Using the extended quarter degree grid cell system to unify mapping and sharing of biodiversity data

R. Larsen[1,2]*, T. Holmern[1], S. D. Prager[3], H. Maliti[2] and E. Røskaft[1]

[1]*Department of Biology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway,* [2]*Tanzania Wildlife Research Institute, PO Box 32, Arusha, Tanzania and* [3]*Department of Geography, University of Wyoming, 1000 E. University Ave. Laramie, WY 82071, U.S.A.*

Abstract

Information on the distribution of animal populations is essential for conservation planning and management. Unfortunately, shared coordinate-level data may have the potential to compromise sensitive species and generalized data are often shared instead to facilitate knowledge discovery and communication regarding species distributions. Sharing of generalized data is, unfortunately, often ad hoc and lacks scalable conventions that permit consistent sharing at larger scales and varying resolutions. One common convention in African applications is the Quarter Degree Grid Cells (QDGC) system. However, the current standard does not support unique references across the Equator and Prime Meridian. We present a method for extending QDGC nomenclature to support unique references at a continental scale for Africa. The extended QDGC provides an instrument for sharing generalized biodiversity data where laws, regulations or other formal considerations prevent or prohibit distribution of coordinate-level information. We recommend how the extended QDGC may be used as a standard, scalable solution for exchange of biodiversity information through development of tools for the conversion and presentation of multi-scale data at a variety of resolutions. In doing so, the extended QDGC represents an important alternative to existing approaches for generalized mapping and can help planners and researchers address conservation issues more efficiently.

Key words: atlasing, biodiversity informatics, geocoding, GIS, QDGC, Tanzania

Résumé

L'information sur la distribution des populations animales est essentielle pour la planification de la conservation et la gestion. Malheureusement, les données partagées au niveau des coordonnées risquent de compromettre les espèces sensibles, et les données généralisées sont souvent partagées pour faciliter la découverte et la communication des connaissances concernant la distribution des espèces. Le partage de données généralisées est, malheureusement, souvent opportuniste et manque de conventions mesurables qui permettraient le partage cohérent sur une plus grande échelle et à des résolutions variées. Une convention commune pour des applications africaines est le système de *Quarter Degree Grid Cells* (QDGC). Cependant, la norme actuelle ne supporte pas l'emploi des références uniques à travers l'Equateur et le premier méridien. Nous présentons une méthode pour étendre la nomenclature QDGC pour soutenir l'adoption de références uniques à l'échelle du continent, en Afrique. Le QDGC étendu fournit un instrument pour partager les données généralisées sur la biodiversité là où les lois, les réglementations et les autres considérations formelles empêchent ou interdisent la distribution de l'information au niveau coordonné. Nous disons dans quelle mesure le QDGC étendu peut être utilisé comme norme, une solution mesurable pour l'échange d'informations sur la biodiversité grâce au développement d'instruments pour la conversion et la présentation de données à échelle multiple à des résolutions diverses. Ce faisant, le QDGC étendu représente une alternative importante aux approches existantes pour la cartographie généralisée et il peut aider les planificateurs et les chercheurs à traiter les problèmes de conservation plus efficacement.

Introduction

Human demographic pressure and consumption levels are causing a rapid loss of biodiversity on a global scale

**Correspondence:* E-mail: ragnvald@ntnu.no

(May, Lawton & Stork, 1995). However, the international community has come to recognize the importance and role of biodiversity and the need to reverse this process (CBD, 1992; UN, 2000). To address these demands, collection and dissemination of biodiversity data are being carried out at an ever increasing rate (Soberon, Llorente & Onate, 2000; Soberon & Peterson, 2004), where such data often are geo-referenced and stored within a geographic information system (Chapman, Muñoz & Koch, 2005), thus allowing more sophisticated analysis of spatial trends and patterns (Balmford *et al.*, 2005).

The collection of spatially referenced biological data is generally at the most detailed level possible given available funding levels and purpose of the collected data. Collection measures may include a variety of approaches including animal counts in georegistered aerial photos (Jachmann, 2002) and line transects on the ground (Treydte, Edwards & Suter, 2005). A common theme in nearly all such data is that geographic location is of principle importance. Management of protected areas relies on understanding where different species are located and the relative distribution of species throughout a study or management area. One common approach is the use of metric squares as enumeration areas and consolidating collected information relative to these squares (Norton-Griffiths, 1973; Campbell & Borner, 1995). In management applications, these squares are often based on the Universal Transverse Mercator coordinate system (UTM). While UTM is ideal for the local management of coordinate data and corresponding metric squares, its zonal nature does not readily extend to large scale, multi-zone analysis. For instance, while positional references in adjoining UTM zones may refer to the same geographic location, both the actual coordinates and constructed reference schemes will differ across zones. Although this issue can be overcome using automated approaches, the problem of creating a set of unique references suitable for larger scale applications remains.

Although UTM-based approaches may be ideal for relatively small-scale applications, the maximum benefit of collected biodiversity data can only be realized if the data are published and shared. Most collected biological data are associated with highly accurate geographic coordinates, but such raw data may have the potential to compromise research integrity or wildlife security. At the same time, if no data collections are made available, the entire conservation community is disadvantaged and, in essence, flying blind with regard to other research activities. Approaches are thus required to facilitate the publication of generalized locational information in a consistent manner without compromising the integrity of any specific location.

Given the highly local nature of most biological data, one pressing challenge is the development of a useful scheme that can allow consistent, generalized reference at a continental scale. Currently there exist few continental scale grid systems with a continuously predictable and scalable nomenclature. Two such approaches include C-Squares (Rees, 2003) usually applied in marine science, and the Quarter Degree Grid Cells (QDGC) system used in several African atlases. For instance, Carswell *et al.* (2005) presents thirteen bird atlases where the presentation level is based on full degrees, half degrees or quarter degrees. While there are other systems in use (e.g. WMO squares), many of these have not typically been used in biodiversity applications and often lack scalability in terms of resolution (Table 1). In addition, there are also a variety of nonsquare systems but these systems tend to be oriented towards more computational approaches (Sahr, White & Kimerling, 2003), and are not represented on any standard hard-copy map products.

In spite of its widespread use in the bird atlases, the original QDGC system has several inherent limitations. Because of its naming convention, the original definition of QDGC is unable to support unique references on a global scale. In addition, the original standard does not accommodate cell references smaller than the originally specified 15′ by 15′ smallest unit. A consequence of this lack of extensibility is that in the preparation of the second edition South African Bird Atlas, the working group has now started to use 5×5-min pentads and 3×3-min triads to supplement for the lack of smaller grid square of the original version of QDGC (Harebottle *et al.*, 2007).

Given the constraints of the original QDGC for continental scale analysis, the desire to prevent further implementation of stop-gap measures, and the emerging emphasis on the use of QDGC as a basis for improving the quality and distribution of biodiversity sampling efforts (Robertson & Barker, 2006), we present an extended model of the QDGC that addresses these issues. In this paper, we develop the QDGC reference system further, providing a suitable standard for both presentation of data on a continental grid system and for making management maps at any level. The following sections describe the original QDGC and the new extensions to the original standard. We examine general issues associated with the use of gridded data, as well as some of the long term

Table 1 Attributes of the extended Quarter Degree Grid Cells compared to other major reference systems

Reference system	Description	Coverage	Advantage (+)/disadvantage (−)	Example of use	Reference
QDGC (original)	Represents a way of making consistently measured squares covering a specific area to represent specific qualities of the area covered. The squares themselves are based on the degree squares covering earth (e.g. 3302 BB)	Regional (Africa south of Equator, east of zero meridian)	+ Easy to produce location strings − Does not support unique references across the Equator and Prime Meridian − Does not accommodate cell references smaller than 15′ × 15′	Bird atlases	Harrison *et al.* (1997)
QDGC (this study)	Represents a way of making consistently measured squares covering a specific area to represent specific qualities of the area covered. The squares themselves are based on the degree squares covering earth (e.g. E033S02BB)	Global	+ Easy to produce location strings + Supports point data and easily convertible to latitude and longitude coordinates + Allows general level analyses without exposing the precise coordinates of potentially sensitive information + Systematic communication of information + Readily backwards compatible with QDGC	Biodiversity mapping	This study
British National Grid (BNG)	The grid is based on the OSGB36 ™ datum (Ordnance Survey Great Britain 1936, based on the Airy 1830 ellipsoid). The maps adopt a Transverse Mercator projection with an origin at 49°N, 2°W. Over the Airy ellipsoid a straight line grid, the National Grid, is placed with a new false origin (to eliminate negative numbers), creating a 700 km by 1300 km grid. A location can be indicated to varying resolutions numerically by adding digits to the coordinate (e.g. NN 166712)	Regional (Great Britain, western Europe)	+ Easily convertible to latitude and longitude using available tools − Limited coverage	Biodiversity mapping	Birch (1949)
Universal Transverse Mercator (UTM)	In this grid, the world is divided into 60 north–south zones, each covering a strip 6° wide in longitude. These zones are numbered consecutively beginning with Zone 1, between 180° and 174° west longitude, and progressing eastwards to Zone 60, between 174° and 180° east longitude (e.g. 4833439 630084 17T)	Global	+ Very good accuracy, reproducibility and flexibility. − In some countries (e.g. U.S.A. and Mexico) maps lack proper grids rendering UTM recording difficult − Accuracy falls off with distance from the central meridian potentially complicating mapping when large features of interest span multiple UTM zones − The existence of separate grids for the 60 UTM zones complicates interpretation and mapping	Biodiversity mapping	Bugayevskiy & Snyder (1995)

Table 1 Continued

Reference system	Description	Coverage	Advantage (+)/disadvantage (−)	Example of use	Reference
Military Grid Reference System (MGRS)	Geocoordinate standard used by militaries for locating points on the earth. In most areas (between latitudes 80°S and 84°N), the MGRS grid is identical to the UTM grid system, but uses a different labelling convention (e.g. 4QFJ12345678)	Global	+ In addition to including the advantages of UTM, it also deals with the problem of distinguishing between zones by assigning letter references and redundant discriminators to the UTM zone number − Tile spaced arbitrarily (leading to problems when important features fall on grid seams)	Military	Kimerling *et al.* (2005)
World Meteorological Organization (WMO) squares	Divides a chart of the world with latitude–longitude gridlines into grid cells of 10° latitude by 10° longitude, each with a unique, 4-digit numeric identifier. Each 10° × 10° square is allocated a number between 1000 and 7817.	Global	+ Easy to transfer between latitude and longitude coordinates and the grid system − Limited scalability (cannot be subdivided)	Metrologic	Curry (2001)
C-squares	C-squares provides a hierarchical nomenclature for dividing 10° × 10° World Meteorological Organization (WMO) squares into smaller units (each an individual 'c-square') of 5° × 5°, 1° × 1°, 0.5° × 0.5°, 0.1° × 0.1°, etc., using an alternating base 2, base 5 linear division, as fine as may be required. Each cell of the resulting subdivision is allocated a unique alphanumeric identifier (c-squares code) (e.g. 1316:225:469)	Global	+ Provides a solution to the limitation of 'bounding rectangles' representation of dataset footprints and associated search procedures + Provides an unlimited, recursive nomenclature for subdividing WMO squares + Allows general level analyses without exposing the precise coordinates of potentially sensitive information	Marine surveys	Rees (2003)
World Geographic Reference System (WGRS)	Defines a system of uniform regional grids (each 100° × °100 km), anchored on and named by prominent cultural and/or physical places. Subsets – called local grids may also be defined (e.g. US.DC.WAS.54. 18.28)	Global	+ Location description make it easily interpreted + Communication with electronic devices (e.g. GPS receivers, GIS) + Localization string focuses on places of maximum interest and use − Overlapping local coordinate systems may cause confusion about precise location if reference feature is unknown or miss specified	Communica-tion between electronic devices	Clarke, Dana and Hastings (2002)

challenges and opportunities associated with the adoption and use of the refined standard.

Materials and methods

QDGC system principle

Atlases with grid cell-based representation can be traced back to early publications by The Atlas of British Flora (Perring & Walters, 1962). The grid cells provided a convenient enumeration unit for indicating species distribution without creating a map that would be too complex to consume.

Several approaches were developed to associate positional reference information with individual cells. One approach to gain traction was the QDGC specification. The QDGC reference string described at the Avian Demography Unit Department of Statistical Sciences consists of four numbers and two letters (e.g. 3302BB) (Harrison *et al.*, 1997). A position given in longitudes/latitudes is readily translated into the QDGC reference string by a simple procedure. To find the code, the African Demographic Unit (ADU) suggests the following guideline (ADU 2001):

1 Each degree square is identified by a four digit number comprised of the values of latitude and longitude at its upper left corner, e.g. 3302, the large square in the diagram below (Fig. 1a).

2 Each degree square then divided into sixteen 'quarter degree squares,' each 15′ × 15′. Put another way, each quadrant of the original one degree by one degree cell is dividing into four quadrants for a total of sixteen cells within the original whole-degree area. These are given two additional letters, the first indicating which quarter degree quadrant, the second indicating which quarter–quarter degree quadrant (e.g. 3302BB, Fig. 1a).

At this stage, the QDGC reference string could point to a generalized location within any single quadrant resulting from the intersection of the Prime Meridian and the Equator (though, there is no specification with regard to which global quadrant any location string might fall within). Given this limitation and the geographic emphasis of the original application, the use of QDGC was restricted to an area covering southern Africa (Fig. 2). Accordingly, the hemispheric and resolution limitations, applications of QDGC outside of southern Africa or those requiring a representation smaller than 15′ × 15′ must use an alternate approach or must implement a customized stop-gap measure.

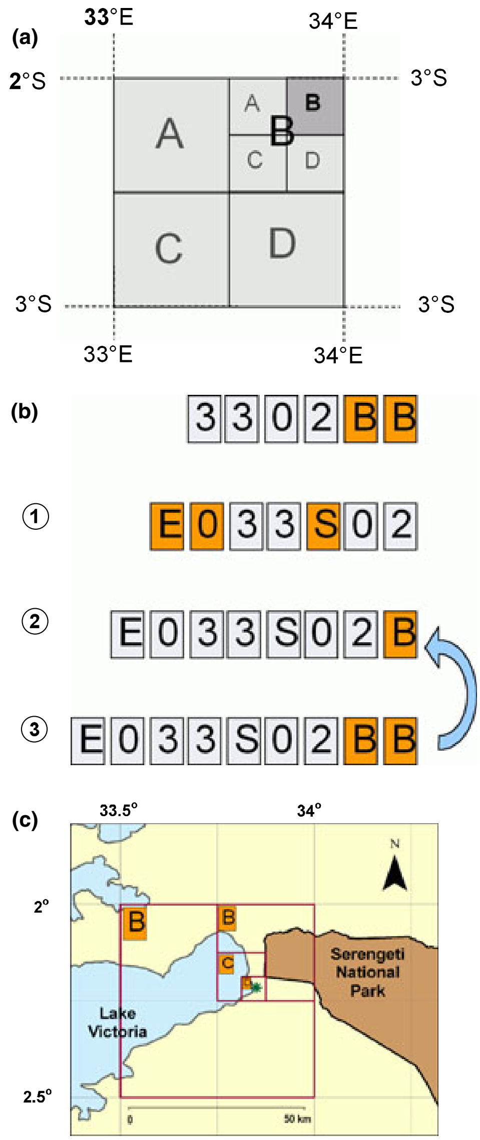

Fig 1 Illustration of Quarter Degree Grid Cells (QDGC) (a) in the old standard in which the QDGC ends at quarter degree squares (the dark grey square) giving a reference string of 3302BB indicated by bold letters, (b) where the recursive nature of QDGC is described, the hemispherical references are indicated and the longitudal value is given a preceding zero to preserve its invariable length and (c) with a sample observation symbolized by * from western Serengeti where QDGC is used to level 4 described with the following QDGC reference string: E033S02BBCD

Extending the QDGC standard

To make the QDGC reference string unique on a global scale and more useful for both larger and smaller scale

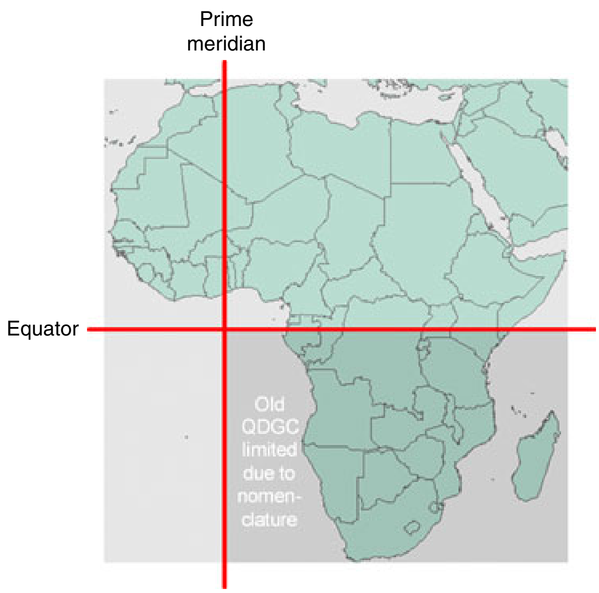

Fig 2 A map showing the African continent where the dark grey area is covered by the old Quarter Degree Grid Cells system. The southern parts of Africa were covered while areas north of Equator and east of the Prime Meridian were not covered

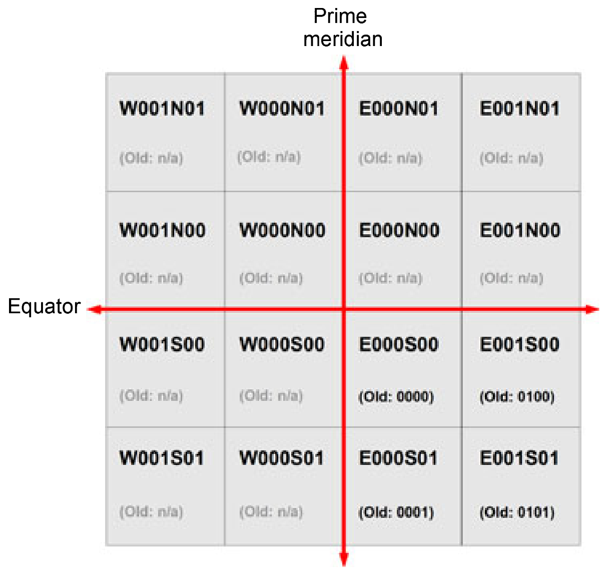

Fig 3 The figure shows how the reference corner is depending on the hemispheres, it also indicates how the old Quarter Degree Grid Cells standard was not applicable in three of the four quadrants

applications, we extend the QDGC reference system in a logical and consistent manner. Philosophically, the extensions to QDGC must address the above mentioned hemisphere issue, must be scalable to accommodate a wide range of possible resolutions (e.g. the size of the individually represented squares), must be easily implementable, and should facilitate ready mapping between the original QDGC standard as well as other generalized geocoding standards.

We begin by addressing the issue of addressing unique locations independent of hemisphere. This change simply requires the addition of characters delimiting north/south and east/west position. As the east/west reference may be up to 180 we add a preceding zero, ensuring that all east and west degree measurements are padded to three characters (Fig. 1b, step 1). We retain the ADU standard for addressing individual degree squares via the lowest longitude/latitude pair, so that the reference corner is that which is closest to the intersection of the Prime Meridian and the Equator. Thus results in a different reference corner depending on the hemisphere (Fig. 3).

The original QDGC standard allows a final division of a full degree square into four by four sub-squares. By allowing the process of dividing a full degree to continue on the resulting square for each level, the extended approach facilitates a set of new squares that are 1/4th of the original square for each previously calculated quadrant. This is in keeping with quad-tree approaches commonly used in GIS (Mark, Lauzon & Cebrian, 1989) and carries with it a variety of analytical efficiencies that may be harnessed as data collections expand.

Formally, the process is described as follows:

1 Each degree square is designated by seven characters. The first character indicates whether the square is east or west of the zero meridian (E or W). The three following numbers indicates at which degree of longitude the square starts. The next character indicates whether the square is north or south of the equator (N or S). Next are the two numbers indicating the latitudinal distance to the equatorial line. This is a reference to a unique degree square and is referred to as QDGC level 0.

2 Each degree square is divided into four squares where the upper left quadrant is designated A, the upper right B, the lower left C and the lower right D (always proceeding in a Z pattern from upper left to lower right). This is QDGC level 1.

3 Each resulting quadrant is subject to the same process as described in the step 2. A character quadrant indicator (A, B, C or D) is added at the end of each iteration (e.g. level 4: E033S02BBCD). The level number is increased by one per iteration (Fig. 1b).

As illustrated, step 1 indicates level 0, and the levels are readily identifiable by the number of characters following the level 0 designation (e.g. two characters following the level 0 designation is level 2). The level 4 application of the extended QDGC designation is illustrated by Fig. 1c. For the mathematically inclined, it is worth noting that the number of squares at each level designation is the level designation as an exponent associated with the base of 4. Thus, for level 0, $4^0 = 1$ square, for level 1, $4^1 = 4$ squares, for level 2, $4^2 = 16$ squares, and so on. This readily allows the calculation of the fraction of the original one degree by one degree area for a square at any level (Table 2).

Results: implementation and use

Generalized geo-referencing systems such as the QDGC have inherent limitations to their use. These can be summarized as problems related to lack of predictable or consistent nomenclature, area distortion, discontinuity at large spatial resolution and restricted area coverage. Here we have further developed one of the most used systems in Africa, QDGC, to include an extended nomenclature, which will allow assigning unique squares at a greater spatial resolution and without geographic restriction. Although the original QDGC was designed for areas relatively near the equator, the system is also well suited for use in other areas as long as the relative differences in cell sizes are taken into consideration (see Fig. 4). This provides a crucial tool for improving conservation planning and prioritization. While the previous section addressed the extensions to QDGC, this section will address the issues regarding area distortion, use and benefits gained by having a scalable, globally consistent implementation.

Table 2 Some examples of the resolution of Quarter Degree Grid Cells square in col/rows per degree square, squares per degree square, decimal degrees, area extent (variable) and sample reference strings using the African continent as an example

Level	Cols/Rows Per degree square	Squares per degree square	Extent in decimal degrees	Min/average/max area (km^2)	Example
0	1	1	1	9780/11,504/12,308	E033S02
1	2	4	0.5	2445/2876/3077	E033S02B
2	4	16	0.25	611/719/769	E033S02BB
3	8	64	0.125	153/180/192	E033S02BBC
4	16	256	0.0625	38/45/48	E033S02BBCD
5	32	1024	0.03125	9.5/11.3/12	E033S02BBCDA
6	64	4096	0.015625	2.4/2.8/3	E033S02BBCDAD
Contd.	…	…	…	…	…

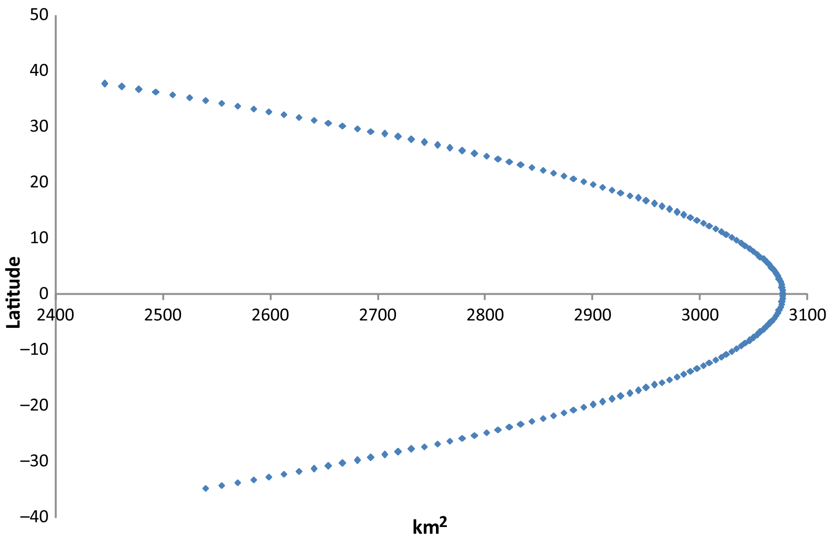

Fig 4 Change in area for QDGC Level 1 squares (0.5° × 0.5° squares). The squares are biggest around Equator and then the sizes becomes smaller the further north or south the squares are located. The graph only depicts squares for a north/south transect on continental Africa from 37N to 35S

Area distortion and calculation

The Atlas of southern African Birds is the most comprehensive atlas of birds in Africa using the QDGC system (Harrison *et al.*, 1997). The publication covers six nations comprising of 7,332,504 individual records.

Although widely used, the difference in QDGC areas as a function of latitude is not always discussed in detail. For example, in their study of spatial patterns of avian assemblages in South Africa, Githaiga-Mwicigi, Fairbanks & Midgley (2002) do not offer details with regard to the effects of this issue. While Wilson *et al.* (2007) mention the grid area as an issue of concern, they explicitly mention that no correction for the QDGC 'box size' is made in their analysis of invasive species.

The issue of defining enumeration areas is a classic problem in the geographic literature. Typically referred to as the modifiable areal unit problem or MAUP, the designation of arbitrary geographic zoning of phenomena has the potential to significantly impact the representation of distribution and the evident spatial relationships (Fotheringham & Rogerson, 1993). The best solution to overcome the MAUP is to represent actual entity-level information or to use a very high resolution QDGC level. When this is not possible, it is important to take into account the variation of QDGC cell size with latitude to avoid the introduction of any additional error or uncertainty.

The squares covering earth in the longitude/latitude system start out being almost 'square' around the equator and become narrower the farther north or south they are placed following the convergence of the meridians towards the poles. As such, proper accounting for the geographic area requires consideration of each QDGC cell in an orthogonal coordinate system wherein area can be accurately calculated. For purposes of maintaining consistent area calculations at a continental scale, we adopt the Albers Equal Area Conic projection (Bugayevskiy & Snyder, 1995) to calculate the squares of all QDGC covering Africa. While other equal area projection could be used, care should be taken to understand how the selection of a projection may bias accuracy towards any particular continent.

For level 1, the size in cells ranges from *c.* 3077 km^2 at Equator to 2445 km^2 (North Africa) and 2539 km^2 (southern tip of Africa). This results in cells at the latitudinal extremes of Africa that range from 79.5% to 82.5% of the area of the equatorial cells. On a country level, there is less variation. For example, we find that the difference for level 4 square sizes in South Africa is *c.* 12.4%, for Tanzania 2.15%, whereas for smaller countries such as Rwanda, the difference is 0.1%. We suggest studies using QDGC should take great care in using the correct area representation and we include the corresponding area calculations in QDGC layers published online (Larsen, 2007a). As differences in area become increasingly prominent further away from equator, correct area representation (e.g. for density calculations) becomes critical for larger scale analysis. When correct areas of the squares are taken into account using an equal area projection, the issue of misrepresentation of relative density becomes tractable.

An example implementation

The Biodiversity Atlas GIS server (Larsen, 2007b) offers a system for presenting biodiversity related information by using QDGC. By combining the QDGC standard and geo-referenced observation data, the system facilitates an all-in-one system suitable for projects or offices where the aim is to document biodiversity at a local, regional, continental and/or global scale. (Fig. 5). The system is embedded into a Linux-based virtual machine, incorporates HostGIS Linux (HostGIS, 2008) and utilizes freely available virtualization software from VMWare to allow the Linux environment to operate in a standard Windows environment (VMWare, 2008). Supported features include data storage, data entry, map viewing, data export such as KML (KML, 2008) and email backup routines.

Discussion: challenges and opportunities

A tool for conservation planning and management

Good management of natural resources is based on sound scientific analysis. The ideal is a process where solid and documented data is accessible for managers and policy makers for analysis. Providing building blocks like the extended QDGC geocoding standard will promote the release of scientific data and thus be an important part of this process.

The publication of generalized data via QDGC is important for two reasons. First data published at any QDGC level may be reinterpreted using dasymetric mapping techniques to address some of the issues associated with the MAUP mentioned above. Dasymetric mapping is an approach that

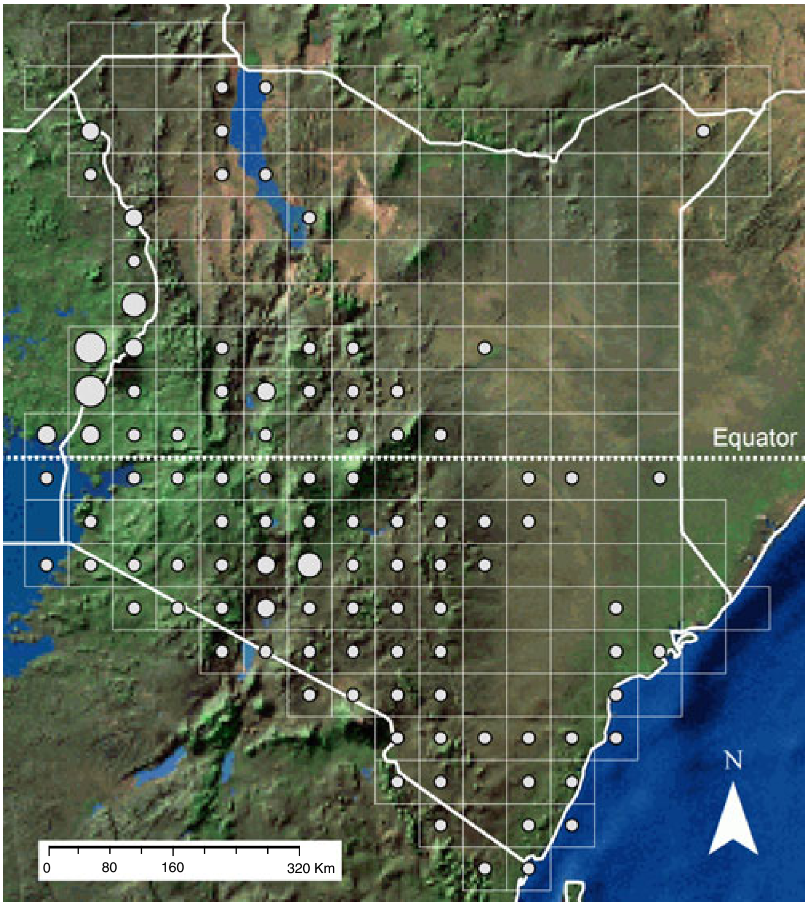

Fig 5 A map showing species richness of birds in Kenya by using data from a database on western Palaearctic migrant birds from the Center for Macroecology, University of Copenhagen (Walther & Rahbek, 2002). The presentation resembles that of the Biodiversity GIS database Server (Larsen, 2007b), but has been modified for clarity. Size of dots indicate relative richness

allows for areal interpolation based on ancillary data (Mennis, 2003). In a biodiversity example, a single QDGC cell might span two very distinct habitat zones. A user with detailed habitat information could use this information to estimate where within the QDGC cell a represented species might actually exist. Users wishing to estimate approximate geographic distributions of species may use information such as remotely sensed land cover data or other appropriate data of their choosing to serve as ancillary information in a dasymetric interpolation process. This approach would allow for a wide range of reuse of the generalized information to produce additional detail, without compromising the original information in any way.

The second reason data published via the QDGC is vital to the conservation community is related to systematic communication of biodiversity information. Already, the Global Biodiversity Information Facility (GBIF) facilitates global scale digitization and dissemination of primary biodiversity information through the use of point data and corresponding metadata (GBIF, 2007). As Yesson *et al.* (2007) illustrate, however, at the continental scale, there is negative correlation between GBIF coverage and species richness. Given the vast biodiversity reserves associated with developing countries in the tropics, these areas are very likely to see intensified mapping efforts in the near future. Given that past efforts, particularly in avian demography projects in southern Africa, have utilized QDGC (Harrison *et al.*, 1997; Carswell *et al.*, 2005), it follows that an appropriately extended QDGC represents a logical, backwards compatible approach to future mapping efforts on the African continent.

In addition, a very important aspect of sharing data in databases such as the GBIF is the standards in which data are published by the producer. Such standards for data exchange are crucial in facilitating the common use of data from a variety of sources. The Access to Biological Collections Data Schema 2.06 (ABCD, 2003) is one such standard, which facilitates exchange of data. In accordance with this and other standards, research data are equipped with metadata such as position, time of collection and ownership. The ABCD facilitates the incorporation of positional information, but we failed to find specifications regarding how highly detailed locational information might be generalized to support broader distribution of potentially sensitive data. The QDGC therefore has the potential to become a critical element of emerging African spatial data infrastructure standards and could serve as a means to present and share data at a resolution not typically facilitated by GBIF.

The burgeoning area of spatial data infrastructure is another important potential application for the extended QDGC. In the context of sustainable land management, Groot (1997) defines a spatial data infrastructure as, 'a tool to facilitate access to, and responsible use of geo-information'. In an early review of eleven different national spatial data infrastructure, Masser (1999) observed that the driving forces behind spatial data infrastructures tended to fall under two important themes including the, 'growing importance of geographical information', and the need for, 'some form of government intervention to coordinate data acquisition and availability'. Given the increasing importance of understanding biodiversity issues in the face of climate change, anthropogenic disturbance and other pressures, the ability to geographically search biodiversity related data holdings will become increasingly important. While Masser (1999) suggests 'government intervention' as a driving factor and, indeed, the United Nations has several related activities under the auspices of the UNEP, FAO and others, much of the data sharing in Africa continues to occur through informal channels and in a relatively unstructured fashion. The extended QDGC provides a grassroots approach to systematically facilitate data sharing and Groot's (1997) call to facilitate access to and use of important geographic information.

The purpose of this paper was to present the extended QDGC standard and to illustrate the value of this standard as it relates to mapping biodiversity data across the African continent. The intent is not to suggest that QDGC based data sets should replace original point level data, but rather that QDGC serves as an ideal platform by which to make it known that particular datasets exist. In this way, using QDGC serves as the basis for exchange of data in a manner that allows for a wider distribution and easier access than otherwise might be the case given project sensitivities and technical constraints. By allowing for initial query and analysis using QDGC, further, more in-depth studies could be pursued based on theses developed at the coarser scale. As such, we therefore recommend the extended QDGC to be used as a standard scalable solution for exchange of biodiversity information.

Acknowledgements

This project was part of 'Biodiversity and the Human Wildlife Interface in the western Serengeti'. We thank the Norwegian Fredskorpset, Julian Blanc, Jan Ketil Rød, Grant Hopcraft, Howard Frederick, David Ericsson, Christine Mentzel, Les Underhill, Simon Mduma for valuable inputs and help, as well as TMAP, TANAPA, NCCA and TAWIRI for permissions and help with logistics, and Liz and Neil Baker for advice and input. Lastly, we thank the reviewers for their constructive comments and suggestions.

References

ABCD (2003) *Content definition: the ABCD Schema. CODATA/TDWG Task Group on Access to Biological Collection Data*. Access to Biological Collections Data. Available at: http://www.bgbm.org/TDWG/CODATA/Schema/ (accessed November 2007).

ADU (2001) *Working with grid cells: The dreaded QDGC or QDS*. African Demography Unit. Available at: http://web.uct.ac.za/depts/stats/adu/qdgc.htm (accessed September 2007).

BALMFORD, A., CRANE, P., DOBSON, A., GREEN, R.E. & MACE, G.M. (2005) The 2010 challenge: data availability, information needs and extraterrestrial insights. *Philos. Trans. R. Soc. Lond.* **360**, 221–228.

BIRCH, T.M. (1949) *Maps, Topographical and Statistical*. Claredon Press, Oxford.

BUGAYEVSKIY, L.M. & SNYDER, J.P. (1995) *Map Projections: A Reference Manual*. Taylor & Francis, Bristol, London.

CAMPBELL, K.L.I. & BORNER, M. (1995) Population trends and distribution of Serengeti herbivores: implications for management. In: *Serengeti II – Dynamics, Management, and Conservation of an Ecosystem* (Eds A. R. E. SINCLAIR and P. ARCESE). The University of Chicago Press, Chicago, IL.

CARSWELL, M., POMEROY, D., REYNOLDS, J. & TUSHABE, H. (2005) *The Bird Atlas of Uganda. British Ornithologists Club*. University of Oxford, United Kingdom.

CBD (1992) *Convention on Biological Diversity*. Convention on biological diversity (with annexes). Concluded at Rio de Janeiro on 5 June 1992. Available at: http://www.cbd.int/doc/legal/cbd-un-en.pdf (accessed February 2008).

CHAPMAN, A.D., MUÑOZ, M.E.S. & KOCH, I. (2005) Environmental information: placing biodiversity phenomena in an ecological and environmental context. *Biodivers. Inform.* **2**, 24–41.

CLARKE, K.C., DANA, P.H. & HASTINGS, J.T. (2002) A new world geographic reference system. *Cart. Geogr. Inf. Sci.* **29**, 355–362.

CURRY, R. (2001) *HydroBase 2—A Database of Hydrographic Profiles and Tools for Climatological Analysis*. Technical Reference, Preliminary Draft, November 2001. Woods Hole Oceanographic Institution, Woods Hole, Massachusetts. Available at: http://www.whoi.edu/science/PO/hydrobase/HB2_DownloadDoc.html.pdf (accessed February 2008).

Fotheringham, A.S. & Rogerson, P.A. (1993) GIS and spatial analytical problems. *Int. J. Geogr. Inf. Sci.* **7**, 3–19.

Gbif (2007) *Global Biodiversity Information Facility*. Global Biodiversity Information Facility. Available at: http://www.gbif.org (accessed November 2007).

Githaiga-Mwicigi, J.M.W., Fairbanks, D.H.K. & Midgley, G. (2002) Hierarchical processes define spatial pattern of avian assemblages restricted and endemic to the arid Karoo, South Africa. *J. Biogeogr.* **29**, 1067–1087.

Groot, R. (1997) Spatial data infrastructure (SDI) for sustainable land management. *Int. J. Appl. Earth Obs.* **3/4**, 1569–8432.

Harebottle, D.M., Smith, N., Underhill, L.G. & Brooks, M. (2007) *Southern African Bird Atlas 2. Instruction manual*. First edition June 2007. Updated November 2007. Available at: http://sabap2.adu.org.za/docs/sabap2_instructions_v5.pdf (accessed November 2007).

Harrison, J.A., Allan, D.G., Underhill, L.G., Herremans, M., Tree, A.J., Parker, V. & Brown, C.J. (1997) *The atlas of Southern African birds*. Vols 1 and 2. Birdlife South Africa, Johannesburg, South Africa.

Hostgis (2008) *HostGIS*. http://www.hostgis.com/ (accessed 7 February 2008).

Jachmann, H. (2002) Comparison of aerial counts with ground counts for large African herbivores. *J. Appl. Ecol.* **39**, 841–852.

Kimerling, A.J., Muehrcke, P., Muehrcke, J.O. & Muehrcke, P.M. (2005) *Map Use: Reading, Analysis, and Interpretation*, 5th edn. JP Publications, Madison, WI.

Kml (2008) *KML Documentation*. Available at: http://code.google.com/apis/kml/documentation/ (accessed February 2008).

Larsen, R. (2007a) *QDGC/QDS Quarter Degree Grid Cells*. Website. Available at: http://www.qdgc.org (accessed December 2007).

Larsen, R. (2007b) *Biodiversity Atlas GIS Server*. Software. Beta version. Available at: http://www.mindland.com/atlas/ (accessed November 2007).

Mark, D.M., Lauzon, J.P. & Cebrian, J.A. (1989) A review of quadtree-based strategies for interfacing coverage data with digital elevation models in grid form. *Int. J. Geogr. Inf. Sci.* **3**, 3–14.

Masser, I. (1999) All shapes and sizes: the first generation of national spatial data infrastructures. *Int. J. Geogr. Inf. Sci.* **13**, 67–84.

May, R.M., Lawton, J.H. & Stork, N.E. (1995) Assessing extinction rates. In: *Extinction Rates* (Eds J. H. Lawton and R. M. May). Oxford University Press, Oxford, United Kingdom.

Mennis, J. (2003) Generating surface models of population using dasymetric mapping. *Prof. Geogr.* **55**, 31.

Norton-Griffiths, M. (1973) Counting the Serengeti migratory wildebeest using two-stage sampling. *East Afr. Wildl. J.* **11**, 135–149.

Perring, F.H. & Walters, S.M. (1962) *Atlas of the British Flora*. T. Nelson Ltd, Norwich.

Rees, T. (2003) ''C-squares'', a new spatial indexing system and its applicability to the description of oceanographic datasets. *Oceanography* **16**, 11–19.

Robertson, M.P. & Barker, N.P. (2006) A technique for evaluating species richness maps generated from collections data. *S. Afr. J. Sci.* **102**, 77–84.

Sahr, K., White, D. & Kimerling, A.J. (2003) Geodesic discrete global grid systems. *Cart. Geogr. Inf. Sci.* **30**, 121–134.

Soberon, J. & Peterson, A.T. (2004) Biodiversity informatics: managing and applying primary biodiversity data. *Philos. Trans. R. Soc. Lond.* **359**, 689–698.

Soberon, J., Llorente, J.B. & Onate, L. (2000) The use of specimen – label databases for conservation purposes: an example using Mexican papilionid and pierid butterflies. *Biodivers. Conserv.* **9**, 1441–1466.

Treydte, A.C., Edwards, P.J. & Suter, W. (2005) Shifts in native ungulate communities on a former cattle ranch in Tanzania. *Afr. J. Ecol.* **43**, 302–311.

UN (2000) *United Nations Millennium Declaration*. United Nations. Available at: http://www.un.org/millennium/declaration/ares552e.htm (accessed February 2008).

VMWare (2008) *Building the Virtualized Enterprise With VMware Infrastructure*. Whitepaper. Available at: http://www.vmware.com/pdf/vmware_infrastructure_wp.pdf (accessed February 2008).

Walther, B.A. & Rahbek, C. (2002) Where do Palearctic migratory birds overwinter in Africa. *Dansk Ornitol. Foren. Tidsskr.* **96**, 4–8.

Wilson, J.R.U., Richardson, D.M., Rouget, M., Procheş, Ş., Amis, M.A., Henderson, L. & Thuiller, W. (2007) Residence time and potential range: crucial considerations in plant invasion ecology. *Div. Distr.* **13**, 11–22.

Yesson, C., Brewer, P.W., Sutton, T., Caithness, N., Pahwa, J., Burgess, M., Gray, W., White, R., Jones, A., Bisby, F. & Culham, C. (2007) How Global Is the Global Biodiversity Information Facility? *PLoS ONE* **2**, e1124. doi:10.1371/journal.pone.0001124

(*Manuscript accepted 18 May 2008*)

doi: 10.1111/j.1365-2028.2008.00997.x