Can the South African address standard (SANS 1883) work for small local municipalities?

Serena Coetzee[1], Antony Cooper[2]

[1]Dept Computer Science, University of Pretoria, Pretoria, South Africa, scoetzee@cs.up.ac.za
[2]Logistics and Quantitative Methods, CSIR, PO Box 395, Pretoria, 0001, South Africa, acooper@csir.co.za

Abstract

The aim of the South African address standard (SANS 1883), currently being developed under the auspices of the South African Bureau of Standards (SABS), is not to devise a new system of addressing or to build a national address database, but rather to enable interoperability in address data, which in turn will facilitate developing a national address database. The standard defines twelve address types that describe all forms of addresses currently in use in South Africa. The UML data model for these includes 60 elements. The question arises as to whether such an all-encompassing model is practical in a local municipality with the responsibility to produce and maintain official addresses for only two of the address types in SANS 1883, namely the Street Address and Site Address types. The objectives of this paper are to 1) point out complexities in SANS 1883 relating to the Street and Site address types; 2) propose a simplified data model for these address types; 3) show how address data based on this model can be maintained in lock-step with other datasets such as the cadastre and municipal boundaries; 4) recommend content for SANS 1883 guidelines for assisting, in particular, small local municipalities implement an address database that can be exchanged according to SANS 1883. While our paper is based on the South African address standard, the Street address type, consisting of a street number, street name and place name, is common to many countries and our findings are thus applicable to an international audience.

1. Introduction

The aim of the South African address standard (SANS 1883), which is currently being developed under the auspices of the South African Bureau of Standards (SABS), is not to devise a new system of addressing or to build a national address database, but rather to enable interoperability in address data, which in turn will facilitate developing a national address database. SANS 1883 consists of three parts:

- SANS 1883-1. *Geographic Information – Address Standard, Part 1: Data format of addresses* (committee draft).
- SANS 1883-2. *Geographic information – Address Standard, Part 2: Guidelines for addresses in data bases, data transfer, exchange and interoperability* (committee draft).
- SANS 1883-3. *Geographic information – Address Standard, Part 3: Guidelines for address*

allocation and updates (committee draft).

SANS 1883-1 defines twelve address types that describe all forms of addresses currently in use in South Africa: the Street address, Site address, Intersection address, Building address, Farm address, Informal address, Landmark address, SAPO box address, SAPO street address, SAPO site address, SAPO post restante address, and SAPO-type village address. The Unified Modelling Language (UML) data model for these twelve address types includes 60 elements. The question arises as to whether such an all encompassing data model is practical in a local municipality with the responsibility to produce and maintain official addresses for only one or two of the SANS 1883 address types. Some data elements of an address, such as the municipality, the province, and the country, are identical for all addresses in that municipality and do not have to be recorded separately for each individual address. This information is necessary only when the address data is shared or exchanged, when it can be added.

Metadata helps to understand and interpret the content of the data being exchanged, and amongst others, SANS 1883-1 provides for the originator and distributors of the address data, the point of observation for recording the location of the address (e.g. center of the property or street front), the life cycle stage of the address (future, active, retired) and the official status of the address. Once again, this information is necessary only when the address data is shared or exchanged, when it can be added.

Coetzee and Cooper (2007b) describe how SANS 1883 aims to provide an all-encompassing description for an address in South Africa. They highlighted current causes for ambiguities in addresses, and described solutions for reducing these causes for ambiguity, such as having a single set of official place name boundaries for the country, and mandating certain authorities to assign addresses. These causes for ambiguities add to the complexities found in SANS 1883 and described in section 2 of this paper. Coetzee and Cooper (2007a) also describe a range of benefits that standardization of addresses would bring to South Africa and its people.

In European countries such as Denmark, where an official national address register is maintained, such ambiguities are being eliminated. A recent study in Denmark analysed the qualitative and quantitative impact of address ambiguities. The qualitative analysis confirmed that the ambiguities affect people every day, sometimes even resulting in life threatening situations. The quantitative analysis proved that in the long term the costs of renaming the streets are covered by the savings arising from eliminating these ambiguities. A Danish statutory order now prohibits address ambiguities (Lind, 2007).

Street addressing plays a key supporting role in municipal development (Farvacque-Vitkovic *et al*, 2005), and due to their service, infrastructure and land administration responsibilities, it is commonly found that a local authority establishes and maintains address reference data for its area of jurisdiction (Coetzee *et al*, 2008). Address standards have been developed and are currently being developed by a number of countries and international organizations. These include Australia and New Zealand (as a joint effort), Denmark, South Africa, the United Kingdom, the United States of America, the Universal Postal Union (UPU), the International Organization for Standardization (ISO) and the Organization for

the Advancement of Structured Information Standards (OASIS), and Coetzee *et al* (2008) analyzed various characteristics of these standards. A European survey on addresses and address data (EUROGI, 2005) gives clear evidence that although address systems exist in European countries, with a long history as well, and although address master files or address registers are available in most countries on certain conditions, only very few published standards for address data exist, making the task of "interoperable and seamlessly accessible" address data sets "across all of Europe" even more difficult.

There are examples, however, of national address standards that have been implemented successfully to enable a national address dataset or register. In the UK, for example, the National Land and Property Gazetteer (NLPG) is the specific implementation of *BS 7666 – Spatial datasets for geographical referencing* and holds addresses of all fixed man-made properties. It covers England and Wales and comprises 27.8 million property records with over 29.3 million associated addresses. Underlying this project is a definitive National Street Gazetteer (www.nsg.org.uk) containing details of 1.5 million records and also based on *BS 7666*. A similar project in Scotland is also approaching maturity. The implementation of *BS 7666* within local government calls for each local authority (376 in England and Wales and 32 in Scotland) to create and maintain a local gazetteer of all land and property within their administrative area for which they have statutory obligations in relation to planning and development control and street naming and numbering. Data entry conventions have been agreed and contractual relationships and timetables have been set to ensure that gazetteers to a common standard are available (Coetzee et al., 2008; www.nlpg.org.uk).

Another example is the Australian Geocoded National Address File (G-NAF®) of the Public Sector Mapping Agencies (PSMA) which follows a semi-automated process of massaging contributor address data into a standardized format that is acceptable for merging into the G-NAF®. Any address data that cannot automatically be converted into the standard address format, is subjected to a manual review process. The PSMA is the custodian of the G-NAF® and acts as a clearinghouse by merging data from as many as 15 government agencies and organizations into the G-NAF® (Paull, 2003). In Ireland a definitive reference directory for addresses is maintained by *An Post and Ordnance Survey Ireland* (OSi). The GeoDirectory (www.geodirectory.ie), as it is called, combines postal addresses (where mail is delivered) and geographic addresses (a geo-code to position the address on a map) in one database which is available to organizations or individuals who require it (Fahey and Finch, 2008).

As can be seen from the British example, an address standard enables the development a national address database and related gazetteers. The Australian example shows how an address standard can facilitate integrating data from different sources. The Irish example shows the benefits of geocoding postal addresses. SANS 1883 aims at facilitating all these benefits for South Africa.

Cooper (2008) points out some commonalities in the address definitions of a number of national and international address standards, and while our paper is based on the South African address standard, the *Street address* type, consisting of a street number, street name and place name, is common to many countries and our findings are thus applicable to an international audience. Thus, the implementation of

an address standard at a local authority is a scenario not only found in South Africa, but also elsewhere in the world.

The objectives of this paper are to 1) point out complexities in SANS 1883 relating to the Street and Site address types; 2) propose a simplified data model for these address types; 3) show how address data based on this model can be maintained in lock-step with other datasets such as the cadastre and municipal boundaries; 4) recommend content for SANS 1883 guidelines that would assist, in particular, small local municipalities to implement an address database of which the data can be exchanged according to SANS 1883.

2. Complexities in the South African address standard

Addresses come in many forms and have a variety of uses: an address is not only a set of directions for delivering post; addresses also facilitate the delivery of a wide range of other services such as utility services such as water, sewerage, telecommunications and electricity supply; refuse collection; billing; postal and courier delivery; to emergency response; goods delivery; serving summonses; household surveys; and visiting (Coetzee and Cooper, 2007b). SANS 1883-1 provides the following definitions:

***Address**: an unambiguous specification of a point of service delivery.*
***Point of service delivery** (or service delivery point): actual location where a service could be provided.*

In contrast to postal mail which is usually delivered by a single agency or institution within a country that can prescribe the format and content of an address on the postal item, the SANS 1883 definition for an address includes service delivery by any institution in any number of ways (by post, by hand, by vehicle, or even virtually for a financial service), adding to the complexity of addresses that have to be represented in SANS 1883. For example, the *Street address* type and the *SAPO street address* type differ only in the locality part of the address. A sample address of each one of these two address types is listed in Table 1 below. These two sample addresses refer to the same service delivery point even though their localities seemingly differ.

Table 1. Sample addresses

Street address type	546 Puccini Street, Constantia Park
SAPO street address type	546 Puccini Street, Glenstantia, 0181

Address databases for South Africa exist, some in the public and others in the private sector, but none of the providers of these address databases have been given a mandate to provide an official register of addresses. The Geographic Names Council (GNC) is responsible for standardizing place names in South Africa, and in this capacity the GNC has delegated the responsibility for allocating addresses to the municipalities. However, municipalities are not concerned with postal delivery: they

use addresses for other purposes such as land administration, road maintenance, tax collection and delivery of water, electricity, sewerage and other services. Thus, a municipality does not have to accommodate all SANS 1883 address types in its address database.

SANS 1883 defines the *official addressing body* as the authority to whom power has been delegated to assign addresses and notes that currently the only such bodies in South Africa are the SA Post Office for postal addresses, and municipalities for all other official addresses in their areas of jurisdiction. In line with this definition, SANS 1883 allows the status attribute to be set to 'Official' only if the address type is *Street address*, *Site address* or *SAPO-type rural village address*. The *Street address* type describes addresses as we commonly know them in the Eurocentric world: street number, street name and a locality (i.e. suburb, place name, neighborhood, and the like). The *Site address* type describes a variety of addresses, especially those that were assigned in the apartheid era to black townships on the outskirts of cities. In these townships street names were not always assigned. A typical township address consists of a number and a locality in the form of the name of the township, suburb or section. Municipalities are now starting to assign street names in these areas, but since it is a political process that involves the local community, progress is slow (Coetzee, 2008). Thus, the South African address standard has to cater for the official status of these township addresses. *Site addresses* are also sometimes used in security estates and similar developments. Finally, the *SAPO-type village address* type describes addresses that are assigned as part of the SA Post Office's rural addressing project (Rossouw and Kgope, 2007). This is the only address type currently in wide use in rural areas and is included as an official address type for this reason and because it is applied systematically and consistently across the country, and in collaboration with the local communities.

While SANS 1883-1 describes all address types currently in use in South Africa, SANS 1883-3 provides guidelines for the allocation and maintenance of street names and numbers in addresses of the *Street address* type, and refers the reader to the relevant documentation from the SA Post Office regarding the allocation of addresses of the *SAPO-type viallage address* type in rural areas. SANS 1883-3 recommends that where addresses do not yet exist, addresses of the *Street address* type in formal areas and the *SAPO-type village address* type in informal areas should be allocated.

The purpose of the South African address standard is to enable the sharing and exchange of address data. For this reason attributes such as the coordinate reference system, point of observation, originator, custodian and resource provider are included in the SANS 1883 data model. Data elements such as the municipality, province, and country are included for the same reason. Since these latter attributes and data elements are identical for all addresses in a particular municipality (except for those few that lie across provincial boundaries), they do not have to be recorded for each individual address separately. These attributes are necessary only when the address data is shared or exchanged, when they can be added in a batch. Thus, the address database that is used for day-to-day maintenance of addresses in the municipality does not have to include these attributes and data elements.

3. Simplification of the SANS 1883 address data model for use at local municipalities

In an informative annex of SANS 1883-1, a simplified data model is included for two of the twelve address types, the *Street address* and *Site address* types, which are the two official address types in formal areas. In this simplified data model the number of elements has been reduced from 60 to 32, and if the *Site address* type is removed, the number of elements can be further reduced to 27. Such a simplified data model still includes attributes and data elements that are identical for all addresses in a particular municipality and thus a further simplification of the data model for day-to-day address data maintenance is possible. In this section we first describe a 'bare bones' data model of only six elements for day-to-day address data maintenance at a municipality. This models includes the *Street address* type only, but can easily be adjusted to include the *Site address* type. In the second part of this section, we show how data in this 'bare bones' data model can be prepared for data exchange according to the simplified data model provided in the informative annex of SANS 1883-1.

Figure 1 shows the six elements of the 'bare bones' data model for day-to-day address data maintenance. In this data model the separate types and unions for attributes and data elements have been condensed into the base type *Address* and its dervived type *Street Address.* The four other elements are a data type for alphanueric text, an enumeration for the street name type, a type for the address location and a code list for the recorded names. These elements are used to define the attributes of the base type *Address* and its dervived type *Street Address*.

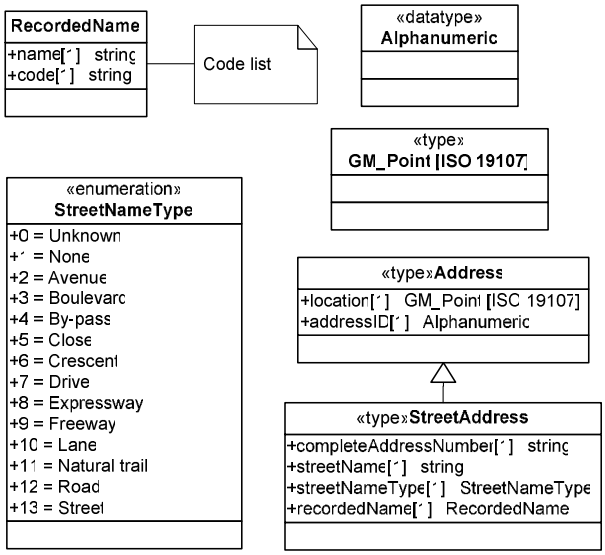

Figure 1. 'Bare bones' data model for day-to-day address data maintenance (adapted from SANS 1883-1)

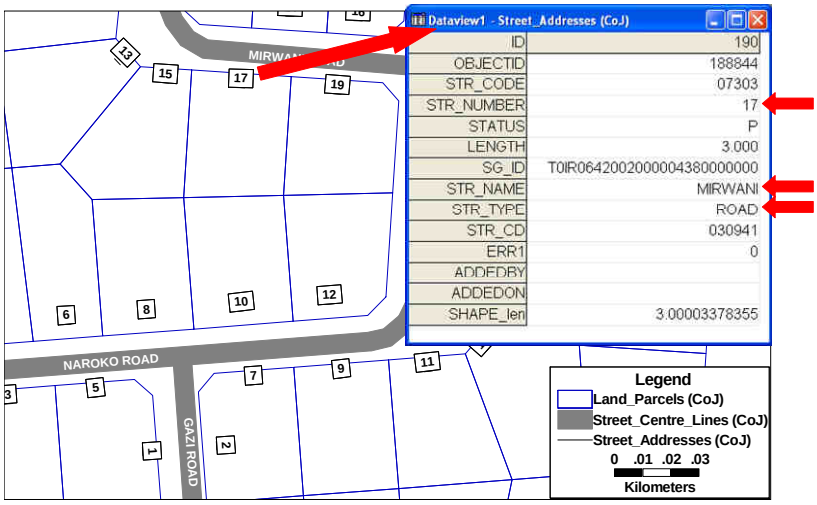

Figure 2. Practical implementation of the 'bare bones' data model

In practice this 'bare bones' data model can be implemented as a single table or layer of point features in which each feature represents an address – refer to Figure 2. Ideally, the link between the address and the recorded name should be derived spatially, but it could also be implemented through a foreign key relationship.

To prepare address data from this 'bare bones' data model for data exchange, a number of attributes and data elements have to be added. In Tables 2 and 3 we list these attributes and data elements respectively and describe how they can be derived or assigned in preparation for data exchange. The second column in each table refers to the relevant SANS 1883 type to which the attribute or data element belongs.

Table 2. Preparation address data elements for data exchange

Data elements		
Data element name	**Type name**	**Description**
recordedName	RecordedNameWithUsedNames	Derive spatially from surrounding recorded name boundary
municipality	Locality	Derive spatially from surrounding municipal boundary
province	Locality	Derive spatially from surrounding municipal boundary
country	Locality	'ZA'

Data maintenance is simplified by maintaining separate datasets for the cadastre, the recorded names, municipalities and provinces. The boundaries of land parcels and recorded names are approved by municipalities and recorded at a Surveyor General's office. Once a boundary is approved, this change is implemented once in the relevant dataset and then derived spatially for all the addresses within that boundary. It is not necessary to update address records individually and there is no need to individually maintain the link between a land parcel or recorded name and an address. In South Africa the Municipal Demarcation Board (MDB) (www.demarcation.org.za) is an independent authority

responsible, in terms of the Constitution, for the determination of municipal boundaries. Municipal and provincial boundary data are freely available from the MDB's website. Provincial boundaries are protected by South Africa's consitution and can only be altered by the National Assembly, with a supporting vote of at least two thirds of its members and by the National Council of Provinces, with a supporting vote of at least six provinces. Provincial boundaries have remained unchanged since the adoption of the Constitution in 1996. Thus, whenever the address data need to be prepared for data exchange, the latest versions of the cadastre, the recorded names, municipalities and provinces are used to fill in the relevant attributes and data elements in the address data.

Table 3. Preparation address attributes for data exchange

Attributes		
Attribute name	**Type name**	**Description**
coordinateReferenceSystem	GeographicalLocation	Set according to the reference system used in the data that are being prepared
pointOfObservation	GeographicalLocation	Set the appropriate enumerator from PointOfObservation according to the convention at the municipality, usually *Centre of land parcel*, *Entrance to the point of service delivery* or *At the street front to the service delivery point*.
custodian	Address	Set to the official name of the municipality
originator	Address	Set to the official name of the municipality
resourceProvider	Address	Set to the official name of the municipality
landParcelKey	Address	If cadastral data are available, derive spatially from surrounding land parcel, otherwise NULL
addressType	Address	*AddressType.StreetAddress*
language	Address	*'EN'*
lifeCycleStage	Address	*LifeCycleStage.Active*
status	Address	*Status.Official*
completeStreetNameElementSequence	CompleteStreetName	*CompleteStreetNameElementSequence. NameAndTypeDirectionalModifier*, assuming that either directionals and modifiers are not used, or this sequence is applied throughout the municipality
streetNameAndTypeSequence	CompleteStreetName	*StreetNameAndTypeSequence.NameType*, assuming the street type always follows the name in this municipality

Synchronization with these 'external' datasets when preparing address data for exchange has the further advantage that anomalies can be detected, such as address data records in the municipal dataset that fall outside the municipal or provincial boundary. An address should also be within the boundaries of a land parcel and recorded name, and address records outside these boundaries are incorrectly geo-

referenced, or indicate an error in the recorded name boundary, or fall within an informal settlement. The number of addresses within a single land parcel is another consistency indicator: while one or even two address records per land parcel (on a street corner or where a subdivision has not yet been recorded in the cadastre) are quite common and allowable, more than two indicate either poor data quality or, once again, an informal settlement. Due to the existence of informal settlements, these anomalies and inconsistencies cannot be removed without closer inspection.

In our model above, the street name details are maintained separately for each individual address point feature. An alternative approach would be to link individual address records through a foreign key relationshop to a line feature in the street layer or dataset. The street name details are then maintained in the street dataset and automatically updated in the address records via this link. While this makes it possible to maintain the street name details in a single layer or dataset, care should be taken not to lose information when the street network is maintained, for example, when line features are added or removed to reflect a change in the current traffic direction.

The latest version of SANS 1883-1 includes an informative annex with a simplified UML data model for the *Street address* and *Site address* types only. In this paper we have illustrated a further simplification of this simplified UML data model and described how it can be practically implemented at local municipalities. We recommend that this description of the simplified, practical implementation of the South African address standard should be made available to the public on the SANS 1883 website (www.cs.up.ac.za/~scoetzee/sans1883) which is accessible by the public. Sample data files of this simplified and practically implementable model should also be available for download in various formats such as DBF, ESRI SHP files, MapInfo TAB files, etc.

4. Conclusion

In this paper we presented a simplification and practical implementation of the UML data model in the first part of *SANS 1883, Geographic Information – Address Standard, Part 1: Data format of addresses*. While the South African address standard has to cater for all address types currently in use in South Africa, we have shown how a small subset of these address types can be practically implemented at local municipalities, and how this address data can be prepared for data exchange. The periodic synchronization with other datasets such as the cadastre, recorded names, municipal and provincial boundaries has the added advantage that anomalies and consistencies can be detected.

Similar simplifications of the UML data model are possible for other address types in SANS 1883 for specific uses, for example, for the five postal address types relevant to organizations working only with the postal addresses of customers. We recommend that this description of the simplified practical implementation of the South African address standard, along with sample data files that can be downloaded, is made available to the public on the SANS 1883 website (www.cs.up.ac.za/~scoetzee/sans1883). We would also like to encourage other communities to publish

similar documentation with practical guidance for the implementation of SANS 1883.

If one could show that our 'bare bones' data model can be implemented for address data based on address standards from other countries or organizations, this would give an indication as to the absolute core requirements for address data at a local authority level, and how such data can be prepared for address data exchange. It would also provide input towards a compendium of best practices for the maintenance of address data. Such best practices could be applied when developing reusable software components for address allocation and maintenance. Such work, albeit beyond the scope of this paper, would support the maintenance of address data, and improve the quality of address data worldwide.

Acknowledgements

The authors wish to thank the SABS for initiating the project for the South African address standard (SANS 1883), and for allowing us to report on it here. We would also like to thank all delegates to the project meetings for their contributions and active participation in developing the draft standard, and also those who provided inputs via correspondence.

References

BS 7666:2006 Spatial datasets for geographical referencing, 2006, British Standards Institution (BSI), United Kingdom.

Coetzee S 2008, 'Address data exchange in South Africa', Proceedings of the *ISO Workshop on address standards: Considering the issues related to an international address standard*, ISBN 978-1-86854-689-3, 25 May 2008, Copenhagen, Denmark.

Coetzee and Bishop J 2008, 'Address databases for national SDI: Comparing the novel data grid approach to data harvesting and federated databases', *International Journal of Geographic Information Science*, accepted for publication, to appear.

Coetzee S and Cooper AK 2007a, 'The value of addresses to the economy, society and governance – a South African perspective', *Proceedings of the 45th Annual URISA Conference*, 20-23 August 2007, Washington DC, USA.

Coetzee S and Cooper AK 2007b, 'What is an address in South Africa?', *South African Journal of Science*, Nov/Dec 2007, 103(11/12), pp449-458.

Coetzee S, Cooper AK, Lind M, McCart Wells M, Yurman SW, Wells E, Griffiths N and Nicholson MJL 2008, 'Towards an international address standard', *GSDI-10 Conference*, Trinidad and Tobago, 25 – 29 February 2008.

Constitution of the Republic of South Africa, 1996, viewed 17 August 2008, http://www.info.gov.za/documents/constitution/index.htm .

Cooper AK 2008, 'Overview of an address and purpose of the workshop', Proceedings of the *ISO Workshop on address standards: Considering the issues related to an international address standard*, ISBN 978-1-86854-689-3, 25 May 2008, Copenhagen, Denmark.

EUROGI, 2005, *Survey of European National Addressing as of May 2005*, Edited by Karen Levoleger and Christopher Corbin, Apeldoorn, NL, viewed 17 August 2008, http://eurog.vbnprep.com/POOLED/DOCUMENTS/a101730/EUROGI_Address_Survey_Resp_V3Final.pdf.

Fahey D and Finch F, 2008, 'GeoDirectory Technical Guide', *An Post / Ordnance Survey Ireland*, Republic of Ireland, viewed 17 August 2008, http://www.geodirectory.ie/getdoc/e99524f1-7dad-49bf-9bc2-7251d81a2905/GeoDirectoryTechnicalGuide-v8.aspx.

Farvacque-Vitkovic C, Godin L, Leroux H, Chavez R and Verdet F 2005, *Street Addressing and the Management of Cities*, Washington DC, US: The World Bank.

ISO 19107:2003, *Geographic information – Spatial schema*, International Organization for Standardization (ISO), Geneva, Switzerland.

Lind M 2007 'Benefits of common address data – experiences and assessments', *45th Annual URISA Conference*, 20-23 August 2007, Washington DC, USA.

Paull D 2003, 'A Geocoded National Address File for Australia: The G-NAF What, Why, Who and When', *PSMA Australia Limited*, viewed 17 August 2008, http://www.psma.com.au/resources/the-g-naf-what-why-who-and-when.

Rossouw P and Kgope K 2007, 'South African Post Office Rural addressing in South Africa', *PositionIT*, Sept/Oct 2007, pp 66-71.

SANS 1883 South African address standard website, available online at www.cs.up.ac.za/~scoetzee/sans1883, accessed 31 July 2008.

SANS/CD 1883-1. Geographic Information – Address Standard, Part 1: Data format of addresses (committee draft), 2008, South African Bureau of Standards (SABS), Pretoria, South Africa.

SANS/CD 1883-2, Geographic information – Guidelines for addresses in databases, data transfer, exchange and interoperability (committee draft), 2008, South African Bureau of Standards (SABS), Pretoria, South Africa.

SANS/CD 1883-3, Geographic information – Guidelines for address allocation and updates (committee draft), 2008, South African Bureau of Standards (SABS), Pretoria, South Africa.