Geo-Mar Lett (2003) 22: 181–187
DOI 10.1007/s00367-002-0111-0

Chris Jenkins

Automated digital mapping of geological colour descriptions

Received: 16 October 2001 / Accepted: 17 September 2002 / Published online: 11 December 2002

Abstract Sediment colour data are delivered by geologists as Munsell codes (Rock Color Chart) and linguistic descriptions. Using new software suitable for very large data sets, the two types can be brought into conformance and mapped together digitally. The native codes are extracted. For linguistic descriptions chromatic terms are identified with Munsell codes, then mixed in a temporary transform of psychometrically linear CIE colour space. Adjustments are made for dark/light and pale/strong modifiers. The output Munsell codes are statistically validated and mapped using special GIS legends to render them in true colour. The output displays provide a new view of marine sediment facies, comparable to remotely sensed colour imagery.

Introduction

Colour is an important character of sediments as it reflects their composition and chemical state. Changes of colour are associated with geological formations, river outfalls, organic carbon contents and reduction states, oxygen fugacity of the overlying waters, living colonisers, and the balance between terrigenous and biogenic provenance (Stanley 1969; Hamilton 2001).

When colour of sediment is described (usually soon after collection), it is either as a Munsell code expressing Hue-Value-Chroma (HVC) or as a type of word-based (linguistic) description. The former are set out in the standard Geological Society of America (GSA) Rock Color Chart (Goddard et al. 1951) – for example, 5GY 4/5 for green. The latter are largely free-form descriptions produced by field geologists – e.g. light greyish green.

Colour descriptions and codes are common place in digital marine geological data sets but hitherto have not been mappable on an automated basis. This article describes a method whereby both the codes and descriptions are used to produce digital (GIS) mappings of seafloor colour. The work is part of the dbSEABED programme for the processing of large marine geological data sets (Jenkins 1997). This style of information processing aims to mine diverse forms of data and produce a conformable, information-rich product which is useful in digital mapping, statistics, input for models and queries.

Procedure

Munsell codes

AH Munsell (1923) formalised a colour space which conveys the common perception of colours through Hue, Value and Chroma (Fig. 1). Hue is the spectral content (red, yellow, green, blue, purple), Value refers to lightness, and Chroma is the vividness or saturation. Schemes using HVC tend to conform to cultural colour vocabularies (see Roget 1852). A later Renotation Munsell colour space (Newhall et al. 1943) corrected some of the distortions of Munsell's system. This modern Munsell system is very widely used for industrial production, and it was formalised in geological sciences by the GSA Rock Color Chart which displays Munsell codes with their verbal descriptive equivalents and sample colour tablets (cf. Nickerson 1940). The three-dimensional Munsell colour space is visualised as a cylinder, with Hues arranged around the radius (colour wheel), Chroma radiating away from the axis, and Values increasing axially from base to top.

Since the publication of the Rock Color Chart, it has been commonplace to include Munsell codes among observations during marine geological research. For the colours to be mapped digitally, the codes have first to be extracted from digital renditions of the various

---

C. Jenkins
Institute of Arctic and Alpine Research (INSTAAR),
The University of Colorado, Boulder, CO 80309-0450, USA
E-mail: chris.jenkins@colorado.edu

expedition data sets and manipulated into a common format. For example, 5GY6/2 and 5-6GY6/2 (both incorrect notation), 5GY 6/2 and 5GY 6/2 to 7.5G 4/5 (both correct) are common variants. In data sets containing over 10^4 attributed sites and to which new, large data sets are continuously being added (e.g. Jenkins 1997), this task needs to be automated. For this project, the software attempts to parse the internal Munsell code format and reports faulty codes to a diagnostics file which is used for subsequent corrections. Several software error-traps are also set against out-of-range inputs and outputs.

The verified and processed Munsell codes are output by the dbSEABED software, alongside other attributes of the sediments such as rock presence, grain sizes and sorting, carbonate and organic carbon contents, physical properties, grain type and feature facies (Jenkins 1997). This allows investigation of their relationships to colour.

Linguistic descriptions

Word-based descriptions of geological materials are almost always in terms of objects, modifiers and quantifiers. Objects convey in absolute terms the value of an attribute, whether grain size, composition or colour. Colour examples are green, greenish, grey and greyish. Modifiers convey a relative meaning, a modification of the attribute, examples being light, bright and dusky. Quantifiers convey the dominance of the attribute, for example, that a sediment is mainly, occasionally or probably a certain colour. dbSEABED employs this division to parse sediment descriptions using a fuzzy set theory formalism and a thesaurus (see Jenkins 1997).

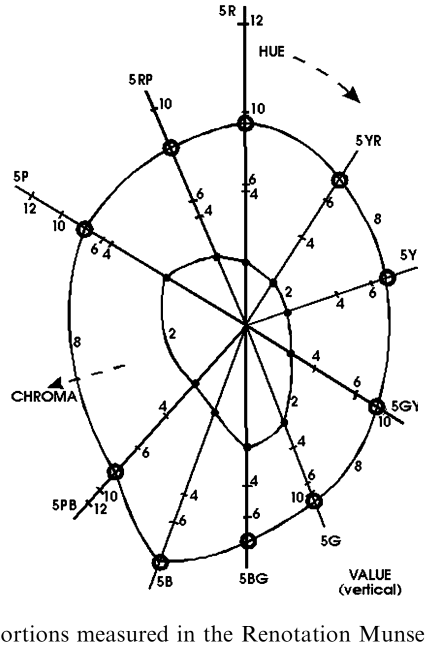
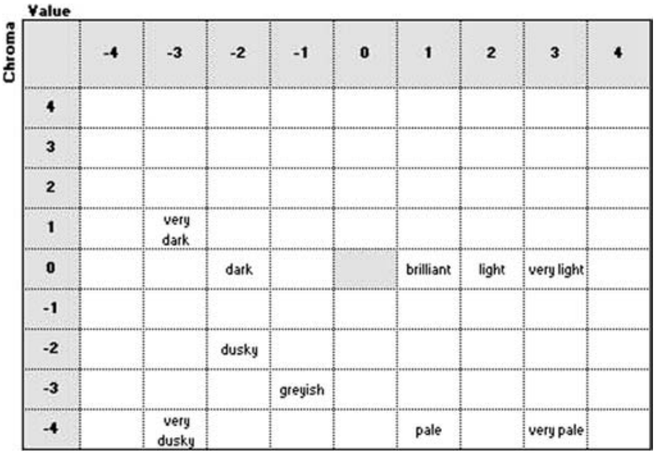


Fig. 1 Distortions measured in the Renotation Munsell by Indow and Aoki (1983). The Munsell colour space is non-linear and unsuitable as a base for performing calculations

Descriptions of colour are extracted from data sets in fields carrying the general lithological descriptions (e.g. olive green muddy sand) and from fields dedicated to colour (olive green). A typical colour description combines several chromatic terms with modifiers and can be quite complex (see Rock Color Chart). The task for a software parser of colour descriptions is to deal correctly with all the terms in a description in terms of visual perception, and then to output a useful and reliable, quantitative expression for the colour.

In order to parse colour descriptions, a model of the psychophysical meanings (see Agoston 1979) of specific terms is required. For the colour objects this is straightforward – we adopt the Munsell Hue/Value/Chroma indices (and then CIE x,y,Y; from Commission Internationale de l'Eclairage) for the terms, using the GSA Rock Color Chart. However, modifiers involve relative adjustments of Chroma and Value from a base colour. In order to deal objectively with these two concepts, Chroma and Value offsets for the terms relative to the base colours were measured using the entries in the Rock Color Chart (Fig. 2). These offsets were then entered in the parsing thesaurus, later to be used as operators by the dbSEABED software.

Unfortunately, the Munsell colour system is not suitable for the combining of colour terms, which is necessary to parse a multi-term description. It is psychometrically non-linear in Hue and Chroma (Fig. 1; Indow and Aoki 1983; Indow 1988) but it is linear in Value. The manipulations of Hues and Chroma in a parser can be performed in an alternative colour space such as CIE (see Agoston 1979; Fig. 3). CIE colour space permits linear arithmetic mixing of colours. It allows for the possibility that an output colour can be achieved by mixing more than one combination of colour terms and also deals faithfully with complementary colours (Agoston 1979). The RGB colour space is unsuitable – it does not represent all natural colours and is non-linear.



Value

Chroma

	-4	-3	-2	-1	0	1	2	3	4
4									
3									
2									
1		very dark							
0			dark			brilliant	light	very light	
-1									
-2			dusky						
-3				greyish					
-4		very dusky				pale		very pale	

Fig. 2 Offsets in Value and Chroma observed for various modifier terms in the Rock Color Chart (Goddard et al. 1951)

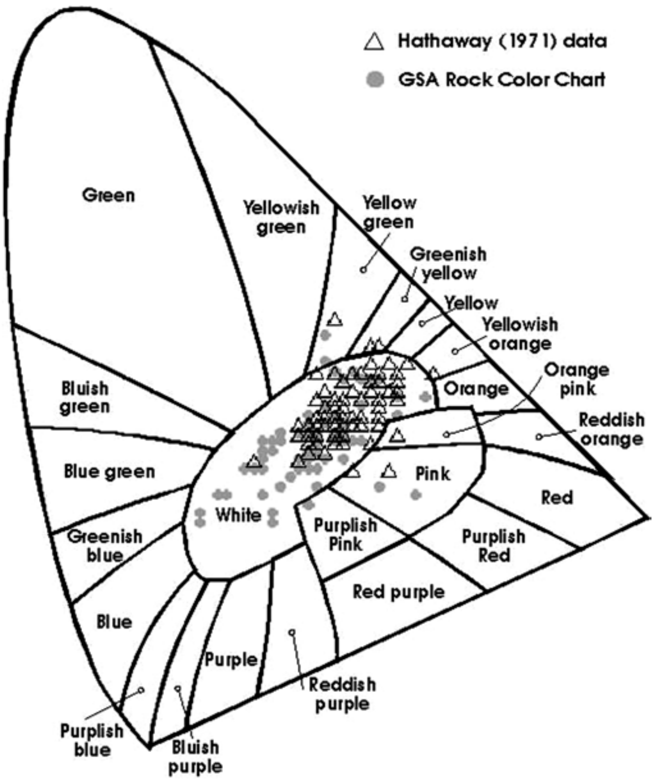


Fig. 3 CIE colour space which is based on human visual response to the colours of light. The CIE (x, y) instances of the Hathaway (1971) marine data and Rock Color Chart are plotted, projected down from their various luminance values

In detail, the calculation of colour output proceeds as follows.

1. Colour objects chromatic (with Hue/Chroma) and neutral (Value only) are each assigned a Munsell code based on the Rock Color Chart and also on calibrations performed by matching single colour terms with Munsell codes using actual marine sediment data sets.
2. The codes' Hue/Value/Chroma co-ordinates (H, V, C) are converted to CIE chromaticity (x, y) and luminance (Y) using a look-up-table of 2,379 colours which was empirically derived by Indow and Aoki (1983).
3. Weightings (l_i, m_i, n_i) to the CIE instances are calculated as follows. Terms with the suffix 'ish' (e.g. greenish) have an implicit weighting, in this case of 50%. Rear significance is the usual syntax in colour descriptions and is applied through a simple dominance table depending on the number of terms. A simple weighted mixing of CIE terms is then performed with output

$$\{\mathrm{x},\mathrm{y},\mathrm{Y}\} = \{\overline{x},\overline{y},\overline{Y}\}$$

4. The output is transferred back to Munsell colour space – (H_c, V_c, C_c) – at the colour of three-dimensional (x,y,Y/100) Cartesian closest approach in the Indow and Aoki (1983) look-up-table.

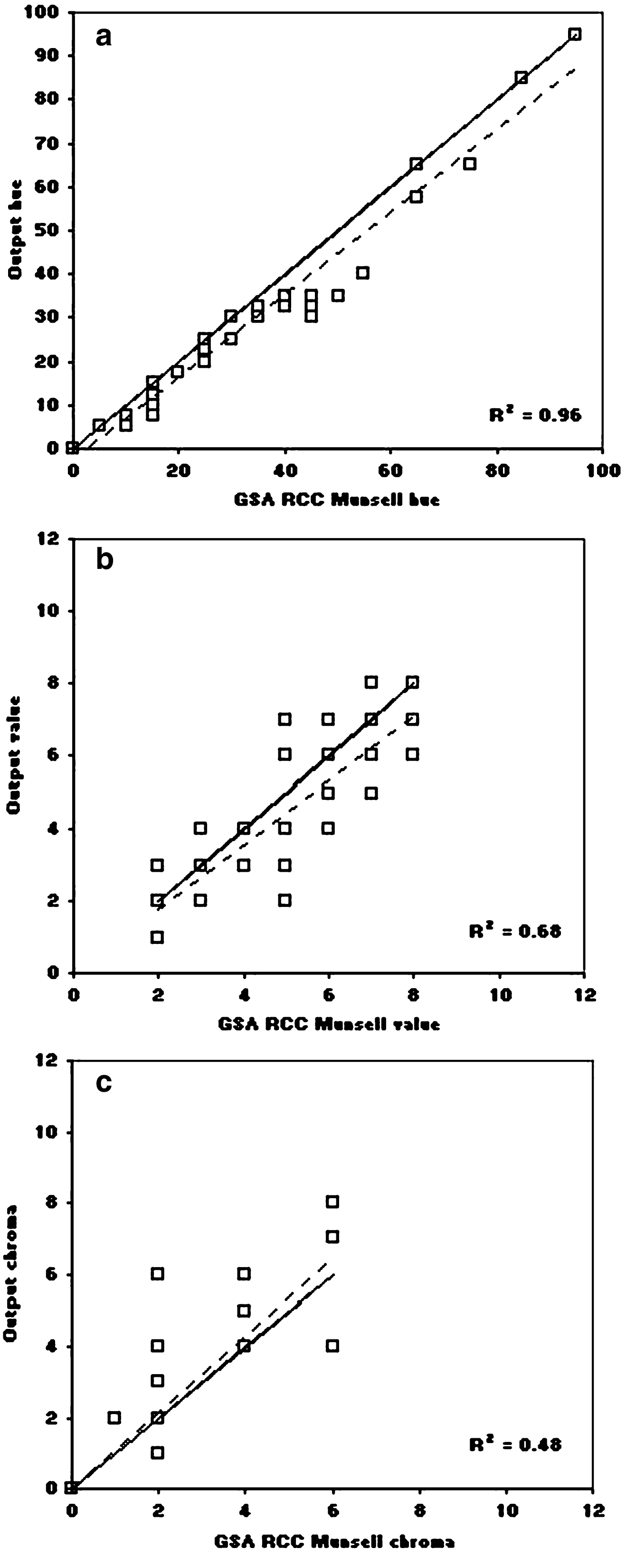


Fig. 4a–c Testing results for the parser using the GSA Rock Color Chart data set. **a** Hue, **b** Value, **c** Chroma. *Solid lines* 1:1 correspondence between input and output. *Dashed line* Linear regression between the inputs and outputs

5. Modifying terms (i.e. very pale, greyish, light, medium, deep and very dark) are given effect on both the Chroma and Value of the output Munsell code, as shown in Fig. 2.

Validation

The reliability of the processing was tested in several ways.

1. The Rock Color Chart provides linguistic colour descriptions (names) and matching Munsell codes, so the original and parsed Munsell codes can be compared statistically (Fig. 4). For Hue, Value and Chroma the linear regression R^2 statistics were 0.96, 0.68, 0.48, which is satisfactory. These R^2 values suggest that colour names are much more disciplined for Hue than in Value (grey levels) and worst for Chroma. A few colour names performed badly, for example, greyish pink (5R 8/2) which outputs too dark (5R 6/1).
2. Sensitivity tests were performed, for example, with and without rear-significance weighting of description terms, without which R^2 statistics (HVC) were 0.96, 0.67, 0.43 – marginally worse than with weighting.
3. Some large data sets of field colour descriptions carry both linguistic and Munsell code colour for each sample. In this case, however, it is difficult to use statistical analysis to rate performance of the parser because field observers tend to adopt a wide range of codes for any one colour. In the Hathaway (1971) data set of the east coast of the USA, olive corresponds to 10 Munsell codes: Hue 10YR to 2.5Y, Value 3 to 5 and Chroma 2 to 4; white to four codes: Hue 2.5Y to 10Y, Value 5 to 6, Chroma 1 to 4. A

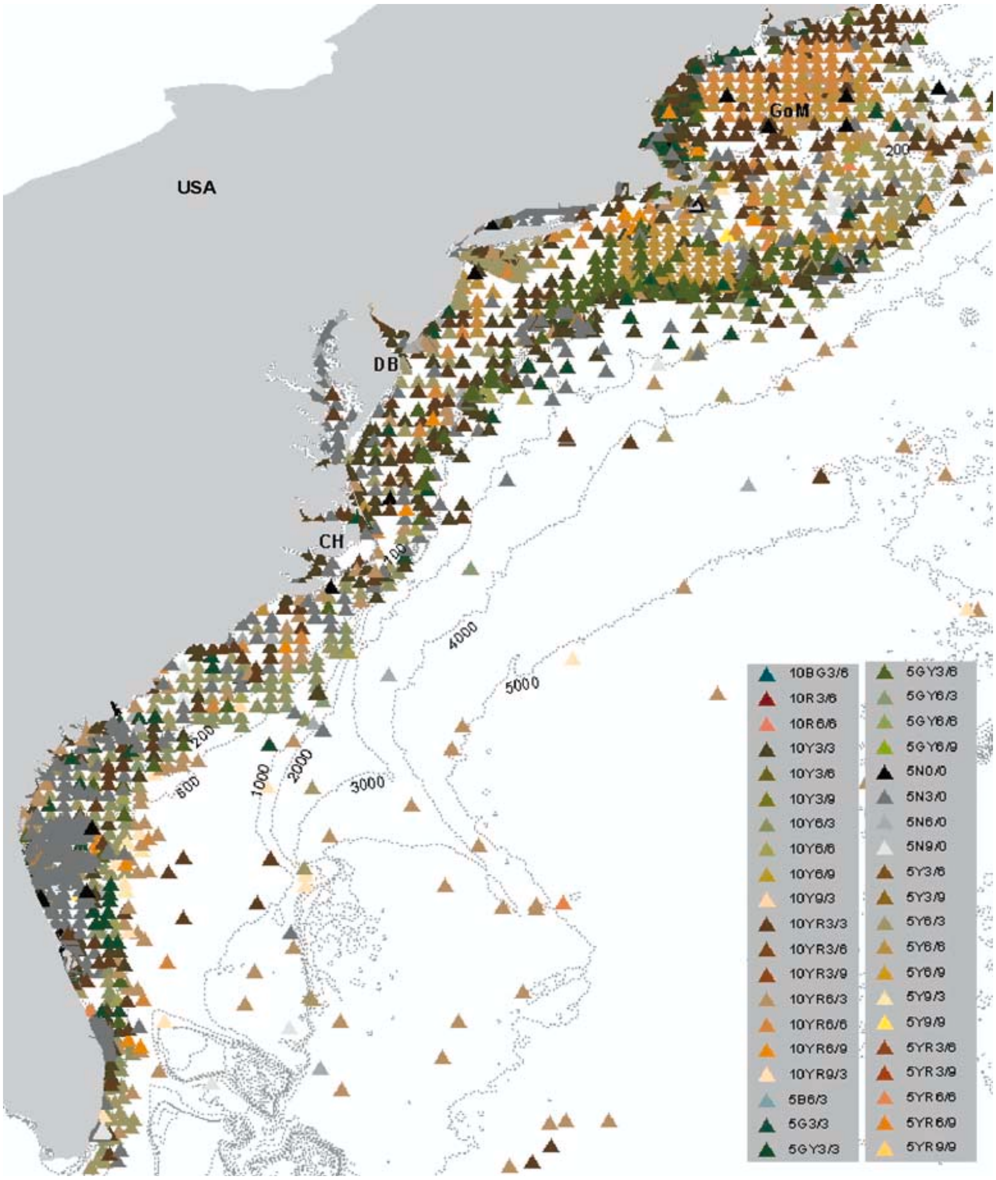


Fig. 5 Large-scale mapping of seabed colour along the US Atlantic continental margin, USA. *Ch* Cape Hatteras, *DB* Delaware Bay, *GoM* Gulf of Maine

better form of validation uses a visual check that processing outputs conform to the original intentions of the colour descriptions. In order to do this, on-screen colour squares of the output Munsell codes are generated using the CMC Munsell display tool of van Aken (1999) and then compared to the input colour names. The technique has confirmed that valid output colours are produced.

Display

In order to strike a balance between proper rendering of the colours and a practical, total number of colour variants, the output codes are rounded to increments of 5 in Hue, 3 in Value and 3 in Chroma. For example, 2.5Y 4/7 is rounded to 5Y 3/6. On data sets which have been processed to date, 125 different codes are produced, almost completely in greys, reds, browns, yellows, and greens (Hues N, R, YR, Y, GY). With rounding, this is reduced to 40 output codes of which only a few slightly exceed the Macadam limits of naturally occurring colours (Agoston 1979).

Symbol colours in the GIS legends are set to approximate the actual colours of the output in two ways: (1) by reference to the Rock Color Chart, and (2) by using the CMC programme (van Aken 1999) which can display onscreen the colour of a Munsell code.

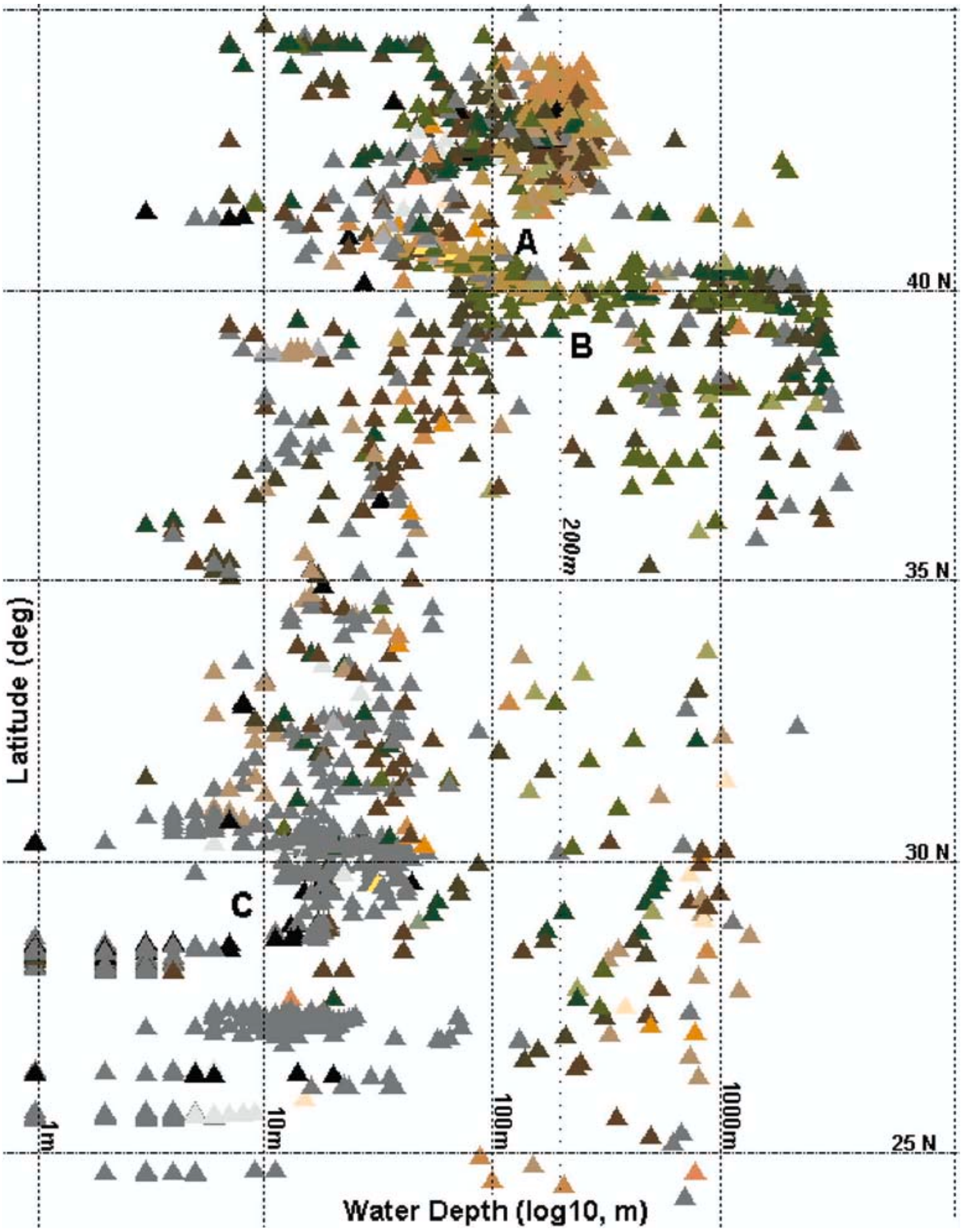


Fig. 6 Re-projection of US Atlantic continental margin seafloor colours by latitude and water depth (logarithmic scale). The visualisation is suitable for investigations of relationships between sediment colour and water mass (temperature, oxygen), and wave and current energy (legend same as for Fig. 5; for *symbols A–C*, refer to text)

Application

The procedure is now routinely applied over regions of the ocean floor where sufficient data are available. The example presented here is of the US Atlantic continental margin, for which many data sets describe the colours of seabed samples in terms of Munsell codes and word-based descriptions. The largest of the data sets is the composite set of Hathaway (1971; Poppe and Polloni 2000) but there are also many new data. The complex colour mapping which results (Fig. 5) coincides with the earlier mapping of Stanley (1969) in its generalities, but it is a digital mapping in which the spatial resolution of the input data is preserved to the final digital map display. Furthermore, since the data are digital, they can be viewed at scales from local to regional and in different co-ordinate frames.

The observed patterns of colour changes are summarised as follows (Figs. 5, 6).

1. A great deal of spatial variability is observed, implying substantial patchiness for colour. Colours which dominate in one zone are also encountered in most other geographic settings. Black, yellow and yellow-brown colours especially are few and irregularly scattered. Nonetheless, several zonal patterns of colour dominance are observed.
2. South of a transition between Cape Hatteras (36ºN) and Delaware Bay (39ºN), grey colours dominate in continental shelf sediments between 1 and 100-m water depths (Fig. 6, near site C; see Fig. 5 for locations). To the north of these latitudes, shelf sediments have greater Chroma, usually in brown (Fig. 6, near site A).
3. Inshore sediments, at depths shallower than 50 m, are most often grey (grey, olive grey or brownish grey) (Fig. 6). This includes those parts of Georges Bank shallower than 100 m.
4. Sediments of intense green colour (Fig. 6, near site B) are most common on the outer shelf and upper slope at depths of 70 to 300 m.
5. The Gulf of Maine (41–45°N) is a deep, partly enclosed basin in which relatively deep water (100 m) occurs in close proximity to the coastline. The distribution of colours deviates from patterns over open-shelf areas. The inshore sediments are dark green whereas basinal sediments are pale to medium brown.

An interpretation of the causative factors in seabed colour is not the goal of this paper. The large-scale interpretation of colour variations provided by Stanley (1969) is essentially unaltered. This includes that colour variations are only weakly and irregularly related to seabed physiography and sediment grain size. The strongest correlations appear to be to water depth and mineralogy (specifically to coloured minerals such as glauconite, dark grains), and also to dilution of colour by pale-toned carbonate materials. Some highly scattered colours such as black and yellow may have local causes such as erosion into older stratigraphy, concentrations of glacial debris, local benthic biologic productivity, and groundwater efflux.

Discussion and conclusions

This paper describes a new procedure which automates the digital mapping of seabed colour using large observational data sets (i.e. more than 100,000 attributed sites) and produces GIS displays in realistic colours. The procedure allows rapid updating of geographic coverages as new data is acquired, and allows both coded (Munsell) and linguistic (description) input data types to be plotted together, both calibrated. It preserves the spatial heterogeneity (patchiness) of seabed colour which is apparently very high in most areas.

Using the outputs, it is possible to produce digital maps of sediment and rock colour for marine and continental areas ranging in scale from local to global – wherever suitable input data exist. These digital products can be visualised and combined with other data types in novel ways. This opens up new opportunities for investigation of the dependencies between seafloor colour, sediment provenance, oceanography and biogeochemistry.

Acknowledgements My thanks go to the GretagMacbeth Inc. for making available online the software tool CMC.exe for Munsell Code display, and to NGDC for the online Hathaway (1971) digital data set. The USGS Coastal and Marine Geology section and the Royal Australian Navy kindly assisted with funding.

References

Agoston GA (1979) Color theory and its application in art and design. Springer, Berlin Heidelberg New York

Goddard EN, Trask PD, de Ford RK, Rove ON, Singewald JT, Overbeck RM (1951) Rock Color Chart. Geol Soc Am, Colorado

Hamilton LJ (2001) Cross-shelf colour zonation in northern Great Barrier Reef lagoon surficial sediments. Aust J Earth Sci 48:193–200

Hathaway JC (1971) Data file, Continental Margin Program, Atlantic Coast of the United States. 2. Samples collection and analytical data. WHOI Ref 71-15

Indow T (1988) Multidimensional studies of Munsell Color solid. Psychol Rev 95(4):456–470

Indow T, Aoki N (1983) Multidimensional mapping of 178 Munsell Colors. Color Res Appl 8:145–152

Jenkins CJ (1997) Building a national scale offshore soils database from both word-based and numeric datasets. Sea Technol 38(12):25–28

Munsell AH (1923) A color notation. Munsell Color Company, Baltimore, MD

Newhall SM, Nickerson D, Judd DB (1943) Final report of the OSA subcommittee on the spacing of the Munsell Colors. J Optical Soc Am 33:385–418

Nickerson D (1940) History of the Munsell color system and its scientific application. J Optical Soc Am 30:69–77

Poppe LJ, Polloni CF (2000) USGS East-Coast sediment analysis: procedures, database, and georeferenced displays. CD-ROM US Geol Surv Open-File Rep 00-358

Roget PM (1852) Thesaurus of English words and phrases, 3rd edn. P Haddock Ltd, Bridlington

Stanley DJ (1969) Atlantic continental shelf and slope of the United States – color of marine sediments. US Geol Surv Prof Pap 529D:1–15

van Aken H (1999) Munsell conversion, version 4.01. Program GretagMacbeth LLC, New Windsor, NY