>
3.1	9/5/06	REMCo	Amendment to Section 2.2 to clarify the format of headers as per rule change C18/05S,
3.2	1/10/10	T Sheridan	Updated to reflect the relevant Market Operator requirements following the transfer of REMCo's SA retail market operations to AEMO.
3.3	31/10/16	D. McGowan	<p>Update to include:</p> <p>WA</p> <ul style="list-style-type: none"> • C02/16C – REMCo to AEMO transition changes. <p>SA</p> <ul style="list-style-type: none"> • IN029/16 – REMCo to AEMO transition

Table of Contents

1. Introduction	6
1.1. Purpose	6
1.2. Audience.....	6
1.3. Scope.....	6
1.4. Related Documents	6
1.5. Overview and Structure.....	6
2. CSV Format.....	7
2.1. Character Set	7
2.2. Header	7
2.3. Footer	7
2.4. Literals treatment.....	7
2.5. Line breaks.....	8
2.6. Values separator.....	8
2.7. Optional elements.....	8
2.8. Numeric values	8
2.8.1. Positive and negative values	8
2.8.2. Leading and trailing zeroes	8
2.8.3. Format.....	8
2.8.4. Dollar amounts	9
2.8.5. Date/time	9
2.9. Tab characters	9
2.10. Special characters	9
2.11. Leading and Trailing Spaces	9
3. File Format.....	11
3.1. File name and extension – Market Operator to Market Transactions.....	11
3.2. File name and extension – B2B Transactions and Market to Market Operator Transactions	11
3.3. End of file marker	12
4. Compression Format	13
4.1. Format	13
4.2. Number of CSV components per file.....	13
4.3. File name and extension	13
4.4. Password protection	13
5. Email Format – B2B Transactions	14
5.1. Subject Line	14

5.2.	Number of attachments	14
5.3.	Size restrictions	14
5.4.	Acknowledgements	14

1. Introduction

1.1. Purpose

This document defines format of Comma-Separated Values (CSV) files for transactions that are to be exchanged between Market Participants.

1.2. Audience

The audience of this document includes IT developers, IT managers, testers, system analysts & system architects involved in the design and development of systems for Gas Retail Markets in South Australia and Western Australia.

1.3. Scope

This document covers the delivery format of CSV components that are either:

- incorporated into an aseXML transaction (aseXML payload)
- communicated via an automated electronic file (see B2M Interface Control Document)
- attached to an e-mail or provided on a disk
- downloaded from a web site

The contents (data elements) of CSV components carried in aseXML transactions are defined the B2M Interface Control Definitions Document or the B2B Systems Interface Definitions Document.

1.4. Related Documents

This document should be read in conjunction with the other documents contained within the AEMO Specification Pack as defined in the AEMO Specification Pack – Usage Guide.

1.5. Overview and Structure

Section 2 of this document defines the requirements to a CSV component format, use of line separators, tabs and treatment of literals and numeric values.

Section 3 specifies requirements to the format of a file that may carry a CSV component.

Section 4 describes the compression format to be used with CSV files.

Section 5 covers requirements to the format of an e-mail that may have compressed CSV file attached.

2. CSV Format

A CSV component contains the values in a table as a series of ASCII text lines organized so that each column value is separated by a comma from the next column's value and each row starts a new line.

This section specifies CSV format details. Most of the sections here particularize requirements to CSV component to be delivered not via aseXML.

2.1. Character Set

The character set that is to be used within a CSV component is ASCII 7-bit. No Unicode characters are allowed.

2.2. Header

The very first line in a CSV component shall contain a set of column designators (column headers) that can be used in further processing. The column designators make CSV files more human readable and facilitate tracking of import problems. For non-aseXML B2M & M2B transactions that are sent through the FTP server, the header row is case sensitive and is all upper case. For aseXML B2M & M2B transactions with a CSV payload, the header row is case insensitive.

Only one header per CSV component is allowed.

2.3. Footer

No footer is required for a CSV component.

2.4. Literals treatment

The CSV import application shall be capable of parsing literals whether they are surrounded by double-quotes or not. A CSV import tool/parser shall interpret the following two rows from a CSV file analogously:

```
123,"This is a sample field",456
123,This is a sample field,456
```

If a literal containing double-quotes needs to be inserted into a CSV, the entire field must be surrounded by double-quotes in addition to prefixing each contained double-quote with another one, e.g. in order to insert a field

```
This is a sample "quoted" field
```

into a CSV row, the following format shall be used:

```
"This is a sample ""quoted"" field"
```

This treatment allows for the incorporation of commas and quotes into a CSV element, if necessary. As the "lowest denominator", the format produced by Microsoft® Excel matches this treatment of literals.

2.5. Line breaks

The CSV components shall have a combination of Carriage Return (CR, ASCII decimal code 13) and Line Feed (LF, ASCII decimal code 10) at the end of each line.

Empty lines, i.e. lines containing just CR and LF, are not allowed.

2.6. Values separator

A comma "," is to be used to separate values in CSV file. If a comma shall occur inside a literal, then the entire literal shall be surrounded by double quotes as per 2.4.

No trailing commas are allowed at the end of each line, i.e. the number of value separators in any one row will always be: *number_of_values* – 1.

2.7. Optional elements

Where a CSV element is defined as "optional" or marked as "not required" or "NR" in the B2M Interface Control Document or the B2B Systems Interface Specification Document, at very least, a placeholder for that element shall be present in a CSV component row.

For example, consider a 7-value CSV row for which 6th and 7th values are defined as optional.

```
Col_Hdr_1,Col_Hdr_2,Col_Hdr_3,Col_Hdr_4,Col_Hdr_5,Col_Hdr_6,Col_Hdr_7
abcdef1,defgh1,1234,123.45,qwertyuio,,
abcdef2,defgh2,6543,234.56,qwertyuio,,
. . .
```

2.8. Numeric values

2.8.1. Positive and negative values

Positive numbers in CSV file shall be unsigned. Negative numbers shall be prefixed with a negative sign.

2.8.2. Leading and trailing zeroes

There shall be no leading zeroes unless a specific data format requires this. Trailing zeroes are allowed only when the *scale* value requires so, refer to 2.8.3.

2.8.3. Format

The numeric format is defined as **numeric** [(*precision* [, *scale*])]. The *precision* (referred to as 'Logical Length' in the ICD and B2B SID) and *scale* (referred to as 'Decimal Length' in the ICD the B2B SID) determine the range of values that can be stored in a numeric field:

- The precision specifies the maximum number of decimal digits that can be stored in the column. It includes *all* digits, both to the right and to the left of the decimal point. Precisions can range from 1 digit to 38 digits.
- The scale specifies the maximum number of digits that can be stored to the right of the decimal point. The scale shall be less than or equal to the precision. You can specify a scale ranging from 0 digits to 38 digits or use the default scale of 0 digits.
- The number of digits to the left of the decimal point cannot exceed $precision - scale$

An *Integer* can be represented as a numeric value with the *scale* of 0, i.e. Numeric(11,0)

Below are examples of **valid** values for a numeric type defined as Numeric(5,3)

12.345	12.000	0
-12.345	12	
12.100	12.0	

Here are some examples of **invalid** values for the type defined as Numeric(5,3):

1,200	12-
12.345678	123456.78

2.8.4. Dollar amounts

Dollar amounts shall not use a dollar sign "\$" either as a prefix or a postfix.

2.8.5. Date/time

In general, date/time and time elements in the body of CSV transactions will be expressed with a Time Zone Designator (TZD).

Date/time in the format ccy-MM-ddThh:mm:ss+hh:mm

Time in the format hh:mm:ss+hh:mm

Fractions of seconds can be included as an option in the format ss.ss... with any number of digits after the decimal point supported.

The time zone selected will be at the discretion of the sending party. The sending party must therefore ensure that the combination of time and time zone accurately communicates the point in time being defined.

For example, if the time to be communicated is 9.00am (Central Australia Standard Time), the data element could contain 09:00:00+09:30 or 09:30:00+10:00. It is then up to the receiving party to ensure that they have the ability to convert this time to another time zone if required.

In the case of the CSV element 'Planned_Outage_Time', as this is only included in a manually-prepared e-mail, it will always be in local time without a Time Zone Designator.

Note: as defined in the B2B-B2M Hub Specifications and Architecture documents all date/time stamps in the messaging layer (ebXML) and in the headers of aseXML transactions will be expressed in GMT+10 (market time).

2.9. Tab characters

Tab characters shall not be used in CSV files.

2.10. Special characters

The use of CDATA (non-parsed character data as defined in Ref.[2]), characters "<", ">", "&" and hexadecimal characters is prohibited.

2.11. Leading and Trailing Spaces

In the case of numeric values the use of a leading, embedded or trailing space is inappropriate. Spaces should not be used where a value has a Numeric characteristic.

Where the value has a “text” characteristic that by its nature it can have a space or spaces as part of the structure, only embedded spaces are permitted. Leading and trailing space-characters immediately adjacent to the comma field separators should not be included in the CSV file. Therefore “John Citizen” resolves to ---,John Citizen,--- or ---,”John Citizen”,---

Where values must have a leading or trailing spaces as a valid part of the data it must be delimited with double-quote characters. In this way it indicates that the leading and trailing spaces are a component part of the data for example:

---,” John Citizen ”,---

3. File Format

For non-aseXML transactions, CSV component may be incorporated into a file. Following the compression, the file can be delivered to the receiver as an e-mail attachment or on a disk.

3.1. File name and extension – Market Operator to Market Transactions

The file name shall be constructed of the following elements (case-sensitive and all uppercase) separated by underscore "_" signs:

1. The Gas Retail Market . For South Australia it shall always be "SAGAS". For Western Australia it shall always be "WAGAS".
2. The name of the transaction to which the CSV file is supplied, as defined in the AEMO Specification Pack.
3. The CSV file originator identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents.
4. The CSV file recipient identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents.
5. A 14 character alphanumeric unique identifier with the following:
 - <datetime><sequence>
 - where
 - <datetime> is YYYYMMDDhhmm
 - <sequence> is a 2 digit value ranging from '00' (zeroes) to 'ZZ'
6. The file extension of "CSV", case-sensitive, all uppercase, separated from the file name with a period "."

Here's an example of a CSV file name for the WA market:

`WAGAS_UAI_REMCO_ALS_2004050114359B.CSV`

3.2. File name and extension – B2B Transactions and Market to Market Operator Transactions

The file name shall be constructed of the following elements (case-sensitive and all uppercase) separated by underscore "_" signs:

1. The Gas Retail Market. For South Australia it shall always be "SAGAS". For Western Australia it shall always be "WAGAS".
2. The name of the transaction to which the CSV file is supplied, as defined in the AEMO Specification Pack.
3. The CSV file originator identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents.
4. The CSV file recipient identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents.
5. A date/time stamp in the format CCYYMMDDHHmmSS when the file has been generated, 24-hour format, local time of the originator.

6. The file extension of "CSV", case-sensitive, all uppercase, separated from the file name with a period "."

Here's an example of a CSV file name for the SA market:

`SAGAS_ENERGYHISTORYRESPONSE_TXUR_ENVSA_20040703131500.CSV`

Note: If, as in the above example, a date/time stamp is used as the unique ID, fractions of seconds will not be included in the date/time stamp.

The following is an example of a CSV file name that is intended for more than one recipient:

`SAGAS_ENERGYHISTORYRESPONSE_TXUR_ALL_20040703151500.CSV.`

Note: not all participants (including tAEMO) may be able to accept transactions with 'all' in the file name.

3.3. End of file marker

The application that parses CSV shall be able to handle End-Of-File mark (EOF, ASCII decimal code 26) at the end of the file, if present.

4. Compression Format

4.1. Format

The supplied file shall be compressed using ZLIB format specification, as per RFC1950. The zip format is defined as being compatible with the PKZIP v2.04G utility. Utilities that support this format include WinZip and PKZIP.

Default level of compression shall be used.

4.2. Number of CSV components per file

There shall be only one CSV component per compressed file.

4.3. File name and extension

The name of the compressed CSV file name shall match the uncompressed CSV file name with extension "ZIP", case-sensitive, upper case, replacing "CSV".

4.4. Password protection

Compressed files are not normally password protected. If they are to be password protected these requirements will be specified in the Non-IT RMR Communications document which is part of the Information Pack. .

5. Email Format – B2B Transactions

A compressed file containing CSV component shall be attached to an e-mail for delivery to a participant. The specifications in this section facilitate internal delivery by participants. Note Email format for AEMO to Market transactions are defined in the ICD.

5.1. Subject Line

The e-mail subject line for the compressed file incorporating CSV component shall be composed of the following parts, separated by a "_" signs:

1. The Gas Retail Market. For South Australia it shall always be "SAGAS". For Western Australia it shall always be "WAGAS".
2. The name of the transaction to which the CSV file is supplied, as defined in the AEMO Specification Pack
3. The CSV file originator identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents
4. The CSV file recipient identifier, valid participant as per the FRC B2M-B2B System Specification and FRC B2M-B2B System Architecture Documents
5. Date/time stamp in the format CCYYMMDDHHmmSS when the file has been generated, 24-hour format, local time of the originator.

Below is an example of a constructed e-mail subject line for the SA Market:

[SAGAS_ENERGYHISTORYRESPONSE_TXUR_ENVSA_20040703131500](#)

5.2. Number of attachments

Each e-mail shall only contain one compressed file attachment to simplify internal routing that may be implemented by some participants.

5.3. Size restrictions

Compressed e-mail attachment shall not exceed 2MB in size to support limitations imposed by some e-mail gateways.

5.4. Acknowledgements

No acknowledgments shall be expected to CSV files delivered by e-mail or on the disk.