INTERNATIONAL STANDARD

ISO 23219

First edition
2022-07

Natural gas — Format for data from gas chromatograph analysers for natural gas — XML file format

Gaz naturel — Format pour les données des analyseurs de chromatographie en phase gazeuse pour le gaz naturel — Format de fichier XML

Reference number
ISO 23219:2022(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO /TC 193, *Natural gas,* Subcommittee SC 1, *Analysis of natural gas*.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

There are multiple suppliers of gas chromatograph analysers for measurement of natural gas composition. If correctly set up, there is no reason to prefer any one in particular, since they give comparable results. However, the situation gives rise to a variety of methods of reporting data, which can create confusion over the use of such data for off-line calculation of properties or evaluation of analyser performance.

Therefore, a uniform method of data presentation, independent of the source of the analyser, as presented in this document is considered valuable.

Natural gas — Format for data from gas chromatograph analysers for natural gas — XML file format

1 Scope

This document specifies a text file format - XML file format - for reporting natural gas analysis results and other data relevant to natural gas. The file name is applicable when it includes the extension of .XML (case insensitive).

The XML file format is useful for output from ISO 6974-1[1] for composition and ISO 6974-2[2] for uncertainty, for input for ISO 6976[3] and for input for ISO 10723[4] for performance evaluation. Typically these would be the gas composition as provided on an analysis certificate, or results from a performance evaluation that would be read into an Excel spreadsheet for data processing.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 8601, *Date and time — Representations for information interchange — Part 1: Basic rules*

3 Terms and definitions

No terms and definitions are listed in this document.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

4 Basic instructions

The following steps shall be obeyed to come to a correct XML representation:

a) Unambiguous, reasonably concise, non-proprietary, and since the chemical formula is part of the InChI it is often clear what the component is. The alternative of using Chemical Abstracts Registry Number (CAS RN) is not so understandable and is proprietary.

b) Attributes in tags should be avoided, e.g. not <amount unit=”mol%”>1,2</amount> but <amount><value>1,2</value><unit>mol%</unit></amount>

c) The XML should have some checksum (at the end), e.g. CRC-16 check (hexadecimal) sum as displayed by ZIP programs, at the end of the file as an XML comment, e.g. <!--D86A3640-->. This is not intended for security purposes, merely as a confidence check that the file contents have not been corrupted.

d) Date and time are tagged <date_time> and use the convention specified in ISO 8601, e.g. <date_time>2018-07-10T13:59:00+01</date_time>

   This data may be used for

   — date and time of sample registration,

— date and time when the results were produced,

— date and time when the results were validated, or

— date and time when the report was issued.

e) Date and time are always useful data and the specified format is any format that is acceptable to Excel and VBA. For example, time may be hh:mm:ss (e.g. 13:45:10), where seconds are optional (assumed to be 00 if omitted, and hours are 24 hour clock). Date may be dd/mm/yyyy (e.g. 27/02/2017) or dd-mmm-yy (e.g. 27-Feb-17).

f) In general XML tags is lowercase (including inchi, which is usually written as InChI)

g) Any multiple word tags to be concatenated with underscore, e.g. <retention_time>

h) In general tags should be complete words, rather short forms or acronyms

i) The kind of numeric datatype Decimal (type xs:decimal or xs:float/xs:double) shall be chosen. Numbers should not contain spaces and use period as the decimal separator; i.e. no grouping of digits, no thousands separator (e.g. comma). E-notation is used for powers of 10, e.g. 34, 12.3, -3.4567, 5.6E3 and 7.8E-2.

j) Component data are in <peak> blocks (see annexes).

k) Units should use a single solidus or -exponent, e.g. MJ/m^3 (see annexes), or $MJ \cdot m^{-3}$ Space between units is not allowed.

l) Properties should be in a <properties> block, with each property in a <property> block, and, to be flexible, the property will be identified with a keyword rather than by an XML tag (see annexes).

m) XML is defined by the tags – keywords enclosed in angular bracket, e.g. <value>, with a terminating tag with / in front of the keyword, e.g. </value>.

n) This format of the XML file format is defined in the schema given in Annex A.

o) The XML tags defined in the schema are case insensitive (they are all lowercase) and have no leading or trailing spaces. The tags have no attributes, i.e. no name="value" within the tag.

p) The content of the tags, e.g. keywords, are case insensitive, and leading and/or trailing spaces are ignored. In the examples (see Annexes B, C and D), the contents are often with trailing spaces in order to improve readability.

q) The only likely need for lowercase is to distinguish the S.I. prefixes of m (milli) and M (mega). Keywords should not contain spaces.

r) Example keywords and units are given in Annexes A, B, C and D. Commonly useful keywords are specified in those same annexes.

s) In general, any keywords and units may be agreed between writing and reading applications (see Clause 5).

t) Reading applications should ignore any keyword sections that are not relevant to their operation.

u) Software, capable of processing XML data to Excel and vice versa is listed in Annex E.

v) Tags that contain values should have unique names.

5 General comments to keywords and units

5.1 Possible keywords and units are given in Annex A (see also examples on commonly useful keywords in Annexes B, C and D).

5.2 Example gas component names are specified in Annex F.

5.3 Amount units may be:

— ppm mol, mol%, mf or mol_fr (mole fraction), or

— ppm mass, mass% or mass_fr (mass fraction),

— concentration units, e.g. mg/m^3, are not recommended, but if they have to be used then the reference conditions for the volume shall be given in the units, and whether the volume amount includes the named component or not.

NOTE PPM is parts per million moles, i.e. moles per million moles (or parts per million ideal gas volume).

5.4 For calorific value (CV) the following should be considered:

— gross CV is the same as superior CV and higher CV, or

— net CV is the same as inferior CV and lower CV.

Possible units: MJ/m^3, kJ/mol, MJ/kg, Btu/ft^3

5.5 According to ISO/IEC Guide 98-3[6] every measurement should have an associated uncertainty. This can be specified using the <uncertainty> tag. This is the standard uncertainty (one standard deviation) in the same units as the value (i.e. not as a percentage), and coverage_factor shall be specified. The form of the distribution can be specified. In omitted the distribution is assumed to be Normal (Gaussian). Other common distributions are Uniform and Triangular – for these the uncertainty specifies the half range.

5.6 Since amounts are in mole fraction then the total amount must sum to unity, hence there will always be correlation between the mole fractions. The correlation_coefficient block is to allow for this. It is usual to assume this is an identity matrix (=1 if row=column, =0 if row<>column) if it isn't provided. In the block it is expected that row < column. The correlation matrix is symmetric.

5.7 Retention time, height and area have no need for units. Only that the units should be consistent for that analyser and analysis. Retention time is to be in seconds.

Annex A
(normative)

<iso23219> XML tags

A.1 <measurements>

A.1.1 <parameters>

Optional parameters to describe the measurements. Examples below.

A.1.1.1 <date_time>

See Clause 4 item d).

A.1.1.2 <cylinder_number>

Text description.

A.1.1.3 <certficate_number>

Text description.

A.1.2 <peak>

A.1.2.1 <component>

Component data.

A.1.2.1.1 <name_local>

Text description of the component.

A.1.2.1.2 <parameters>

Optional parameters to characterize the component.

A.1.2.1.2.1 <k_name>

Text description of the property. Example names (and units) :-

molar_mass	[kg/kmol]	
boiling_point	[K]	(normal, at 101,325 kPa)
specific_gravity	[-]	(60/60ºF relative density)
critical_temperature	[K]	
critical_pressure	[MPa]	
critical_volume	[m^3/kmol]	

critical_density [kg/m^3]

acentric_factor [-]

Note that the InChI descriptor contains the chemical formula.

A.1.2.1.2.2 <k_value>

Numerical value.

A.1.2.1.2.3 <k_units>

Text description of standard units.

A.1.2.1.3 <inchi>

Text InChI descriptor of the component (see Annex F).

A.1.2.1.4 <amount>

Data for the component's amount of substance.

A.1.2.1.4.1 <value>

Numerical value.

A.1.2.1.4.2 <units>

Text description of amount units (see 5.3).

A.1.2.1.4.3 <uncertainty>

Data for the component's amount uncertainty.

A.1.2.1.4.3.1 <u_value>

Numerical value.

A.1.2.1.4.3.2 <u_coverage_factor>

Numerical value, e.g. for a 95 % confidence interval of a normal distribution the value is 1,96; which is often taken as 2.

A.1.2.1.4.3.3 <u_distribution>

Text description of the distribution (see Clause 5 e). If absent assumed normal.

A.1.2.1.4.3.4 <u_measurements>

Integer value. The number of measurements used in the calculation of <u_value>

A.1.2.1.4.3.5 <u_correlation_rc>

Integer value. Unique value for <correlation_coefficients> <c_row> or <c_column>.

A.1.2.2 <retention_time>

Numerical value in seconds.

A.1.2.3 <peak height>

Numerical value with arbitrary units (consistent throughout the file).

A.1.2.4 <peak area>

Numerical value with arbitrary units (consistent throughout the file).

A.1.3 <correlation_coefficients>

See 5.5. If absent assumed to be an identity matrix.

A.1.3.1 <element>

A.1.3.1.1 <c_row>

Integer value (referenced by <u_correlation_rc>).

A.1.3.1.2 <c_column>

Integer value (referenced by <u_correlation_rc>).

A.1.3.1.3 <c_value>

Numerical value (in the range -1 to +1).

A.2 <properties>

A.2.1 <method>

Data for the method for the property calculation.

A.2.1.1 <m_name>

Text description of the method used, e.g. ISO 6976[3], ISO 20765-2[5].

A.2.1.2 <parameters>

Parameters that are relevant to the method. Examples below.

A.2.1.2.1 <combustion_temperature>

Numerical value: < 32 assumed °C, 32 to 100 assumed °F, >100 assumed K.

A.2.1.2.2 <metering_temperature>

Numerical value: < 32 assumed °C, 32 to 100 assumed °F, >100 assumed K.

A.2.1.2.3 <metering_pressure>

Numerical value: > 25 assumed kPa , <25 assumed psiA.