WORLD METEOROLOGICAL ORGANIZATION

GUIDE ON WORLD WEATHER WATCH DATA MANAGEMENT

Geneva, April 1992

PART 1.
CHAPTER 1 - The WWW Data Management Concept
CHAPTER 2 -Requirements for the Development of Specific Data Management Functions
CHAPTER 3 - Meteorological Data Representation: Overview of Codes
CHAPTER 4 - Meteorological Data Representation: Alphanumeric (character) Codes
CHAPTER 5 - Meteorological Data Representation: Binary Representation Form

PART 2.
CHAPTER 6 - The Use of Computer Graphics for Meteorological Data Representation
CHAPTER 7 - Data Bases in a Meteorological Environment
CHAPTER 8 - The Distributed Data Bases Concept
CHAPTER 9 - Data Monitoring
CHAPTER 10 - Computer Software Exchange
CHAPTER 11 - Endnote
LIST OF ACRONYMS

TABLE OF CONTENTS

CHAPTER 1 - The WWW Data Management Concept

1.1 Purpose and Scope of WWWDM
1.2 Principal Long-term Objectives of the WWWDM
1.3 Programme Organization
1.4 Major Influences
1.5 Programme Components and Plans
1.6 Status of the WWWDM Activities

CHAPTER 2 -Requirements for the Development of Specific Data Management Functions

2.1 Introduction
2.2 Data Representation
2.3 Data Exchange
2.4 Data Monitoring

CHAPTER 3 - Meteorological Data Representation: Overview of Codes

3.1 Introduction
3.2 The Manual on Codes
3.3 The Numbering System of Code Forms and Binary Representations
3.4 Code Issues

CHAPTER 4 - Meteorological Data Representation: Alphanumeric (character) Codes

4.1 Introduction
4.2 Code Forms
4.3 Coding Procedures

4.4 Aeronautical Meteorological Codes
4.5 Observing Station Identification
4.6 Dates and Changes in Meteorological Codes

CHAPTER 5 - Meteorological Data Representation: Binary Representation Form

5.1 Introduction
5.2 BUFR (FM 94)
5.3 GRIB (FM 92)
5.4 Issues Raised by the use of Binary Representation Codes

CHAPTER 6 - The Use of Computer Graphics for Meteorological Data Representation

6.1 Introduction
6.2 Current Graphics Systems and Techniques
6.3 The User Interface
6.4 Standards for Meteorological Graphics
6.5 Representation of Image Data
6.6 Standard Environments

CHAPTER 7 - Data Bases in a Meteorological Environment

7.1 Introduction
7.2 Database Technology
7.3 Query Languages
7.4 Data Base Technologies Employed by Meteorological Services
7.5 Storage of Quality Control Information in Meteorological Data Bases

CHAPTER 8 - The Distributed Data Bases Concept

8.1 Introduction
8.2 The Relationship of the DDBs to the Existing WWW Infra-structure
8.3 Implementation of the DDBs
8.4 A Possible Implementation Strategy
8.5 The Role of the DDBs in Data Collection

CHAPTER 9 - Data Monitoring

9.1 Introduction
9.2 The Identification of Errors and Deficiencies within the WWW
9.3 The Role of the GOS
9.4 The Role of the GTS
9.5 The Role of the GDPS
9.6 The Role of Data Management

CHAPTER 10 - Computer Software Exchange

10.1 Introduction
10.2 Software Exchange
10.3 Software Standards

CHAPTER 11 - Endnote

LIST OF ACRONYMS

CHAPTER 1

THE WORLD WEATHER WATCH DATA MANAGEMENT CONCEPT

1.1 Purpose and Scope of World Weather Watch Data Management

World Weather Watch Data Management (WWWDM) is the component within the World Weather Watch (WWW) system which provides those support functions needed for the overall management of WWW data and products, the most economical use of the resources of the WWW system components, and for monitoring data and product availability and quality. The underlying principle in the WWWDM design is the integration of the Global Observing System (GOS), the Global Telecommunications System (GTS) and the Global Data Processing System (GDPS) facilities, services and functions into an efficient system which works as a single entity.

The concept of WWW Data Management can be illustrated through consideration of the Open Systems Interconnection (OSI) 7 layer reference model (shown in Figure 1.1). The GTS provides the functionality of the lowest 3 or 4 layers of the model (depending on the particular circuit being considered). That is, the GTS consists of physical links across which data frames may be transmitted and acknowledged, a system to determine how messages are routed, and possibly (but not always) offering host-to-host error free data transmission. The GOS and the GDPS provide the functionality of the Applications Layer.

Figure 1.1 - The OSI Reference Model

ISO	LAYER
7	Application
6	Presentation
5	Session
4	Transport
3	Network
2	Data Link
1	Physical

As already noted, the concept of WWWDM is one of carrying out those activities required to optimise the integration of the GOS, GTS and GDPS. In essence this is to be done by carefully examining the ways data are presented (or represented), and the "session level" procedures of the WWW in the light of changing technologies and requirements. Presentation Level issues include consideration of codes issues and more broadly how data are represented at various stages in their transport and manipulation. Session Level issues include consideration of how a session is established between WWW centres. Currently bilateral communication between centres (usually by mail) is used to establish that one centre holds a bulletin required by an adjacent centre which then undertakes to routinely on-forward that message as soon as possible after receipt. A limited range of ad-hoc queries (generally limited to requests for re-runs of missed data) are catered for using service messages.

WWWDM is to define and design proper procedures and interfaces, particularly in the area of data processing and telecommunications, to allow Members to obtain the coherent and appropriate sets of data and products required, despite the disparity in the levels of sophistication of technology and techniques of various WWW centers. An important objective for WWWDM is to allow for suitable interfaces, transition arrangements and support as the system components evolve, with a view towards enabling all Members to participate in the WWW at a level commensurate with their abilities and requirements, and to attain the appropriate level of technology as is necessary.

WWWDM is to provide specifications for data representation, including codes and exchange formats, guidelines for the design of data bases and storage of observational data and processed information. Standards in data representation and operational procedures will be introduced on a broad scale. The monitoring of WWW operations and the quality of basic data and output products are activities essential to the development of better Data Management (DM) practices. Regular information will be provided to Members on the status of operation of the WWW system and methods will be developed to correct deficiencies promptly.

1.2 **Principal Long-term Objectives of the WWWDM**

The principal long-term objectives of WWWDM are to:

(a) Fully integrate WWW operations and monitoring activities including methods to correct deficiencies in the WWW system;

(b) Establish common procedures for management and handling of data and products within the WWW system in order to meet effectively and efficiently Members' individual requirements;

(c) Co-ordinate and support Data Management (DM) issues for the participation of Members in the technologically advancing WWW system.

1.3 **Programme Organization**

The WMO Second Long-term Plan 1988-1997 (SLTP) introduced the WWWDM concept into the WWW Programme. It was adopted by the tenth session of the WMO Congress (Cg-X) in 1987. The WMO Congress (Cg-XI) in 1991 extended and refined this concept in the WMO Third Long-term Plan 1992-2001 (TLTP). The Commission for Basic Systems, in its ninth session in 1988 (CBS-IX), organised the programme by establishing the CBS Working Group on Data Management (WGDM) together with its Subgroup on Data Representation (WGDM/SGDR) and its Subgroup on Codes (WGDM/SGC). The latter emerged from the previously existing CBS Working Group on Codes. The Commission, through the Working Group on Data Management, is responsible for:

(a) Keeping under constant review the provision of services of meteorological data management supporting the main WWW elements (GOS, GDPS and GTS) in both real-time and non-real-time for observational data and products;

(b) Developing new or adjusting existent meteorological data management functions and services to meet new, revised or specialised requirements and to ensure that mutually compatible and internally consistent subsets of data emerge from data which are normally obtained in different manners on different time and space scales;

(c) Consolidating and co-ordinating statements received from other bodies, Members, regional associations, other technical commissions and appropriate international organizations on the need for new international forms of presentation of observational data and products within the WWW system using suitable code forms, formats and data representation forms (binary, character and graphics);

(d) Keeping abreast of the activities of the International Organization for Standardization (ISO) on matters relating to international standards;

(e) Publishing appropriate regulatory and guidance material on data management.

Resolution 5 (CBS-IX) (1988), for the first time, defined the Terms of Reference of the Working Group on Data Management which are reviewed by the regular sessions of CBS. The WGDM/SGDR deals mainly with the WMO binary data representation and graphical representation forms/standards, data base management issues and the impact of industry computer standards on meteorological graphics, while the WGDM/SGC handles all

issues related to WMO character codes. Expert meetings and seconded consultants provide the necessary input of technical expertise.

At global level, the implementation of DM functions is co-ordinated through co-ordination/implementation meetings on the Main Telecommunication Network (MTN). The regional associations play an active role through regional co-ordination/implementation meetings, through specific regional DM training events and through their respective Working Groups on the Planning and Implementation of the WWW.

1.4 **Major Influences**

1.4.1 ***Anticipated Future Needs***

The doubling time for the volume of data generated and exchanged within the WWW system is currently around five years. This is caused by the progressive use of automation in environmental observing systems, such as the surface-based remote-sensing systems Radar, Sodar, Lidar, and the automated in situ observation systems ASAP, ASDAR and drifting buoys, etc. More satellite data will be produced due to improved retrieval procedures which yield higher resolution and frequency of vertical soundings. Oceanographic satellites will add a new load of wave and surface wind data to the data traffic on the GTS. The data volume generated by numerical weather prediction (NWP) will expand sharply due to the increasing number of types of analysis and forecast products, extended forecasting periods, enlarged geographical coverage and better horizontal and vertical resolution.

The total amount of generally available data and products is expected to exceed, in many cases, the requirements of any individual Member.

Uncontrolled injection of information sets into the GTS can soon lead to an overloading of communications links and national data bases and will thus hinder the Members' timely access to those sets of information needed for the efficient performance of their operations.

The necessity to cope with the ever growing data volume and, at the same time, stringent constraints in funds will spur many Members into an accelerated cycle of employing more automation. This trend may deprive many developing countries of the benefits to be gained from meteorological improvements, because access to the necessary information requires suitable computer facilities and powerful communication links which may not be readily available and/or affordable.

1.4.2 ***Technological Advances to be Exploited***

The technology evolving for data processing and telecommunications makes it possible to design systems for the management of meteorological data which will enable better co-operation between Members who are operating at different levels of technological sophistication. The 25 Regional Specialised Meteorological Centers (RSMCs) with geographical specialisation, and four RSMCs with activity specialisation in concert with the three World Meteorological Centers (WMC), provide an ideal starting point for the co-ordination of data storage and data exchange on a higher level between each other and the efficient and effective support with data and products of all National Meteorological Centers (NMC).

The rapidly growing appreciation of international standards for telecommunications and computer technology in the meteorological community, particular for computer graphics, software management and information exchange, opens new and promising prospects for the WWW system.

A dual system of meteorological formats is available for the exchange and storage of data and products, i.e., binary data representation forms for exchanging large data volumes between automated centers, and a system of character-oriented codes for the data exchange with non-automated centers or where bit-oriented codes are not practical.

Other WMO Programmes (e.g., World Climate Programme, Global Atmospheric Water, Hydrology and Water Resources Programme, Environmental Programme, etc.) should also benefit from the basic systems for the

support of their operational requirements.

1.5 **Programme Components and Plans**

The WWWDM Concept has been set up in five major programme segments. Cg-XI approved the general policies for their realization and requested the relevant bodies of the Organization to proceed accordingly. The programme components are:

a. To specify, and to encapsulate in a clearly defined concept the global and regional Data Management requirements for the development of common procedures for handling data and products, and for procedures for monitoring the WWW operations;

   This should enable the WMO to obtain a fully integrated system of WWW operations and monitoring activities, including methods to correct promptly deficiencies in the WWW System. For instance, the store-and-forward methodology is the predominant mechanism to exchange data on the GTS, however the GDPS and other WMO programmes have begun to express the need to issue ad hoc requests for data and metadata (= data about data) which cannot easily be met by the existing GTS infra-structure. Requirements have emerged for access to new types of data and auxiliary data sets which are at present not available for exchange within the WWW system. The need for more sophisticated means of data exchange and storage has led to the formulation of the Distributed Databases (DDBs) concept. For these new services and data storage and exchange mechanisms to gain widespread, early acceptance it is essential that they comply with pertinent international standards, and will most probably be based on the principles underlying the 7-layer OSI structure;

b. To further develop or adopt suitable data representation forms for exchange and storage of data and products, including graphics, such as the WMO binary representation forms for data and products, to keep under review the existing WMO character-oriented codes, and to develop techniques for the transformation between binary data representation forms and character codes;

   This should provide an improved data representation system within the WWW that will meet the requirements for the transmission and storage of meteorological data and products both for automated and non-automated centers, and that will cope with the need for the transmission and storage of data sets which are considerably larger than those traditionally handled within the WWW. Agreed forms for the representation of meteorological data (observations, gridded products, data and products in graphical form, metadata, informational messages) need to be established;

c. To enable all Members to participate in the monitoring of the WWW system, including the activities of all designated GDPS centers, with a view to improving its overall efficiency;

   Successful performance of this activity should ensure the integrity of data from the point of creation, through its conversion to information and its long-term archival. The monitoring procedures implemented will require reviewing on a regular basis, as it may be expected that the evolution of the WWW may lead to a changing pattern of system deficiencies;

d. To provide other WMO Programmes access, as appropriate, to the WWWDM practices and principles;

   This should allow other WMO Programmes to take advantage of the WWW basic system in support of their operational requirements and will, in the longer term, help to achieve consistent data handling across the Programmes;

e. To co-ordinate the activities to implement the Data Management functions at global, regional and national levels, and to provide training events;

This should ensure that Members' individual requirements are met in an efficient and effective manner with minimal redundancy and duplication. Considering the differing levels of technology of Members' weather services, it is likely that the implementation phase will extend over a time period of many years and be an iterative process.

1.6 **Status of the WWWDM Activities**

Although WWWDM is a new concept within the WWW system several important functions and services have already been implemented. Schemes for real-time and non-real-time monitoring of the quality and availability of data and the verification of products are in place in many WWW centers, and so-called Lead Centers have accepted the responsibility for the monitoring of specific categories of observational data and for the regular feedback to Members. The DDBs Concept is beginning to take shape. Realization of the concept is expected to lead to an improvement in the overall exchange of, and access to, data. Binary representation forms are being generally accepted, but only a small number of WWW centers actually command the technology to use them to exchange data on the GTS or to employ them within their GDPS centers. The WWW is now in an open-ended transition period from character codes to binary representation forms and the first proposals are emerging for an interface between the representation forms. There is an urgent requirement to adopt a standard for the representation of image data. Automation is steadily advancing into the developing and less-developed countries allowing more efficient data handling and exploitation. The exchange of meteorological computer programs among the Members, which was re-activated by CBS-IX, supports this trend. Much more work is necessary to develop, co-ordinate and implement the data management concept as an integral part of WWW operations.

WWW DATA MANAGEMENT PRINCIPAL LONG-TERM OBJECTIVES

1. To fully integrate WWW operations and monitoring activities including methods to correct deficiencies in the WWW system
2. To establish common procedures for the management and handling of data and products within the WWW system in order to meet effectively and efficiently Members' individual requirements
3. To co-ordinate and support Data Management (DM) issues for the participation of Members in the technologically advancing WWW system

WWW DATA MANAGEMENT PROGRAMME COMPONENTS AND PLANS

1. Development of the Data Management Concept
2. Development of Codes and Exchange Formats
3. Monitoring of the Operation of WWW
4. Extension of WWWDM Principles/Practices to Other WMO Programmes
5. Implementation of the Data Management Functions and Services

---

CHAPTER 2

REQUIREMENTS FOR THE DEVELOPMENT OF SPECIFIC DATA MANAGEMENT FUNCTIONS

2.1 Introduction

The role of DM is to develop a set of requirements that will ensure that functions and services related to DM within each WWW component are adequately provided. The functions to be provided naturally fall into one of

several categories; data representation, data exchange, and data monitoring. Each of these categories is considered in turn in this chapter.

2.2 Data Representation

The first function of DM is to establish agreed-upon forms for the representation of meteorological data; observations, observational metadata, gridded products, image products, and informational messages. Much effort has long been devoted to some of these functions, while others have received less attention. For example, alpha-numeric representation forms for observational data have been established and maintained for decades, while binary forms for observational data have existed only for a few years. At the other end of the spectrum, there is as yet no international agreement on a standard form for vector graphic and image data. Although the binary forms are sufficiently general to describe all meteorological data and remain our goal for the future, not all centers can handle the binary forms as yet. Thus, both binary and alpha-numeric representation forms will have to be maintained for the foreseeable future, and the issue of transformation between binary and alpha-numeric forms will remain with us as well. The following requirements may therefore be recommended:

Existing Requirements:

a. Maintain the alpha-numeric forms for the representation of meteorological observations;

b. Maintain and enhance the binary universal form for the representation of meteorological observations (BUFR);

c. Maintain the alpha-numeric forms for the representation of gridded meteorological fields;

d. Maintain and enhance the binary form for the representation of gridded meteorological fields (GRIB);

e. Centres continue to meet their requirements to exchange basic observational data in a format which ensures maintenance of existing services. BTAB should be re-examined with a view to increasing its efficiency, both generally, and particularly with regard to reducing the length of a BTAB message in order to use this form for new data only available in binary representation form. During the transition period between character and binary representation, the requirement for products may need to be met with some duplication of products in each of these formats.

New Requirements:

a. Establish the policy that all meteorological data should be exchanged in the appropriate binary form, and such forms be self-defining, whenever possible;

b. Expand BUFR to permit encoding of observational metadata;

c. Establish a mechanism to permit the encoding of information necessary to classify characteristics of the observing platforms;

d. Expand the adoption of standards for the representation of vector graphic and image data beyond facsimile group 3T4;

e. Establish a means for the identification of the data contents (as opposed to the addressing) of messages.

2.3 Data Exchange

The best approach to accomplishing the rapid distribution of meteorological observations from the GOS to the GDPS and routinely disseminated analysis and forecast products from the GDPS to users, is the store-and-

forward methodology. This approach is currently used on the GTS. In this methodology, the GTS can be viewed as one continuous circuit with many nodes. Each node receives all its information from the previous node, stores the information it receives, and then sends the information required by the next node to it according to the instructions stored in the node's switching directory.

On the other hand, the GDPS has begun to express the need to issue ad hoc requests for meteorological observations, observational metadata, and new forecast products from NWP centers and send ad hoc notifications of erroneous observations to the observational sites. These ad hoc transmissions are not well served by the store-and-forward switching directory methodology because by their very nature the GTS nodes are unable to efficiently respond to ad hoc messages. Rather, ad hoc requests are best accommodated by request/reply mechanisms.

Presently only re-transmission of previously sent messages are permitted on the GTS. The re-transmission function is limited to messages that previously were sent over the link and does not satisfy the ad hoc request/reply need. A different mechanism is needed so that the GTS node can recognize an ad hoc request, satisfy the request from its own limited database, if possible, or forward the ad hoc request to the associated GDPS facility for a response.

In order to maintain current functions and add the needed new ones, certain requirements must be satisfied by the GTS:

Existing Requirements for the GTS:

a. The GTS should continue to deliver observational data as quickly and reliably as possible;

b. The GTS should continue to deliver routinely distributed processed data as quickly and reliably as possible;

c. The GTS should prevent unauthorized access to the communication facilities and to the databases;

d. The same information should not be sent in more than one code form on the same communication link, unless it cost-effectively suits centers' requirements under bilateral or multilateral arrangements.

New Requirements for the GTS:

a. The GTS should establish and maintain a comprehensive catalogue of meteorological messages for observational data and processed information, (an improved WMO-Publication No. 9, Vol. C) updated in near real-time;

b. The GTS should develop a comprehensive mechanism to route ad hoc requests to the appropriate center (GTS or GDPS) for response. Although it is the responsibility of the GTS to develop the request/reply mechanism, it is the responsibility of the WGDM to develop the specific request/reply message format in co-ordination with the WG/GTS;

c. Standard interfaces for queries to and replies from databases should be adopted with due consideration of international standards;

d. A mechanism for the transmission of information on the status of a GDPS or other appropriate center's processing and database should be established;

e. A mechanism for the transmission of information on the contents of a GDPS or other appropriate center's database should be established;

f. Transmission of larger volumes of information and larger message sizes should be supported. The GTS should consider alternative types and levels of service that would be required for several ranges of traffic volume in the future;

g. The increased traffic volume must not be allowed to adversely impact the current rapid distribution of observational data. The GTS should consider the implementation of a flow-control mechanism to ensure the timely delivery of observational data and the non-interference by the transmission of products and replies to ad hoc requests.

However, these requirements are not sufficient. Before the request/reply capability can be used to satisfy the needs expressed by the GDPS, both the GOS and also the GDPS must satisfy certain requirements as well:

New Requirements for the GOS:

a. The GOS should liaise with Data Management concerning the data representation forms to be adopted by new observing systems;
b. The GOS should establish and maintain a comprehensive catalogue of meteorological observation stations and schedules (an improved WMO-Publication No. 9, Vol. A) updated in near real time;
c. The GOS should establish and maintain the capability to recognize and respond to notifications sent from the GDPS regarding erroneous meteorological observations;
d. The GOS should establish and maintain the capability to recognize and respond to requests for retransmission of meteorological observations.

New Requirements for the GDPS or Other Appropriate Centers:

a. The GDPS or other appropriate center should establish and maintain database catalogues;

b. The GTS nodes have limited storage capacity to support their store-and-forward capability, typically equivalent to a volume of data ranging from one hour to a few hours. The GDPS or other appropriate center should therefore establish and maintain the capability to recognize ad hoc requests and issue replies for information that does not exist at the GTS nodes;

c. The GDPS or other appropriate centers responsible for maintaining databases should be able to respond to requests for large portions of their databases in the form of large files for efficient exchange between GTS centres.

2.4 Data Monitoring

There are two distinct aspects of data monitoring; data quantity monitoring and data quality monitoring. Data quantity monitoring is concerned with the timely receipt of observations and products, while data quality monitoring is concerned with the "goodness" of the observations and products. In organizations where the responsibility for communications is separate from the responsibility for processing and product creation, the two aspects of data monitoring are similarly split. Typically, the communication function is concerned with the quantity of data while the computation or processing function is concerned with the quality of the data. In the WMO context of a node on the GTS, the Regional Telecommunication Hub (RTH) is typically concerned with quantity of the data while the National Meteorological Center (NMC) is typically concerned with the quality of the data.

Data quantity monitoring has many different forms depending primarily on the timing of the monitoring information that is collected and the timing of the analyses performed on the collected data. The various forms of monitoring fall into four basic categories:

1. Real-time Collection, Ad Hoc Reporting - General traffic logging - Receipt of one message type;

2. Near Real-time Collection, Near Real-time Reporting:

   Hourly observations;
   Aviation surface observations;
   Radar observations;
   Local and regional weather summaries;
   Synoptic observations;
   Surface synoptic observations;
   Upper-air synoptic observations;
   Unscheduled observations;
   Climate observations;

3. Near Real-time Collection, Retrospective Reporting:

   WMO annual monitoring program;
   Ad hoc analyses of data availability;

4. Retrospective Analysis:

   Time histories;
   Histograms by hour of the day or day of the week;
   Seasonality comparisons;
   Year-to-year comparisons;
   Find what data is lost;
   Many others.

Data quality monitoring is performed at the observation sites and at the NMCs. The observation sites monitoring programme is performed to ensure that "gross" errors are eliminated from data transmitted on the GTS. Data quality monitoring is typically performed at an NMC for two purposes. First, it is desirable to insure that only correct observed values are used in data disseminated to field forecast offices or to the general public. Incorrect observations not only distract from the quality of plotted charts, they can distort contours that might have been generated for the display by objective methods. Such problems attract attention and can be quite noticeable. Second, it is not only desirable, but crucial to insure that only unique, high quality observations are submitted to objective analysis schemes whose task it is to provide initial conditions for numerical prediction models. Such data problems are not seen by the forecast offices or general public, for they are only manifested through a degraded objective analysis and ensuing forecast. Since numerical prediction models have deficiencies unrelated to incorrect data that also lead to incorrect forecasts, it can be difficult to determine whether a forecast problem was due to a data or model deficiency, even for personnel of the NMC.

Because of its importance, much attention is paid to data quality monitoring at a typical NMC, and a wide variety of checks have been developed to insure the highest quality observational database is prepared. These checks fall into one of several categories:

- Elimination of duplicates;
- Internal consistency checks;
- Comparisons with climatology;
- Comparison with a background field;
- Comparisons with other observations;
- Manual monitoring.

Parts of a rather comprehensive monitoring program are thus already in place. The task for DM is to:

- Ensure that this monitoring program is complete and consistent throughout the WWW;

- Facilitate the collection and exchange of monitoring information;

- Establish a system for in-depth analyses of the monitoring information;
- Disseminate the analyses of monitoring results to the appropriate WMO Department/Division.

In order to accomplish these objectives, we are led to a number of requirements:

- The ensemble of monitoring procedures should cover the entire life-cycle of all meteorological information, from the taking of an observation to the transmission of a product to users;
- Each individual monitoring procedure should be standardized with respect to the quantity being monitored, the time it is monitored, and the processing level at which it is monitored;
- The monitoring procedures should be performed on a regular basis, and results transmitted to a central collection center in a timely manner. Automated means of performing and transmitting the results are desirable;
- Central collection center(s) should be established that are staffed with people who have both the time and knowledge to adequately analyze the monitoring results;
- The analyses of the monitoring results should be promptly transmitted to the WMO Secretariat for appropriate action.

CHAPTER 3

METEOROLOGICAL DATA REPRESENTATION: OVERVIEW OF CODES

3.1 Introduction

An important aspect of data management for the WWW is to establish common procedures for data representation, i.e. "character codes" and "binary representations". Such procedures serve to facilitate the timely national and international exchange of the vast volume of meteorological, geophysical, and environmental information required by individual Members to meet their specific operational responsibilities. National and international research and application programmes are also served by the availability of observations in common data forms.

These common procedures for international data representation are based on the concept of using codes to describe weather conditions, reports of instrumental readings, and processed data, thereby considerably reducing the length of messages, avoiding language problems and facilitating automatic processing.

Coded bulletins are used for the international exchange of meteorological information comprising observational data provided by the WWW Global Observing System and processed data provided by the WWW Global Data-processing System. Coded messages are also used for the international exchange of observed and processed data required in specific applications of meteorology for various human activities and for exchanges of information related to meteorology.

Messages may take the form of either a set of code forms, defined by standard procedures, for alphanumeric data exchange (character codes) or a set of representation forms with their specifications and associated code tables for binary data exchange (binary representations). Rules concerning the selection of codes and representation forms for international exchange are specified in the WMO Technical Regulations, Volume I, Chapter A.2.3 (WMO Publication No.49, Ed.1988).

3.2 The Manual on Codes

The code and representation forms are described in the Manual on Codes (WMO Publication No. 306), which comprises Volumes I (Parts A and B), and II, dealing with international and regional codes respectively.

The Manual on Codes, Volume I - International Codes - consists of Part A (Alphanumeric codes) and Part B (Binary representations). Together they form part of the Technical Regulations and are referred to as Annex II to the Technical Regulations. They contain international regulatory material stemming from recommendations of the WMO Commission for Basic Systems (CBS) and decisions taken by the Executive Council and the President of WMO.

National practices regarding the coding of certain elements in reports, analyses, or forecasts for international exchange are included in Volume I as an appendix. In addition, Volume I contains three attachments, printed on yellow paper, which are included for information only and do not have the status of WMO Technical Regulations.

The Manual on Codes, Volume II - Regional codes and national coding practices - consists of seven chapters, six of which are devoted to the six WMO Regions, and the seventh to the Antarctic.

Volume II is not part of the WMO Technical Regulations. It contains regional procedures for the use of international code forms as well as regional code forms which are intended only for exchanges within a given WMO Region, as formally adopted by the regional association concerned. Procedures and code forms for use in the Antarctic, are adopted by the WMO Executive Council on the advice of the EC Working Group on Antarctic Meteorology.

Volume II also contains national coding practices relating to the use of international or regional codes and information on national code forms which might be of interest to other countries. Some special codes which are used in messages exchanged over the WWW Global Telecommunications System circuits, which include ice coverage and satellite ephemeris, are included in Volume II as an appendix.

3.3 The Numbering System of Code Forms and Binary Representations

Each international code form and binary representation bears a unique number from 10 to 99, preceded by the indicator letters FM. This number is followed by a Roman numeral to identify the sessions of CBS (or CSM prior to 1974) which either approved the code form as a new one or made the latest amendment to the previous version.

A code form approved or amended by correspondence after a session of CBS receives the number of that session.

An FM number is allotted to an international code form or binary representation in accordance with the following categories:

10 to 29	code forms for surface reports
30 to 43	code forms for upper-air reports
44 to 59	code forms for analyses and forecasts
60 to 69	code forms for oceanographic and hydrological reports and forecasts
70 to 79	code forms for climatological reports
80 to 89	code forms for atmospherics and satellite reports
90 to 99	binary representation forms (reports, analyses, forecasts)

For regional code forms a different numbering system is used. It starts with the indicator letters RF and a number 1 to 6 corresponding to the WMO Region in which the code form may be used (or the number 7 if the

code form is intended for use in the Antarctic), separated by a solidus (/) from a two-figure number, allotted in numerical order from 01 upwards for each WMO Region or the Antarctic.

This numbering enables the code forms and binary representations to be distinguished from one another and from the code tables, which are numbered with a four-figure number for international character codes and a three-figure number for regional character codes.

Furthermore, an indicator term consisting of one or a maximum of two words (preferably one word of no more than five letters) is used to designate the code form or binary representation colloquially and is therefore called a "code name". This abbreviated code name should, as far as possible, succinctly reflect the basic purpose of the code. In some cases it is included as a symbolic prefix in the code or representation form to assure identification in transmissions, e.g. CLIMAT or GRIB.

A reference list of the international code forms and binary representations, together with the corresponding FM numbers and the full code names (reflecting in a comprehensive manner the basic purpose of the code), is included in Volume I of the Manual on Codes. This list also indicates the decisions of the Executive Council or the President of WMO relevant to the particular alphanumeric code or binary representation.

A listing of regional code forms, together with the corresponding RF numbers and the full code names, is included separately in the chapters devoted to the particular Regions, or the Antarctic, in Volume II of the Manual on Codes.

3.4 Codes Issues

As noted earlier, the WWW is now in an open ended transition period that began with the first use of a binary format for data transmission and will end when the last message is transmitted in a character code format, if ever. The WMO recognized the continued need for character codes for an unspecified period of time, stating in its long-term plan: "A dual system of meteorological formats will be used for the exchange and storage of data and products, namely bit-oriented codes for exchange of large data volumes between automated centres, and a system of character-oriented codes for the data exchange with non-automated centres where bit-oriented codes are not practical."

There is a potential demand for some centres to ingest BUFR (binary) messages and then reformat them back into (current) WMO character code forms for local transmission on character lines. Before undertaking such a course of action serious thought must be given to the purposes of such retransmissions. If they are undertaken to supply non-automated centers with "human-readable" information, then it makes more sense to drop the use of the sometimes complex WMO codes and instead send the information out in a simple tabular form, provided, however, that the telecommunications lines and receiving equipment can accommodate the larger volume of traffic that would result. Eventually, the adoption of suitable tabular formats would eliminate the need for training humans to decode the sometimes obscure WMO code forms. Programming to generate tables, for transmission and display, is much simpler than generating the current WMO character codes.

CHAPTER 4

METEOROLOGICAL DATA REPRESENTATION: ALPHANUMERIC (CHARACTER) CODES

4.1 Introduction

This chapter provides a detailed description of the alphanumeric codes used to transmit meteorological data. The material presented is not intended to supersede the Manual on Codes (WMO Publication No. 306) but rather provide an explanation of how these codes have evolved to their present formats and the nature of the inter-relationships between the various codes. This chapter also consolidates in a single location references to the source material required by a practitioner intending to use these codes in an operational centre.

4.2 Code Forms

4.2.1 Composition

Each alphanumeric code is designed to provide a particular type of information, which may be data obtained from an observing program or system as in SYNOP and TEMP, processed data in the form of statistics as in CLIMAT, processed data in the form of an analysis or forecast as in IAC and GRID, or forecast information as in TAF and ARFOR.

Coded bulletins may be divided into well defined Parts and/or Sections containing appropriate components of the information and identifying information included for regional or national use only. In the Manual, Regulations contain standard coding procedures to be observed and Notes provide additional guidance on the use of the code.

The code forms are designed to be as brief as possible, commensurate with relatively easy manual or automatic encoding and decoding. Unnecessary repetition of data, or the unnecessary inclusion of indicator or control groups or letters is avoided. Only letters from the Latin alphabet, Arabic numerals, and the solidus (/) are used in code forms. The ellipsis (...) may be used in the description of code forms to indicate omitted groups but will not be found in the transmitted codes themselves.

4.2.2 Symbolic Words and Figure Groups

Symbolic words, symbolic figure groups, and code groups, including those which are included in the report only in certain circumstances, are placed in an appropriate position in the code form, so as to avoid the risk of confusion during transmission and to obviate the possibility of ambiguity during decoding. In accordance with WMO Technical Regulation [A.2.3.]1.2.2, symbolic words, groups, and letters (or groups of letters) required for regional or national purposes only are selected so as not to duplicate those used in international code forms.

Symbolic words and symbolic figure groups are used as code names, symbolic prefixes or indicator groups. Symbolic words (and indicator letters) are made up of upper case letters only. Whenever symbolic words and groups in regional codes are already used in international codes, they retain their international character.

Some code names, e.g. SFAZI, have a twofold use:

- As true code names, i.e., as a convenient means of designating each code form without referring to its FM or RF number;
- As a symbolic prefix to the report, analysis or forecast, when necessary, to provide ready identification of the code form.

Code words are included in a coded report, analysis, or forecast either as symbolic prefixes to identify particular information, e.g. ICE, NORMAL, or in place of certain code groups when specified conditions occur, for instance the code word CAVOK.

The meaning of symbolic words, figures, and groups used in international codes, together with reference to the particular code forms in which they are used, is contained in Volume I, Part A of the Manual on Codes. The meaning of symbolic words, figures, and groups used only in a WMO Region or the Antarctic, together with references to the particular code forms in which they are used, is contained in the chapter devoted to that Region or the Antarctic in Volume II of the Manual on Codes.

4.2.3 Code Groups

Code groups may be described solely with symbolic figures, or symbolic figures and symbolic letters (or groups of letters). In the latter case the symbolic figures are used to indicate that particular information follows. The symbolic letters (or groups of letters) in the code groups represent meteorological or other geophysical elements.

In messages, these symbolic letters (or groups of letters) are transcribed into Arabic numerals indicating the value or the state of the elements described in accordance with the standard coding procedures contained in the relevant regulations and specifications.

Code groups put in round brackets are optional items that may or may not be included in the report, analysis, or forecast, depending on specific conditions. The absence of round brackets means that the inclusion of the groups concerned is determined by international or regional decision and is mandatory; these decisions are indicated in the regulations included with the description of each code form. When a code group, or a number of code groups, is used more than once under certain circumstances, it (they) are followed in the code form by a group (groups) of dots.

Except in aviation and radiological reports, analyses, or forecasts (see paragraph 4.4), code data groups are generally composed of not more than five characters, either figures or upper case letters. These limitations stem from historical ITU Telegraph Regulations for traffic subject to charge per word, in which code data groups in excess of five characters were counted as two (or more) words and that the mixed use of figures and letters in one code data group was prohibited. In 1981 these restrictive ITU Telegraph Regulations and the related CCITT Recommendations were eased to allow single "word" code data groups of up to 10 characters. This prompted the WMO Commission for Marine Meteorology to recommend the use of ten-figure groups for ship-to-shore transmissions of ships' weather reports (see WMO Publication No. 386, The Manual on the GTS, Attachment I-1, paragraph 2.2).

4.2.4 Parts and Sections

Code forms are built up from a number of well-defined components, each comprising a different type of coded information. Components which can be transmitted as a separate report are called parts and carry special identification groups. Code forms, or their parts, can be further divided into sections which must not be transmitted separately.

The first section of a code form, or its parts, should contain at least identification data (indicating the type of coded information) and begin with the code MiMiMjMj or the abbreviated name of the code. Position and time groups may, depending on the nature of the report, be included in the first (identification) and/or in the subsequent section(s).

Each subsequent section begins with one or more symbolic indicator figure(s) or an indicator group of at least two figures, preferably in some logical order. Sections which may be omitted from the report under certain conditions are placed in round brackets. In order to facilitate automatic processing, code forms do not normally provide for any written text unless in a separate section.

4.2.5 Notes to the Code Form

Brief explanations of the code form are included in a number of Notes under the code form, usually indicating the following: the parts and sections comprising the code form, what kind of data are to be transmitted in these parts and sections, as well as explanations and special features of the code form.

4.3 Coding Procedures

4.3.1 Regulations

Regulations governing coding procedures and Notes on the use of the code are an essential part of the code form. Regulations indicate the circumstances under which particular parts, sections, code groups, and words should (or should not) be included in the report, analysis, or forecast. Regulations also contain procedures regarding the standard practice in the transcription of symbolic letters (or groups of letters) into figures indicating the value, or the state of, the element described.

For international codes, these regulations carry the same status as the WMO Technical Regulations. Consequently, the standard coding procedures are distinguished by the use of the term "shall" in the English text, and by suitable equivalent terms in the French, Spanish and Russian texts.

A similar practice is followed in regional coding procedures with regard to international code forms as well as in coding procedures related to regional code forms, although here the word "shall" in the English text (and the equivalent term in the French, Spanish and Russian texts) has its dictionary meaning and does not have the regulatory character as in the WMO Technical Regulations.

Notes providing additional explanation, information examples or cross-references, may be appended to the regulations. The word "shall" is not used in these notes.

Where national practices do not conform with the standard international or regional coding procedures, Members concerned shall formally notify the Secretary-General of the WMO for the benefit of other Members. Information on these practices is including in the WMO Manual on Codes, Volume II.

4.3.2 Specification of Symbolic Letters

Symbolic letters (or groups of letters) are made up of either lower case or upper case letters, with or without subscript and/or superscripts. Whenever possible, the symbolic letters (with their subscripts) suggest the element or phenomenon being described. Subscripts may be made up of either lower case or upper case letters or figures, or any combination thereof but with a maximum of two characters. The use of superscripts is (generally) limited to the apostrophe (') to indicate an element or phenomenon which is (or was historically) related to another element or phenomenon indicated with a similar symbolic letter (or group of letters) but with some difference in its observing and/or reporting practice, e.g. V_S - Visibility seawards (from a coastal station), and V'_S - Visibility over the water surface of an alighting area. Superscripts are also used in code forms to indicate the repetitive use of certain code data groups or data lines, allowing the message, when in printed form, to have the characteristics of a table, e.g., the WINTEM message. In this case the superscripts used in the first code data group or data line(s) is the figure 1, in the subsequent data group or data line(s) it is the figure 2, to be followed by a group or line of dots and completed with the last code data group of the data line and/or the last code data line(s) with superscript lower case letters i,j,k,m or n, as the case may be.

The transcription of a symbolic letter (or group of letters) used in code forms into figures requires its unambiguous specification, indicating the meaning of the symbolic letter (or group of letters) and the units in which the element concerned is to be coded and/or reference made to the code table for the element or phenomenon in question. Only those units of measurement which have been approved by the WMO Congress are used in the codes.

In some cases the specification of the symbolic letter (or group of letters) is sufficient to permit a direct transcription into figures. However, when the number of significant figures of this value (expressed in the units given in the relevant specification) is lower than the number of symbolic letters reserved for this element, one or more zeros, as appropriate, must be inserted at the left of the significant figure(s) of the reported value.

Additional explanations, references, and/or standard coding procedures relating to the specification concerned are added, where appropriate, to the specification in the form of notes. Notes indicating standard coding procedures are distinguished from other notes by the use of the word "shall" in the English text, and by suitable equivalent terms in the French, Spanish and Russian texts (see paragraph 4.3.1).

A similar practice is followed in coding procedures related to specifications of symbolic letters (or groups of letters) used in regional code forms, although here formally the word "shall" in the English text (and the equivalent term in the French, Spanish and Russian texts) has its dictionary meaning and does not have the regulatory character as in the WMO Technical Regulations.

Clearly the same symbolic letter (or group of letters) cannot be used for specifying different types of information in one code form and, preferably, should have a unique definition for all code forms. However, in some cases the

same symbolic letter (or group of letters) is used for different types of information in different code forms.

The specifications of symbolic letters (or groups of letters) used in international codes, together with reference to the code table for the element or phenomenon in question, where appropriate, and references to the particular code form (or group in that code form) in which these symbolic letters (or groups of letters) are used, is contained in Volume I, Part A of the Manual on Codes.

4.3.3 Code Tables

Where symbolic letters (or groups of letters) represent non-numeric coded information, code figures are required, the specifications of which are given in special code tables for each element or phenomenon.

Each code table bears a unique number. Code tables corresponding to symbolic letters (or groups of letters) used in international character codes are numbered with four figures from 0100 up to 5299 and allotted in the alphabetical order of the symbolic letters (or groups of letters) in accordance with the following scheme:

The first two figures represent the number of the main letter of the symbol in alphabetical order, where upper case letters are given an odd number, and lower case letters an even number: 01 for A, 02 for a, 03 for B, 04 for b ... 51 for Z and 52 for z;

The last two figures are allocated in accordance with the following scheme:

00 to 01 are reserved for code tables corresponding to a symbol composed of one letter only (X or x, for instance);

02 to 30 are reserved for code tables corresponding to symbols of the forms X_A to X_Z, x_A to x_Z and derived symbols such as X_{Ao} or x_Ao;

31 to 60 are reserved for code tables corresponding to symbols of the forms X_A to X_Z, x_A to x_Z and derived symbols such as X_{Ao} or x_Ao;

61 to 70 are reserved for code tables corresponding to symbols of the forms X_o to X_n or x_o to x_n, n being any number;

71 to 99 are reserved for code tables corresponding to symbols of the forms X', XX, XXX, x',xx, xxx or any similar forms such as X_bX_b, $X_oX_oX_o$, x_bx_b, $x_ox_ox_o$.

The numbers attributed to code tables corresponding to symbolic letters (or groups of letters) used in international character codes, are given in a table preceding the relevant individual code tables, which are included in Volume I, Part A of the Manual on Codes in accordance with WMO Technical Regulation [A.2.3.]1.3.1.

Code tables related to symbolic letters (or groups of letters) for regional use only, are numbered with three figures ranging from 120 to 799 as follows:

Code table number	Reserved for use in
120 to 199	WMO Region I
220 to 299	WMO Region II
320 to 399	WMO Region III
420 to 499	WMO Region IV

520 to 599	WMO Region V
620 to 699	WMO Region VI
720 to 799	Antarctic

Within the range of code table numbers available for a Region or the Antarctic, individual regional code table numbers are allocated in the alphabetical order of the symbolic letters (or groups of letters) used in that particular Region or the Antarctic, although due to the limited numbers (80) available for allocation in each Region or the Antarctic, no rigid scheme has been developed for this purpose.

The numbers attributed to code tables corresponding to symbolic letters (or groups of letters) used only in a given WMO Region or the Antarctic, are listed in a table preceding the relevant individual code tables in the chapter devoted to the Region concerned or the Antarctic in Volume II of the Manual on Codes.

The code tables indicate the symbolic letter (or group of letters) with its specification, followed by a listing of code figures with their corresponding specifications, i.e. value of elements, type and/or state of phenomenon, etc. When appropriate, explanatory notes may be added to a code table. In principle, these notes should not contain regulations for the coding of phenomenon or elements.

All possible code figures are included in the tables. Code figures not presently used but available for future allocation are labeled with the word "reserved". Code figures which are not used and for technical reasons are not available for future allocation, are labeled with the words "not used". Code figures in some tables are supplemented, when appropriate, with a solidus (/), usually to indicate that for some reason particular information is not available or is undetermined. Otherwise the solidus is used in transmissions to indicate missing data.

4.4 Aeronautical Meteorological Codes

Aeronautical meteorological codes (METAR, SPECI and TAF) have been developed, in conjunction with the CAEM and the ICAO, from specific aeronautical requirements as contained in WMO Technical Regulations [C.3.1] and with the realization that they are being used by both meteorological and non-meteorological personnel involved in aeronautical operations. Aeronautical meteorological codes, therefore, need to have a simple, self-evident and unambiguous direct-reading quality.

In view of the above, the structure of aeronautical meteorological codes is such that the individual code groups, following each other in the prescribed order, each contain a specific piece of information regardless of the number of characters required.

To allow for easy identification and to avoid misinterpretation, code groups or extensions thereof may include identifier letters (rather than figures) or standard abbreviations, when appropriate, e.g., the identifier "R" stands for runway visual range, "MPS" for meters per second, etc. Consequently, and contrary to the practice in other meteorological codes (see section 4.2.3), code groups in aeronautical meteorological codes contain a non-uniform number of characters (ranging from 2 to 15 characters) and the code data groups in these reports or forecasts may consist of figures or letters or any combination thereof. In addition, code data groups in aeronautical meteorological reports or forecasts indicating significant present or forecast weather (w'w') may be preceded by a plus (+) or minus (-) sign, indicating the intensity of a significant present or forecast weather phenomenon, as appropriate.

When a phenomenon does not occur, or the state of a phenomenon or the value of an element is of no significance to aeronautical operations, the corresponding group, or the extension of a group, is omitted from a particular report or forecast. Again contrary to the practice in other meteorological codes, these optional groups (or extensions of groups) are not distinguished in the code form. Round brackets put around groups in

aeronautical meteorological codes serve a different purpose, i.e., to indicate that the use of the group(s) concerned is determined by regional or national decision.

Differences in coding procedures as just described with respect to the METAR, SPECI and TAF codes, are also applied in related meteorological codes (ARFOR, ROFOR), as well as in AMDAR and WINTEM, and the radiological codes (RADREP, RADOF).

4.5 Observing Station Identification

Each code form provides for proper identification of the coded message and includes code groups for its positioning in space and time. For fixed observing stations, whose position cannot change with time, station index numbers permit the identification of the stations, including their location (and other particulars). For mobile observing stations, whose reports clearly require the full inclusion of

positioning groups, additional station identification allows meteorological services or centres to follow and recognize the successive reports from those stations.

4.5.1 Fixed Meteorological Observing Land Stations

Except in aeronautical meteorological codes (see paragraph 5.4), fixed land stations at which meteorological observations are made are identified by a five figure group consisting of a two figure block number (II) followed by a three figure station number (iii).

The block number defines the area in which the reporting station index is situated. The station index numbers have been allocated as follows:

Region I: Africa	60000 - 69999
	20000 - 20099
	20200 - 21999
	23000 - 25999
	28000 - 32999
Region II: Asia	35000 - 36999
	38000 - 39999
	40350 - 48599
	48800 - 49999
	50000 - 59999
Region III:South America	80000 - 88999
Region IV: North and Central America	70000 - 79999
Region V: South-West Pacific	48600 - 48799
	90000 - 98999
	00000 - 19999
	20100 - 20199
	22000 - 22999
Region VI: Europe	26000 - 27999
	33000 - 34999
	37000 - 37999
	40000 - 40349
Stations in the Antarctic	89000 - 89999

Block numbers are allotted to the services within each Region by regional agreement.

Station numbers (iii) corresponding to a common block number (II), except 89, are usually distributed so that the zone covered by this block number is divided into horizontal strips; e.g., one or several degrees of latitude. Where possible, station numbers within each strip increase from west to east and the first figure of the 3-figure station number increases from north to south.

Station index numbers for stations in the Antarctic (Block 89) are allocated by the Secretary-General in accordance with the following scheme:

a. Each station has an international number 89xxy, where xx indicates the nearest 10 deg meridian which is numerically lower than the station longitude. For east longitudes, 50 is added to the xx value. The figure "y" is allocated roughly according to the latitude of the station with "y" increasing towards the south;

b. For stations for which international numbers are no longer available within the above scheme, the algorithm will be expanded by adding 20 to xx for west longitude (range of index numbers 200-380) and 70 for east longitude (range of index numbers 700- 880) to provide new index numbers;

c. Antarctic stations which held numbers before the introduction of this scheme in 1957 retain their previously allocated index numbers.

Station index numbers consisting of one figure repeated five times, e.g. 55555, 77777, etc., or ending with 000 or 999, or duplicating special code indicators, e.g. 10001, 77744, 19191, 89998, etc., are not assigned to meteorological stations. The general list of station index numbers, together with information on name and location of the observing stations and their observing programmes, is published in Volume A of WMO Publication No. 9.

In aeronautical codes, fixed observing stations are identified by a 4-letter station indicator in the form CCCC, i.e., the ICAO international location indicator. ICAO Publication Doc. 7910 contains a description of the system of assignment and a complete listing of the ICAO international location indicators.

4.5.2 Mobile meteorological observing land stations

Mobile land station making an upper-air observation or issuing a radiological report on a routine or special basis, are identified by a call sign consisting of three or more characters. It is recommended that this group should be encoded in the form A1A2DDD, where A1A2 are the two-letter geographical designators related to countries or territories as specified in Table C1, Part I of Attachment II-6 of Volume I of the Manual on the GTS (WMO Publication No. 386). The DDD is a location designator composed of the first three letters of the name of the town or commune where the mobile land station made its observation.

4.5.3 Fixed Meteorological Observing Sea Station

Fixed sea stations (lightships and fixed platforms) at which meteorological observations are made and reported using land code forms, are included in the system of station index numbers described in paragraph 4.5.1, and are thus identified by the 5-figure station index number in the form IIiii. Otherwise, fixed meteorological observing sea stations are identified by either the call sign of the ship in the form D...D or an identification group in the form $A_1b_wn_bn_bn_b$, as appropriate (see paragraph 4.5.4, below).

4.5.4 Mobile meteorological observing sea station

Mobile sea stations (ships, buoys, rigs and platforms) making meteorological observations, are identified by either the call sign of the ship in the form D...D or, in the case of a drilling rig, oil- or gas-production platform, and drifting or moored buoys, an identification group in the form $A_1b_wn_bn_bn_b$. The call sign of a ship is composed of three or more alphanumeric characters.

In the identification group $A_1b_wn_bn_bn_b$, A_1b_w normally corresponds to the maritime zone in which the observing station has been deployed. A_1 being the number of the WMO Region (1 - Region I, 2 - Region II, etc.), and bw

the number of a sub-area belonging to the area indicated by A_1 (see Manual on Codes, Volume I, Part A, code table 0161).

The WMO allocates to Members, who request and indicate the maritime zone(s) of interest, a block or blocks of serial numbers (nbnbnb) to be used by their drilling rigs, oil- or gas-production platforms and drifting or moored buoys. The Member concerned registers with the WMO the serial numbers actually assigned to individual stations together with their geographical position of deployment. The WMO then informs all concerned of the allocation of serial numbers and registrations made by individual Members.

In reports of sea stations other than buoys, drilling rigs and oil- or gas-production platforms, and in the absence of a ships' call sign, the word SHIP is used for D...D.

4.5.5 Hydrological observing stations

An international hydrological observing station identification number in the form (OOOAC_i) BB$i_H i_H i_H$ is included in the reports of hydrological observation for a hydrological station and in a hydrological forecast. The two groups permit the identification of the WMO Region (A), country (C_i), river basin or group of basin (BB) and the station ($i_H i_H i_H$). The allocation of identification numbers is the responsibility of regional associations for C_i and BB, and Member countries for $i_H i_H i_H$.

A Region may have a maximum of 99 indicators for large basins or groups of small basins. The number BB=00 is not used. If a country straddles several basins (BB), it should nevertheless have only one and the same figure for C_i. If a basin BB comprises all or part of the territory of more than ten countries, C_i should be allocated starting with the largest countries, giving joint national numbers to others (the smallest). In the latter case the national identification numbers of the station ($i_H i_H i_H$) should be allocated by regional agreement.

Alternatively large river basins composed of more than nine countries may be divided into several sub-basins, each one of which may be allocated a separate BB; thus the number of countries will be less than ten in each BB. In each country and for a portion of a basin BB, the national identification numbers of stations ($i_H i_H i_H$) increase from 010 to 999 from west to east and from north to south. The numbers from $i_H i_H i_H$=000 to $i_H i_H i_H$=009 may be reserved to designate the identification of hydrological forecast centres.

The lists of Ci and BB are published in Volume II of the Manual on Codes and the lists of $i_H i_H i_H$ will be published in a separate volume (Operational Hydrology Report No. ..., WMO Publication No. ...). (This publication will appear at a later stage.)

4.5.6 Airborne observing stations

Transport aircraft in which upper air meteorological observations are made and reported using the AIREP message format, are identified by the call sign of the aircraft as the first group of the report. This call sign is normally composed of an ICAO three-letter designator indicating the aircraft operating agency, and a three- or four-figure flight number. In the absence of an aircraft call sign, the indicator letters ARP are used as the first group of the report. The ICAO designators for aircraft operating agencies are published in ICAO Publication Doc. 8585.

The AIREP message content is given in WMO Technical Regulation [C.3.1.]5.8.5.

4.5.7 Space-based Observing Stations

In reports of data derived from observations by meteorological satellites (SATOB, SARAD, SATEM), the group $I_1 I_2 I_2$ (completed with $I_3 I_4$ or solidi) is used to identify the particular meteorological satellite. The symbolic letter I_1 gives the name of the country or international agency which operates the satellite (see Manual on Codes, Volume I, Part A, code table 1761) and the symbolic letters $I_2 I_2$ is the indicator figure for satellite name

(supplied by operator I_1) in which even deciles are used for geostationary satellites and odd deciles for polar-orbiting satellites.

In satellite ephemeris messages which are issued in national code forms by the USA (TBUS) and the (erstwhile) USSR (FANAS) to transmit information for predicting the path, or locating the position, of polar-orbiting environmental satellites, there is no standardized procedure for identification of meteorological satellites. In TBUS messages the meteorological satellite concerned is indicated by its national operational serial number (NOAA-N) in the first (identification) section of the message, whereas in FANAS messages the meteorological satellite concerned is indicated by the first group in Section 1 of the report, in the form $11L_sS_sS_s$. Following the indicator figures 11, symbolic letter L_s indicates the launching country in a code similar to that of symbolic letter I_1 and symbolic letters S_sS_s give the number (series of the satellite) as follows:

21, 22, 23 ATS series (USA)

04, 05, 06 NOAA series (USA)

02, 03, 04 METEOR-2 series (USSR)

A full description of the TBUS and FANAS codes is contained in the appendix to Volume II of the Manual on Codes.

4.5.8 Meteorological Data-processing Centres

In messages containing processed meteorological data in the form of grid-point values (GRID, GRAF) or radiological trajectory forecasts (RADOF), the first line of the text of the coded meteorological analysis or forecast contains an identification group in the form F_1F_2NNN or $F_1F_2Y_rY_rG_rG_r$ respectively, consisting of an indication of the data-processing centre originating the product and a reference to grid system used (NNN) or the date and time of issue of the forecast ($Y_rY_rG_rG_r$). The allocation of centre identifiers F_1F_2 is listed in the Manual on the GTS, Volume I, Part II, Attachment II-9, Table A (WMO Publication No. 386).

4.6 Dates and Changes in Meteorological Codes

4.6.1 Procedures

There is a need for developing new or updating existing meteorological codes to meet new or changed requirements. For international meteorological codes it is the responsibility of CBS and its WGDM to review the stated requirements and to recommend appropriate action. It is the CBS policy not to change codes, or introduce new code forms, unless there are pressing needs.

Regional procedures for the use of international meteorological codes and meteorological codes which are intended only for exchanges within a given WMO Region or the Antarctic, are the responsibility of the regional associations concerned and their Rapporteurs on DM or, in the case of the Antarctic, the WMO Executive Council upon the advice of the EC Working Group on Antarctic Meteorology, to review the stated requirements and to decide on appropriate action.

4.6.2 Date of Introduction

The relevant decisions, taken by a regional association or, in the case of recommendations by the Commission for Basic Systems or the EC Working Group on Antarctic Meteorology, taken by the Executive Council or the President of WMO, as appropriate, also include a date of introduction for the new or amended code, which normally takes due account of the time frame required for proper world- or region-wide implementation, such as updating of national publications of coding instructions, training of personnel, adaptation of computer programmes, etc. This date of introduction should preferably be the first day of a month, but not coincide with an internationally recognized public holiday like New Year's Day, 1 January.

4.6.3 METNO Messages

In order to provide Meteorological Services with advance notification of changes of operational importance (which include important changes in meteorological codes), the WMO Secretariat issues at weekly intervals (on Thursdays) telegraphic messages containing advance notification of such changes. These weekly telegraphic messages are identified by the code name METNO. Detailed description of the distribution, format and contents of METNO messages are given in the relevant parts of the Introduction to WMO Publication No. 9, Volumes A and C.

4.6.4 Monthly Letter on the WWW and MMS

In addition to the METNO weekly telegraphic notifications, information on important changes, including meteorological codes and (binary) data representation forms, is included in the monthly letter on the operation of the World Weather Watch (WWW) and Marine Meteorological Services (MMS). The monthly letter is issued on the last working day of each month in English, French, Spanish and Russian.

4.6.5 Manual on Codes

Eventually, of course, all changes to meteorological codes are incorporated in the Manual on Codes, either through numbered supplements to the current (loose leaf) edition, containing replacement pages or instructions for manuscript corrections, or by the issuance of a new edition.

CHAPTER 5

METEOROLOGICAL DATA REPRESENTATION: BINARY REPRESENTATION FORM

5.1 Introduction

The Character Code Forms provided the means to overcome the barriers of language, and enabled meteorological information to be interpreted correctly and accurately by any person with sufficient knowledge, regardless of race or native tongue. With the advent of automated data processing it has been recognised that a corresponding need exists to represent meteorological data in a way which can be readily interpreted by any computer, irrespective of make, operating system, or internal data representation. Whereas the most generally recognised form of data representation among the peoples of the world is the numerical system based on the so called "Arabic" notation, for computer applications the most recognised form is based on the binary representation of whole numbers.

In developing binary representation forms for the representation of meteorological data it has been recognised that:

a. Data have more value if correctly defined within the representation form;

b. Data volumes will continue to increase, thus efficiency of representation is important;

c. Whereas meteorological observed data usually relate to a point source, analysis and forecast products are normally presented for selected parameters at selected levels for designated geographical areas.

Ideally, a self defining binary representation form should be universal - capable of representing a required set of data. In practise, the last consideration in the previous paragraph has resulted in two forms being developed for the representation of meteorological data - FM 92 GRIB, used mainly for the representation of products, and FM 94 BUFR, used mainly for the representation of observational data.

5.2 BUFR (FM 94)

Over the past few years the WMO, with assistance from a number of the world's operational meteorological/oceanographic processing centers, have been collectively developing a new system for the exchange and storage of meteorological, hydrological and oceanographic observations, known as BUFR, for Binary Universal Format for Representation of data. In so doing, they have developed a very flexible data representation system that is suitable for any kind of data, not just meteorological information.

5.2.1 The Essence of BUFR

The key to understanding the power of BUFR is the code's self- defining nature. A BUFR message not only contains meteorological data; it also includes a complete description of what those data are: the description includes identifying the parameter in question (height, temperature, latitude, whatever), the units, any decimal scaling that may have been employed to change the precision from that of the original units, any data compression that may have been applied for efficiency, and includes specifying the number of binary bits used to contain the numeric value of the observation. This information, the "data description", is all contained in tables which are the major part of the BUFR documentation.

The strength of this self-defining feature is in accommodating change. For example, if new observations or observational platforms are developed, there is no need to invent a whole new code form to represent and transmit the new data; all that is necessary is the publication of additional data description tables. Similarly for the deletion of possibly outdated observations - instead of having to send "missing" indicators for a long period while awaiting a change to a fixed format code, the "missing" data are simply not sent in the message and the data description section is adjusted accordingly. The data description tables are not changed, however, so that archives of old data may be retrieved.

This self-descriptive feature leads to another major advantage over character oriented codes - the relative ease of decoding a BUFR message. Where a large number of specialised and complex programs are now needed to decode the plethora of character codes in current use, it is entirely feasible to write a single "universal BUFR decoder" program capable of decoding any BUFR message. Such has been done at ECMWF, NMC Washington and presumably at other centers. It is not a trivial exercise to write such a BUFR decoder, but once done, it is done virtually for all time (except for extensions which may be needed to accommodate new features). The program will not have to change with changes in observational practices; only the tables will be augmented, a relatively straight forward task.

The development of BUFR has been synonymous with the development of the data description language that is integral to it. Indeed the major portion of the full description of BUFR is a description of the vocabulary and syntax of the data description language. These are manifested in the tables of the BUFR documentation. The definition of the data description language, and the "descriptors" that are its vocabulary, are what give BUFR its "universal" aspect; any piece of information can be described in the language, not just meteorological observations. The full details of the data description language are available in WMO Publication No. 306, Volume 1, Part B.

Another major aspect of BUFR is reflected in the first initial, "B": BUFR is a pure binary or bit oriented form. All the numbers in a BUFR message, whether data descriptors or the data themselves, are binary integers; a paper representation of the contents would consist of a long string of 0 and 1 values, almost unintelligible to humans. Thus BUFR can only be assembled with the aid of a computer, and requires a computer for its meaning and contents to be displayed. Thus, some degree of automated data processing is essential for the support of binary representation forms such as BUFR.

Given that some form of automated data processing is widespread in the developed world, and seen as highly desirable in the developing nations, the true binary nature of BUFR has been deliberately developed to make it highly machine independent. Since BUFR is comprised entirely of binary integers any brand of machine can handle BUFR as well as any other.

The binary nature of BUFR leads to another advantage over character codes: the ease and speed of converting the message into an internally useful numeric format. With character codes the conversion from ASCII (or

EBCDIC) to integer or floating point is expensive relative to the conversion from binary integers to floating point. The latter is all that BUFR requires. In some recent tests, the ECMWF found a speedup of better than 6 times in decoding BUFR messages over the corresponding TEMP (WMO radiosonde character code) messages. The BUFR data also required about half the machine memory as the character data.

All of this does assume the availability of well designed computer programs that will be capable of parsing the descriptors, which can be a complex task, matching them to the bit stream of data and extracting the numbers from it, responding properly to the arrival of new (or the departure of old) data descriptors, and reformatting the numbers in a way suitable for subsequent calculations. The bit oriented nature of the message also requires the availability of bit transparent communication systems such as the appropriate levels of the X.25 protocol. Such protocols have various error detecting schemes built in so there need be little concern for the garbling of information.

5.2.2 The Vocabulary of the Data Descriptors

The vocabulary of data descriptors is given in terms of the tables that are the bulk of the BUFR definition. The table entries for each descriptor are:

A table reference (a six digit number which uniquely identifies the table entry);

The name of the parameter;

The units of the parameter;

A decimal scale factor that the element value is multiplied by prior to encoding (to change precision);

A reference value that is subtracted from the (scaled) element value prior to encoding (to eliminate negative numbers);

The length, in bits, that the value, as scaled and with the reference subtracted out, will take up in the data section of the BUFR message.

The "parameters", as understood in the BUFR context, include not just the meteorological variables, but also such things as the four-dimensional location of the observation, the block/station number, the instrumentation, the "significance" of the observation, the quality, etc. In other words, all of the pieces of information, all of the "co-ordinates" in multidimensional space that describe an observation, are parameters embraced by BUFR.

As an example of the steps necessary to prepare a data element for a BUFR message, consider the latitude of an observation point, having a value of -45.76 degrees (south). The appropriate descriptor table reference (005012 for this example) shows that the units are degrees, and that the number is to be scaled by a factor of 100; thus the scaled value becomes -4576 centidegrees. The reference value for latitude is -9000, thus the element with the reference subtracted from it is 4424. The particular reference value was selected to ensure that all possible latitude values end up as positive, thus avoiding any machine dependency in the representation of negative numbers. The last step is then to place the 4424 value on a "word" 15 bits in length. The bit length was selected by considering the largest value that the scaled and referenced value might ever take on, and selecting a word-length sufficient to represent that value. This is trivial for latitude; some thought had to go into selecting scale factors, reference values, and word-lengths for other (e.g., meteorological) variables.

In BUFR, at present, there are some 22 separate classes within the table of descriptors, classified by their content. The first seven of the tables deal, in effect, with the "coordinates" of observations: their identification, instrument types, four dimensional location, and "significance", the last including such things as indicating mandatory or significant levels in RAOBs, land/sea distinctions, etc. The remaining tables specify actual measured parameters in logical groupings: vertical elements and pressure, wind, radiation, clouds, etc. Not all of the elements described are measured values; many of the "data" are references to code tables or bit flag tables,

which allow current qualitative observational practices to be included in the essentially quantitative BUFR representation. The latter tables are also included in the BUFR documentation.

The use of code tables allows BUFR to describe qualitative information in a numeric form. The scaling employed within the BUFR tables has been selected to assure adequate precision for the values of the parameters in the BUFR message. If an alternative unit of measurement for a parameter, degree of precision desired, or other variables wanted, all that is necessary is to use operators to re-define precision, or add new descriptors to the table. Once a descriptor has been placed in a table it normally will not be changed; this ensures that archives will remain accessible.

In its most elemental form, then, the essential parts of a BUFR message are a collection of descriptors matched up one-for-one with the data elements described. The descriptors are all fixed length; the data elements have variable length, as specified in the tables. This could be very inefficient as the descriptor list could easily exceed the length of the data described, This potential for inefficiency is recognised in the BUFR system, which leads directly into the next topic.

5.2.3 The Syntax of Data Descriptors

A single descriptor is two octets or 16 bits long, and contains three values:

F X Y

F, with length of two bits, specifies the nature of the descriptor.

If F equals:

> 0 - the descriptor is an "element descriptor" describing, on a one-for-one basis, a single data element. X (six bits) indicates the classification table, Y (eight bits) denotes the element within the table. These are typified by the Class 11 Table mentioned above.
>
> 1 - the descriptor is a "replication operator" stating that the following X descriptors should be repeated for a total of Y replications. The data elements, of course, repeat in the same pattern. Special case: if Y=0 then the replication count is embedded in the data themselves. This "delayed replication" is very useful when a set of general purpose descriptors are being assembled and the number of replications is not known ahead of time. This would be the case for RAOBs, for example, with an a priori unknown number of significant levels.
>
> 2 - the descriptor is a "special operator" (other than replication) denoting some specific action, such as changing the data width by a specified number of bits or changing the reference value.
>
> 3 - the descriptor is a "sequence descriptor" which, by means of table references, specifies a collection or list of other descriptors (of type 0, 1,2, or 3) for common reporting sequences.

The use of descriptors of type 1 and 3 make it possible to reduce greatly the overhead of including many descriptors in every message when the layout of the data is the same from observation to observation. One "standard" observation, no matter what its length or number of elements, can be described with one (type 3) descriptor. Indeed, that is the definition of "standard": the presence of a sequence of descriptors in the type 3 table establishes the standard. (Just how many such standard sequences should be defined has been a matter of some discussion.)

There are, naturally, a rather large number of rules that specify the use of these descriptors and how they interact. Reference to the Manual on Codes (WMO Publication No.306) provides details. The rules were crafted with some considerable care and seem to be free of inconsistencies. Taken together, they are a specification sufficient for writing computer programs to encode or decode BUFR messages.

There is a lot of room in the descriptor tables: X, the class, can range up to 63 (currently tables are defined through X = 31), and Y, the element entry, up to 255 (the largest in use is 63). The regions in the table from X = 54 to 63 and Y = 192 to 255 (inclusive) are reserved for local use and may contain any descriptors that the local center wants to put there; the contents of the rest of the table will be agreed upon by international convention and will be universally available. BUFR even includes a set of descriptors describing the descriptor tables themselves so that updates can be sent around in BUFR code. It is a "universal" format, after all.

5.2.4 Compression of Reports

BUFR makes quite efficient use of space by virtue of its use of binary numbers that take up only as many bits as are necessary to hold the largest expected value. However, when many reports all with the same layout of individual observations, are available, a further compression is possible. The technique similar to that used in GRIB (the WMO code form for GRIdded Binary fields) in that all the like elements from the full set of observations are collected together, their minimum values subtracted out, and the residuals are then represented as binary integers each with a bit length selected to hold the largest residual.

The number of separate reports (called data sub-sets in the BUFR definition) is then placed in the data description part of the message and the data descriptors are not repeated. Often the data descriptor is of type 3, so there will be but one two-octet descriptor in the message used in conjunction with tens or hundreds of compressed observations.

5.2.5 A Complete BUFR (Edition 2) Message

One BUFR message is divided into logical sections without any special characters or bit configurations to separate them. The contents are:

SECTION	CONTENT
0	Indicator Section
1	Identification Section
2	Optional Section (local use)
3	Data Descriptors Section
4	Data Section
5	End Section

Section 0, the indicator section contains the letters BUFR (character coded according to the International Alphabet No. 5), total length of the message (including 'BUFR') and the BUFR edition number.

Section 1, the identification section, contains among other things, an indication of the particular BUFR message type (surface, upper air, etc.) to enable a decoding program to select a general category for information storage prior to looking in detail at the descriptors. The section also contains identification of the originating center, the date, a flag indicating the presence or absence of section 2, and other useful information.

Section 2, the optional section, is completely undefined at present, except for indicating its own length. It is intended for local use such as some sort of data base key.

Section 3, the data descriptors section, prefaces the collection of descriptors with the number of data sub-sets (used in the compression algorithm) and a flag to indicate whether the data are indeed compressed. They do not need to be.

Section 4, the data section, contains only its own length and the actual data described in section 3

Section 5, the end section contains the numbers 7777 (character coded according to the International Alphabet No.5)

As well as the total length of the message in section 0, sections 1 to 4 contain, in their first three octets, a count of how long the section is. Thus it is no problem for a computer to find its way through a BUFR message using 'BUFR' and '7777' as checkpoints.

5.3 GRIB (FM 92)

5.3.1 The Development of GRIB

GRIB (GRId in Binary) was designed to provide a common, global binary representation form for processed data, enabling faster transmissions and reducing storage requirements by packing those data types which are already arranged in a regular grid format, in a highly compacted form.

The WMO Commission for Basic Systems ISS Expert Meeting on Exchange Formats defined the first version of the GRIB in 1984, and this was then used experimentally by a number of centres. Following these trials, some changes were made and the code was adopted in 1985 by CBS. At its ninth session, in 1988, CBS adopted amendments and extensions to generalise the representation of spherical harmonic components and enable polar stereographic representation to be included. At an Expert Meeting of the CBS Working Group on Data Management Sub-group on Data Representation (October 1990) further amendments to GRIB were considered and subsequently approved for use by the president of CBS. They include the provision of decimal scaling, representation of a matrix of values at each grid point and the representation of image data such as satellite pictures.

5.3.2 A General Description of GRIB (Edition 1)

Each field encoded in GRIB consists of one meteorological parameter at one level in the atmosphere. As well as the data, sufficient information is included to define the product and describe the data representation used. As with BUFR the representation form is independent of any particular machine representation. All lengths are given in octets (8 bits). A GRIB coded product consists of 6 logical sections, two of which are optional.

SECTION	CONTENTS
0	Indicator Section
1	Product Definition Section
2	Grid Description Section (optional but recommended)
3	Bit Map Section (optional)
4	Data Section
	End Section

Section 0, the indicator section contains the letters GRIB (character coded according to the International Alphabet No. 5), total length of the message (including 'GRIB') and the GRIB edition number.

Section 1, the product definition section identifies the centre originating the data and the model used in generating the data. The data itself is described: time, parameter, level etc. A flag field is also included to indicate the presence or absence of sections 2 and 3.

Section 2, the grid description section is optional, but it is recommended that it usually be included. It may be omitted only if the grid description is published in Volume B of WMO Publication No. 9, and the appropriate catalogue number of the grid in this publication is included in the Product Definition section. The contents of this block vary, depending on the way data in the Data section (section 4) is represented. For a regular latitude/longitude grid it contains the latitude and longitude of the origin and the extreme point, the number of points along a latitude and along a meridian, and optionally the direction increments. It also identifies the grid point scanning mode e.g. West to East, North to South.

Section 3, the bit map section, if used, can contain either a bit map or a reference to a pre-defined bit map provided by the originating centre. The bit map consists of contiguous bits, with a bit to data point correspondence, ordered as defined in the grid definition.

Section 4, the data section contains the packed data. A number of formats are possible, depending on whether the data is in grid point or spherical harmonics and whether simple or complex packing methods are used. However, the fundamental principle in packing the data remains the same - data is coded in the form of non-negative scaled differences from a reference value. The reference value (R) is stored in four octets as a floating point number. The high order bit is a sign bit (sb), O indicating positive, 1 negative. The next seven bits are the characteristic (C) and the low order 24 bits the mantissa (B).

$$R = (-1) ** sb * 2 ** (-24) * B * 16 ** (C-64) \quad (1)$$

A scale factor (s) is stored in 16 bits (sign bit and 15 bit integer).

Data values transmitted (P_j) are coded in the number of bits required to achieve the required accuracy.

The real data values (Y_j) are unpacked as follows:-

$$Y_j * 10 ** D = R + (X_j * 2 ** E) \quad (2)$$

where: Y_j (j=1..n) are the real unpacked values

D is a decimal scale factor (signed integer)

R is a reference value

X_j (j=1..n) are packed values (positive integers)

E is a binary scale factor (signed integer)

To optimise accuracy with respect to packing density, it is necessary to choose a suitable scale factor, s, give a packing density of i bits per packed value (see 6.3.3 below).

Section 5, the end section contains the numbers 7777 (character coded according to the International Alphabet No. 5).

As well as the total length of the message in section 0, sections 1 to 4 contain, in their first three octets, a count of how long the section is. Thus it is no problem for a computer to find its way through a GRIB message using 'GRIB' and '7777' as checkpoints

5.3.3 Computation of the Scale Factor

Let A_j (j=1...n) be the n real values to be packed

For max (A) = $A_{m'}$ we have

$U_m = r + (P_m)2^s$

But $P_m < 2^i-1$ where P_m contains i bits

Therefore $U_m = r + (2^i-1)2^s$ if ALL bits of P_m used.

This represents values of A such that

$r + (2^i-1)2^s - 0.5(2^s) < A < r + (2^i-1)2^s + 0.5(2^s)$

Thus, A_m is less than

$r + (2^i-1)\ 2^s + 2^{s-1} = r + 2^{s-1}(2^{s+1}-1)$

We require s to be the least integer such that

$r + 2^{s-1}(s^{i+1}-1) > A_m$

That is, $2^{s-1}(2^{i+1}-1) > A_m - r$

But $2^{i+1}-1 > 0$

Therefore $2^{s-1} > (A_m-r/2^{i+1}-1)$

$s > \log_2((A_m-r)/2^{i+1}-1)) + 1$

Thus s must be fixed such that

$s = \text{floor } [\log_2((A_m-r)/2^{i+1}-1))] + 2$

where floor [ ] = greatest integer not exceeding [ ].

5.3.4 Representation of Reference Values

Equation (1) above defines the representation used in GRIB for floating point reference values. It is necessary to transform the floating point notation used within the data processing computer in which GRIB is being encoded into this representation.

The basic transformation required is:

F = 16 ** N * X

where: F is the real number to be stored.

N is an 8 bit integer exponent (with a bias of 64 included).

X is the 24 bit normalised mantissa (1/16 < |X| < 1).

The formula:

N = INT(ALOG(F)*(1.0/ALOG(16.0)) + 1.0 + eps)

gives good results with respect to the normalised conditions on machines with word length > 60 bits for EPS equal to 1.0E-12, and on 32 bit machines for EPS equal 1.0E-8. Note that normalisation of the mantissa is essential to ensure maximum accuracy.

5.4 Issues Raised by the use of Binary Representation Forms

It is not likely that BUFR and GRIB will replace the current WMO character oriented codes in any wholesale manner in the short term. There are simply too many character oriented communications lines in use, in both the developing and developed communities, for change to come rapidly.

One possible scenario for the development of the WWW is for regional centers to take in individual observations from an assortment of observation sites, combine them, convert them to the BUFR format, and pass them on to national or international centers for world wide distribution. At the same time these regional centres would be able to use their computing infra-structure to decode and display gridded fields transmitted in GRIB code.

It is to be hoped that rather than having each centre develop its own code for the manipulation of BUFR and GRIB messages, one, or at most two centres take on the responsibility for the development and maintenance of such software.

One issue, raised earlier, is the possible necessity for some centers to ingest BUFR messages and then reformat them back into (current) WMO character code forms for local transmission on character lines, However, before this gets extensive, serious thought should be given to the purposes of such retransmissions. If they are undertaken to supply non-automated centers with "human-readable" information, then there may be more sense in dropping the use of the sometimes complex WMO character code forms and instead send the information out in a simple tabular form; alternatively, graphical products may be generated and transmitted. Eventually, the aim should be to eliminate the need for training in human decoding of the WMO code forms. Programming to generate tables, for transmission and display, is much simpler than generating the current WMO character codes. If the transmissions are to automated centers, then they should switch over to the binary representation forms as soon as possible. Of course, there will be an interim period when both binary and character data are flowing on the GTS, but every effort should be made to avoid converting binary data, particularly BUFR encoded messages, back into the character codes.

Back to Part 1
CHAPTER 6 - The Use of Computer Graphics for Meteorological Data Representation
CHAPTER 7 - Data Bases in a Meteorological Environment
CHAPTER 8 - The Distributed Data Bases Concept
CHAPTER 9 - Data Monitoring
CHAPTER 10 - Computer Software Exchange
CHAPTER 11 - Endnote
LIST OF ACRONYMS

CHAPTER 6

THE USE OF COMPUTER GRAPHICS FOR METEOROLOGICAL DATA REPRESENTATION

6.1 Introduction

Computer Graphics is used to present meteorological data in a visual form. Examples include: observations, time series and other graphs, contours, satellite and radar images, animations of contours and images, and 'three dimensional' pictures of storms and other output from computer models. Generally there are two objectives in displaying the development of meteorological phenomena in space and time. Firstly, to explore and understand the data, usually using the traditional complex charts, and secondly, in presenting that understanding to other people, usually with simpler charts, such as a significant weather chart. In general, the former visualisation techniques aim at using the largest possible amount of information, within time and performance constraints and the limits of human capabilities.

Before the advent of computer graphics, meteorological graphics consisted of various hand drawn two dimensional representations: e.g. weather observations plotted on charts, analyses of various fields (e.g. pressure or wind speed and direction) or cross-sections. Computer graphics allows:

- The automation of the drawing of these two dimensional graphics, whether on a computer screen or on paper;
- The rapid overlaying of different data, and the successive presentation of picture sequences (animation);
- The construction of 'three dimensional' representations, which may present to the user a much larger quantity of information than is currently shown by two dimensional methods;
- The interactive and flexible display of both two dimensional and 'three dimensional' representations.

This chapter focuses on the issues raised by the use of screen based technologies for meteorological data visualisation and presentation. As discussed above, the current generation of graphical systems makes it possible to display interactively meteorological data in two and three dimensions. Future development in technology should make possible the widespread use of such systems in the l990s. Also numerical weather prediction models and remote sensing platforms are evolving, generating an increasing volume of data that requires more and more powerful data exploration and presentation tools.

6.2 Current Graphics Systems and Techniques

Approaches to obtaining meteorological graphical systems range from purchasing complete systems and maintenance from one of a number of commercial vendors to developing and maintaining all of the software for your own hardware (nobody builds their own hardware!). The latter approach has the advantage of being more flexible, but the cost of developing the system is not shared with other people.

6.2.1 Graphics Hardware

Computer graphics systems can be classified into two categories:

a. Vector-based systems, where the graphical picture is defined in terms of lines. Pen-plotters are the only common devices in this category. Their main characteristic is that pictures can be readily scaled in size without changing their appearance, and drawing time is proportional to the complexity of the picture;

b. Raster-based systems, where the picture is composed of a regular array of dots (called pixels). Nearly all screen based devices are in this category, as well as hardcopy devices such as fax machines, electrostatic plotters and matrix printers. The main characteristic of these systems is that the time to display a picture is largely independent of the picture's complexity but the resolution is lower than that achievable in vector systems.

Most modern systems have components from both categories, and a vector picture can be readily converted to a raster picture, but not vice-versa.

Raster-based systems can be classified into those that only accept raster pictures, such as fax machines, and those that can accept a vector picture and can convert it for display, such as most workstations and personal computers. In these latter systems, the vector to raster conversion may be done by the general purpose computer hardware, as in personal computers, or done by dedicated, specialized graphical hardware, separate from the main processor, as in most workstations.

Nearly all graphical displays are two dimensional screens, but some of the specialized graphical hardware can handle data as if it were three dimensional and project it onto the two dimensional screen, giving a perspective picture. Such devices are slightly misleadingly called three dimensional.

Software reflects the distinction between vector and raster based systems, and older proprietary software may be of one or the other kind. Much, but not all, modern software is designed for both type of graphics.

At present, hardware is developing very rapidly, but software cannot be developed rapidly enough to keep pace. As in all computer systems, the cost of developing and maintaining software is becoming the most significant part of the cost of a computer graphics system over its lifetime. A successful approach to this problem is to ensure that any software developed can be moved to new hardware quickly and easily. A corollary of this is that the data must be readily portable to the new hardware also, and so must the skills of the people maintaining the system.

6.2.2 Graphical Software

A survey of the current technology for meteorological graphics leads to the following classifications:

6.2.2.1 Two-dimensional Graphical Libraries

These systems represent the current mainstream of operational meteorology, and usually consist of graphical subroutine libraries which usually includes collections of FORTRAN callable subroutines. Such systems often make use of the GKS and CGM graphical standards (see Section 6.4). MAGICS and NCAR Graphics are meteorologically orientated examples of such libraries.

The general graphical functions available perform actions such as contouring, shading, wind field plotting, map and text representation and cross-sections. Observation plotting, thermodynamic diagrams and meteograms are specific to meteorology.

Whilst two-dimensional graphics packages are now commonplace, there is still room for improvement, especially in the area of contour labelling and overlaying, e.g. 100-500 hPa thickness on a 1000 hPa chart. Also,

current contouring algorithms and techniques can only deal with smooth fields, but not abrupt changes of gradient such as occur with fronts.

Observation plotting in most existing systems is based on WMO recommended practices, which have been devised for manual plotting but are also suitable for machine plotting. However, there are many situations where observations would normally be plotted too close together for readability and where special provisions for plotting should be taken.

6.2.2.2 Two-dimensional Interactive Systems

The evolution of computer technology is enabling the widespread use of interactive two-dimensional systems. One of the important features is the ability to create and/or display an animation sequence within acceptable limits.

Some systems have evolved from satellite image processing, e.g. McIDAS from the University of Wisconsin, or from existing two-dimensional graphical libraries, e.g. MicroMAGICS and SIGMA/NCAR. Non-meteorological packages for the display of scientific data are also becoming more common and may be useful.

A general system should allow for the manipulation of fields, observations and images, both satellite and radar. The availability of interactive graphics and image editing functions are also essential and all data should be geometrically transformed to a consistent geographical base.

6.2.2.3 Three-dimensional Static Systems

These systems are characterised by being able to produce three-dimensional pictures of phenomena to be used both as a diagnostic tool and as a means of portraying and communicating results. Their effectiveness for routine forecasting tasks has not been demonstrated yet, but they are undoubtedly suited to the portrayal of three dimensional flow. Their "static" nature is because they generate animation sequences from computed pictures. Batch processing is often used in such systems.

Typical applications involve the use of a 50 x 50 grid point data at 30 levels and by showing 2 or 3 variables.

To improve performance, various distributed processing options exist, including:

> Generation of pictures on a large mainframe or supercomputer with display on a mainframe graphics terminal or a workstation connected over an Ethernet type network;
>
> Partial graphical functions on a supercomputer with additional computing performed on a workstation containing specialised graphical hardware processors;
>
> Rendering pictures (a display process which makes an image look realistically three dimensional) on a workstation, subject to current restrictions in performance.

In the first and second above mentioned cases, data compression techniques are used to achieve acceptable transmission speed over a network, especially in the case of Ethernet. Such methods are, in general, reversible and techniques such as differencing schemes (delta modulation) combined with run-length encoding have proven to be quite effective. Benchmarks tests have reported satisfactory results, achieving compression factors of up to 4:1.

The functions needed for this class of system are general purpose graphical methods, ranging from simple surface rendering with hidden line and hidden surface removal to more elaborate methods such as transparency, volume rendering and texture mapping.

Techniques to produce pictures of near photographic quality such as ray tracing and radiosity are not widely used because of practical performance limitations.

6.2.2.4 Three-dimensional Interactive Systems

The interactive control and manipulation of large data sets in meteorology represents a promising approach to the visualisation of meteorological information. In such systems, the user controls the display by means of a graphical device, and the system is able to give prompt response to commands such as rotate, zoom and pan, i.e. the frame rate gives a perception of continuous motion.

It has been observed that static three-dimensional graphical images are sometimes visually ambiguous. Use of interactive rotation and animation is an effective means to resolve that ambiguity. To this end, an environment with real-time rendering capability is required. Such environments may be achieved by very fast graphics workstations or by a supercomputer with a fast link to a workstation or framebuffer.

Interaction also allows the user to control the information content of the animated display in order to rapidly search for relevant information in large data sets.

Three-dimensional interactive systems use the same kind of functions as three-dimensional static systems, subject to the restrictions of real-time picture display. One effective function of three-dimensional interactive systems is the depiction of trajectories.

Current technology is only able to deal with limited subregions of model output data sets, and therefore the issues of selection from a large database plays an important role.

Due to the rapid evolution of systems in this area no consensus has yet been reached as to what proposals seem the most promising or fruitful.

A desire was expressed for the more widespread availability of application level interfaces. Whilst a tool-kit like AVS is considered to have the correct balance between a high-level interface and functionality, its dependence on a particular vendor limits its more widespread use.

The current status in three-dimensional software environments is seen as not providing sufficient evidence as to which graphics support packages should be used for visualisation software. A general guidance to encapsulate the non-portable part of the system in separate entities is the only option available for system designers at present.

As meteorological data sets are usually simple in structure compared to CAD/CAM structures, Meteorological systems should be able to survive the current "shake-out" in graphical systems.

6.3 The User Interface

In interactive systems, the user interface is considered to be as important as the graphical functionality. Therefore, great care should be taken to design and implement effective interfaces as these are seen as a crucial factor affecting the usability of the system.

A general preference for "pull-down" menu systems has been expressed and "Macintosh-like" window/mouse systems are becoming commonplace. The use of multiple windows for displaying related types of data is seen as a useful adjunct to the simple overlay of graphical and image information. Other approaches such as the use of subordinate screens displaying menus at all times are also appropriate.

As meteorological graphics are generally complicated, their use in windowing systems is more effective on larger computer screens.

The choices should be "data-driven" where the actions the user can take is driven by the meteorological data to be manipulated. The alternative, a "function-driven" system, is one where the choice of actions is made before data selection is seen as inappropriate for operational environments.

The system should also have default parameters that vary according to the meteorological variables being portrayed. There should be both user definable and system defined default values to which the user may return if desired.

A combination of both ease-of-use and flexibility is needed. Systems developers should be aware that user interface design is still a new topic in computer science. Most experts recommend a prototyping phase before the final system design is completed. Users should be allowed to interact with a pre-release version of the system in order to provide feedback.

Due to the very specialised nature of some graphical techniques it is recommended that the user is confronted with meteorologically relevant choices. This is seen as a crucial step towards making systems acceptable to the meteorological community.

Animation is considered an essential tool, with three performance thresholds:

> Acceptable minimum: circa 1 frame per second;
>
> Motion tracking: 2 to 4 frames per second and the user is able to track the motion of elements in the scene;
>
> Smooth fusion: more than 8 frames per second and the images are still perceived as an animation sequence but they appear as a continuous smooth motion.

The latter is at the limits of today's technology.

User interfaces should present meteorologically meaningful parameters that minimise complexity and they should be designed for portability over a wide variety of computing environments and for easy and flexible use by the meteorological community. To achieve this, it is preferable that the software should have a clear internal interface to separate the graphical display from meteorological data handling systems.

When developing user interfaces there is a need for high level tool-kits and the use of industry standards such as OPEN LOOK(UI) or Motif (OSF) is considered highly desirable.

6.4 Standards for Meteorological Graphics

Meteorological standards such as GRIB and BUFR represent the data and its geographical coordinates, whereas graphical standards are only concerned with the presentation of pictures on graphical devices, whether screens or hardcopy, and the device's coordinate system.

After ten years of effort by many people, computer graphics acquired its first international standard GKS in 1985. GKS is the main building block of a set of inter-related standards covering the whole area of graphics. GKS concentrates on standardising the interface between application software and a two dimensional graphics system, thus allowing portability of applications across different graphics devices and computers. This saves money in the longer term, as application software is usually in existence for longer than any hardware.

Most graphical standards are defined functionally with a separate specification of how they should be realised in terms of a programing language interface to the application or a protocol between the functions and a device. How to invoke graphical functions are prescribed in a series of separate standards called language bindings. The 'binding' of functions into a language can only be specified in a standard way if the programing language itself is standardised. For example, GKS, at present, has agreed bindings to FORTRAN, Ada, Pascal and C. Protocols may have separate versions, according to need, for a given functionality. These are known as encodings.

Because standards are independent of computer vendors and are achieved by consensus, they may take years to be produced. Appendix A describes the various stages involved. These standards are known as 'de jure'

standards, or international standards, and must be distinguished from an alternative meaning of the word 'standard' that is becoming commonplace.

'Standard' often means a 'de facto' standard - something implemented by one particular vendor, and has become so prevalent that it has a virtual monopoly of the marketplace. These 'standards' are prone to alteration by the vendor to suit their purposes, and may not assist in the task of moving software to different hardware.

6.4.1 Graphical Kernel System (GKS)

The major features of GKS are:

a. 2D graphics only;

b. Supports both vector and raster devices, and multiple devices at once;

c. User defined rectilinear coordinates;

d. Comprehensive description of input;

e. Grouping of parts of pictures into segments, but no segments within another;

f. Different levels of support: with or without input, with or without segments;

g. Supports both long term picture storage and audit trail as virtual devices;

h. Supports both default 'bundled' or explicit setting of specifications of graphical elements such as colour, size and style;

i. No current position.

Its advantages are:

1. It is widely available;

2. It has established a consistent comprehensive terminology for graphics;

3. It can be implemented efficiently;

4. It has established a consistent model of the graphics processes.

Its disadvantages are:

1. The single level segment hierarchy is not sufficient for many applications;
2. Initially there were too many implementation differences between different vendors.

GKS provides a standard for graphics in two dimensions (both input and output). A GKS implementation without input is called level 'a', with synchronous input is called level 'b' and with asynchronous input is called level 'c'. An implementation without segmentation is called level '0'. With segments on specific devices, it is level '1', and if the segments can be moved between devices (using a workstation independent segment store) it is level '2'. The most common implementations are levels 0a, 2b and 2c.

The 'bundle' concept allows applications to support very disparate devices efficiently, such as a monochrome pen plotter and a full colour workstation.

The 'current position' is the location of a conceptual pen that draws the picture. Such a concept imposes an unnecessary, and arbitrary, sequential order on the graphics, which may inhibit multiprocessing or

parallelisation. This is why the concept was removed from GKS.

The philosophy for GKS is that the functions requested by the application are for almost immediate action. The segmentation facility provides an on-line method of storage of transient graphical information but is not designed for longer term storage between sessions. Once the workstation is closed, the segment store ceases to exist.

GKS recognised the need for the storage of graphical information between sessions and initially included within it a GKS Metafile facility which allowed an audit trail of GKS commands (used to create and manipulate pictures) to be stored and later retrieved and executed.

Metafiles are called such because they can contain metadata, information associated with a picture, but not actually necessary to construct the picture (e.g. a list of stations not plotted).

Annex E to the standard defines this GKS Metafile. The annex is not an integral part of the standard but, if present, will allow communications between GKS systems or long-term storage and auditing within a GKS system.

Once it became clear that there was likely to be more than one graphics standard at the functional level and all would have a need for long term storage and retrieval, it was decided to define the metafile function as a separate standard, the Computer Graphics Metafile (CGM). CGM is a facility for picture storage independent of, but still closely related to, the GKS standard.

The GKS standard contains a set of functions for reading and writing metafiles. The intention is that these functions could be used also to read and write CGM Metafiles.

6.4.2 Graphical Kernel System Metafile (GKSM)

From the programming point of view, the GKS Metafile looks very much like a workstation. Once the special workstation defined as a metafile is opened, any graphical commands obeyed are stored in the metafile. This continues until the metafile is closed. This similarity between a workstation and a metafile implies that there is a close relationship between the protocol used to define the metafile and that required to define the interface between GKS and the virtual device. (The standards activity to provide an interface to the graphical device is the Computer Graphics Interface, CGI.)

GKSM only defines a sequential file. Direct access facilities could be simulated with indexed sequential files in a reasonable manner, but the GKS standard only defines an interface for sequential access. Further, the interpretation of a GKSM metafile is dependent on the exact GKS implementation in use.

6.4.3 Computer Graphics Metafile (CGM)

CGM is the standard for 2D picture storage and transfer. CGM is defined as being compatible with GKS but allows a wider range of functionality so that it can be used for interchange of graphical information between other systems, not just GKS. A key element in the philosophy is that the process creating the information in the CGM can be separated in space and time from the process using it. Thus CGM could be used to generate a magnetic tape to be read at a remote installation many weeks later using a different type of graphics system from the one that generated it.

CGM is effectively transporting a virtual picture and, consequently, defines all picture elements in Virtual Device Coordinates which are closely linked to the Normalised Device Coordinates of GKS. CGM used to be known as Virtual Device Metafile (VDM).

The CGM description includes the coding of how the information is formatted. The CGM standard document now consists of four parts. The first contains the functional specification of any conforming metafile. The other three parts contain specifications of three methods of encoding, each with its own particular goal.

Character Encoding.

This is intended for use where it is important to minimize the size of the metafile; where necessary, this is regarded as more important than processing speed. This encoding makes it suitable for transmission through 'ASCII' networks.

Binary Encoding.

This aims to minimize the processor effort required to generate and/or interpret the metafile. It is therefore highly suitable for storage and retrieval of graphical data within a computer system.

Clear Text Encoding.

This encoding is aimed at the requirement of having a metafile that can be read and edited by people. It is also very safe to transport, even between systems with different character sets.

There are no facilities defined for direct access. A sequential format is obviously more straight forward and compact for both storage and transfer of data. Direct access could be defined but the overheads of direct access may be unacceptable. However most of the facilities of direct access could be supplied by indexed sequential formats without contravening the standard by using the facility to embed 'user data' within the graphics file.

Software to generate CGM files and to interpret them are widely available commercially, especially on PCs. CGMs are also recognised as a specific file type by the ISO standard File Transfer And Management (FTAM).

As the understanding of the graphics standards has improved over the last several years, extra features are now required to be transported or stored in CGMs, so a series of additions to the standard have been developed. These are called Addenda and there are three at present.

Addendum 1 supports segments, so that, for example, a map background need only be stored once at the beginning of a metafile, and is then invoked in each of the pictures within the metafile that require it. Segments were deliberately excluded from CGM to ensure compatibility with various levels of differing graphics standards. Attempting to superimpose segmentation structures onto a basic CGM file is a non-trivial exercise, requiring the recording of various segment transforms and attributes.

The Addendum 1 also supports Pixel Arrays. Standards such as GKS and PHIGS support Cell Arrays, which allow the display of fully transformable, device independent 'raster' images. As such, they are too expensive for most hardware and realistic applications. They have been shown to be useful for low resolution images, such as radar, and they do allow accurate overlays with vector graphics. The addition of Pixel Arrays allows higher resolution imagery to be displayed efficiently, albeit only in a device dependent way. For example, it would allow map coastlines, drawn as vectors, to be precisely overlaid on a satellite image, or a raster image from a scanner, PC painting program or PC screen capture program to be combined with vector graphics, though only correctly for one specific device at a time.

Addendum 1 is fully approved and published, and software is starting to be available.

Addendum 2 is a 3D extension and makes CGM able to support GKS-3D or PHIGS. The draft standard has been finalised and will be issued not as an addendum, but as parts 5 to 8 of the CGM standard and known as CGM-3D.

Addendum 3 will support extra features required by the more sophisticated hardware starting to become available, such as Non-Uniform Rational Beta Splines (NURBS) and sophisticated fonts for text. Externally defined libraries of symbols, such as the WMO weather symbols, will be usable. It is still a draft proposal.

ISO are proposing to update and revise the CGM standard, incorporating Addenda 1 and 3, and correcting a few errors and misprints. The revised standard will be completely compatible with the existing CGM standard (i.e.

existing CGM metafiles will still be interpreted and displayed correctly).

6.4.4 Computer Graphics Interface (CGI)

The existence of the GKS standard, the interface between an application and the graphical system, implies an interface (or interfaces) between the graphical system and the hardware, at the 'bottom' of GKS (or any other multi-device standard). This used to be known as the Device Dependent Interface, or Virtual Device Interface (VDI). Because the range of hardware to be supported is vast, from pen-plotters to workstations, CGI is a complex standard, with many options (which probably accounts for its unfinished state after 4 years!).

CGI only supports one device at a time, and does not have comprehensive error checking specified. It is appropriate, for example, for systems designers building graphics cards for PCs.

CGI is quite closely related to the basic CGM standard, which implements a well defined subset of CGI.

6.4.5 Graphical Kernel System - Three Dimensional (GKS-3D)

GKS-3D is a standard for the generation of pictures in a 3D viewing space. It is completely compatible with GKS (2D). It also supports 3D display devices, should these ever become widespread. Its main features are, in addition to the GKS (2D) features; (a) hidden line and hidden surface removal; and (b) perspective and orthographic views.

Its main advantage is in addition to the GKS (2D), multiple views of the same conceptual picture (unlike GKS).

Its main disadvantages are the same as GKS (2D).

6.4.6 Programmer's Hierarchical Interactive Graphics System (PHIGS)

PHIGS is a standard for 3D graphical applications, but has a more sophisticated hierarchical segmentation system than GKS-3D. The main feature is that as well as performing the graphics, it models the solids being portrayed. They are constructed hierarchically in terms of, typically, polylines and polygons. This modelled structure can be dynamically altered, independently of the graphics display. The model structure can be stored for archive purposes, whereas CGM would be used for capturing and storing the resultant picture.

It is suited to rapid dynamic interaction with models/pictures, such as simulated robot arms, or flight simulation. The storage and manipulation of the model causes performance problems on anything other than a very powerful workstation. PHIGS also requires very skilled programmers. It is hoped that in the future, software in the form of higher level tools will make programming and debugging of PHIGS programs easier.

PHIGS has language bindings to FORTRAN, Pascal, Ada and C in progress.

PHIGS was designed to be as compatible as possible with GKS and GKS3D.

6.4.7 PHIGS Plus (PHIGS+)

PHIGS+, pronounced PHIGS PLUS, is an extension to PHIGS which has now reached the Draft International Standard stage. The graphical primitives in PHIGS are identical to those in GKS, GKS-3D and CGM. However, some three dimensional graphics hardware can now support many more sophisticated entities, such as strips of triangles, lit and shaded polygons and splines. PHIGS+ will support these extra facilities, which would typically be used to construct pictures that approach photographs in their realism and quality ('photo-realism'). The PHIGS component is of course totally compatible with PHIGS, and has been shown to be relatively portable software. At present, however, the PLUS part, to do with lighting and shading, is not yet very portable.

6.4.8 PEX and X-Windows

X is becoming a widespread networking protocol amongst workstations connected by Ethernet. X recognises that most workstations support raster and vector graphics and windowing systems. PEX is a proposal to integrate PHIGS+ and the X protocol. It may be some time before PEX becomes an international standard, because of some fundamental technical issues involved in making X an international standard.

Programming X directly is extremely complicated, and is definitely not recommended. For graphical applications, it is much more preferable to program using GKS or PHIGS that have been implemented on top of X. Such systems are commercially available now. The current release of X is X11.5, and this included a PHIGS implementation, and X11.4 included a GKS implementation. For non-graphical applications, other software toolkits are available for use on top of X. Some of these can be used interactively to generate applications, or to configure sophisticated user interfaces, but without conventional programming. Very many are now available, and no specific recommendations can be made.

6.4.9 GKS9X

Because all international standards are reviewed every 5 years, GKS was reviewed in 1990 and a decision was taken to update, rather than continue it unchanged. This was because experience, the evolution of the other standards, newer hardware technology, and a better theoretical understanding of graphical issues have identified certain shortcomings in GKS. The revision is still at a Draft Proposal stage, but will be compatible with the current GKS and will support some of the features introduced in CGM and PHIGS. It will also introduce some features more powerful than those in PHIGS, on both output and input.

6.4.10 Postscript

Postscript is a proprietary format owned by an American company, Adobe. Strictly speaking, it is a low level programming language, but is invariably generated from higher level graphics packages (numerous GKS implementations have Postscript drivers). Numerous graphics devices have specific hardware to interpret the format. It can produce very sophisticated effects, but is very much less compact than CGM for meteorological graphics. It is unsuitable for storage or transmission over medium or low speed lines.

6.4.11 CCITT Recommendation T.4

T4 is a CCITT standard for the transmission of Group 3 facsimile. Its advantages are that fax machines are now becoming very widespread and are a very convenient method of dissemination of graphical products. T4 has the disadvantage that it can only transmit binary images (i.e. black and white, without even greyscales), though CCITT are working on a future standard for the transmission of greyscale and coloured images. T4 images are also much bigger than CGMs, and therefore unsuitable for long term storage of graphical pictures or transmission where bandwidth is a problem. Also, because they are bit maps, all structure in the original picture has been lost and it cannot easily be reconstructed for further processing.

T6 is a similar standard for higher resolution images, with very similar characteristics.

Software to generate T4 or T6 formats from high level graphics packages are only just becoming available.

6.4.12 ISO 10646

As more countries start to use computers widely, there is a greater and greater need to represent all of the alphabets, syllabaries and ideographic character sets properly and effectively.

ISO started working on standard 10646 (the original 7-bit internationalized ASCII is IS-646) allowing up to 32 bits to represent each character, with options for switching into 24, 16, 8 and 7 bit modes. In response to this very complex proposal, a consortium of computer companies, originally mainly American, but now world wide, developed Unicode, a fixed width, 16 bit encoding of all the required characters. To ensure that all the world's characters fitted into just 64K entries, the duplicate entries between Chinese, Japanese and Korean character set were eliminated after these countries agreed.

In 1991, the appropriated ISO committee agreed that Unicode could be encompassed by the DP 10646 standard, with it forming one 16 bit plane of the full 32 bit character space.

6.5 Representation of Image Data

Ideally, there should be only one standard that should cater for different image sources (satellite and radar) and for mapped or un-mapped imagery. In this respect, the proposal to extend the BUFR format to account for radar imagery is seen as conflicting with this basic conception. A possible solution would be to extend GRIB by including provisions for features such as run-length encoding, since there is no conceptual difference between image data representation and model output.

The International Standards Organisation (ISO) has defined a format for the storage and transmission of graphics and image files, called the Computer Graphics Metafile (CGM). CGM has the provision for including other data (metadata) in conjunction with the picture data, such as instrumental data or geographical coordinates. A standard format for embedding this meteorological data has not been defined.

An addition to the CGM standard has been defined to allow the efficient representation of imagery, and the efficient storage of repeated components (known as segments) of pictures, such as coastlines. This is known as Addendum 1 to CGM. CGM and Addendum 1 have been shown to be at least as compact as many proprietary file formats for imagery, contour and plotted charts.

Plotted charts are likely to become even more compactly stored when Addendum 3 is finalised in 1992. This will allow the storage of the meteorological weather symbols outside of the CGM in a library of symbols and text fonts.

6.6 Standard Environments

Two environments are emerging as the principal platforms for two-dimensional interactive systems: PC type systems, and UNIX workstations.

The typical PC system is based on the DOS operating system, and the VGA/EGA display boards. It is possible to enhance PCs with higher resolution boards, but these must be seen as alternatives rather than replacing the standard displays. In the case of developing countries, this environment is seen as particularly suitable due to the increasing availability of various types of PC. The use of a standard such as GKS is seen as very useful, since it provides a common ground for developing a variety of device drivers.

UNIX workstations are seen as becoming more and more widespread as, in general, they have better performance and screen resolution than a PC. In this environment the X-Windows system is the most widespread tool for system development, combined with higher level user interface tool-kits such as OPEN LOOK or Motif. Another important feature of UNIX workstations is the availability of common file systems such as NFS and network protocols such as TCP/IP (in the future, to be complemented by OSI standards).

High performance of three-dimensional interactive visualisation is only available from a limited number of vendors, all of whom offer UNIX-based workstations. There is a definitive preference for UNIX workstations and systems for three-dimensional graphics environments, subject to the problems of standardising the development environment as discussed above.

APPENDIX A

A standards document goes through various stages within ISO before becoming an international standard. These are:

> Work Item, recently renamed Working Draft or Committee Draft. An official project within ISO with agreed scope and goals, assigned to an appropriate committee;

Draft Proposal. The appropriate committee produce successive Working Drafts until it is considered mature enough to be registered as a draft proposal for a standard. DPs are circulated within ISO for technical review and ballot. If it passes the ballot it goes on to be a draft standard, otherwise it is revised in light of the review and becomes a further draft proposal;

Draft International Standard. When sufficient agreement is reached on the DP document, it is registered as a DIS and made available publicly. This should indicate that technical agreement has been reached. The document is circulated within ISO for editorial review and a DIS ballot;

International Standard. The DIS, revised in the light of the DIS ballot becomes the Final Text. After another final ballot, the Final Text is published as an IS. The document will be reviewed after five years and will then be either endorsed, revised or abandoned.

In theory, all technical changes should be complete by the DIS stage and this is, therefore, a reasonable time to start using the standard. However the move from DIS to IS sometimes includes some minor technical revisions, if only to remove ambiguities. If a proposal has only reached the DP stage, it is quite likely that significant changes will appear before it reaches DIS status.

The complete process, from WI to full IS may take at least two or three years. If standards have been developed outside of ISO by another acceptable international body, such as CCITT or IEEE, they may follow a 'fast track' and be adopted directly at the DIS stage.

CHAPTER 7

DATABASES IN A METEOROLOGICAL ENVIRONMENT

7.1 Introduction

WWW DM is concerned not only with ways of representing data for transmission and display, but also with identifying efficient methods of storage and retrieval. Through the 1960s and 70s the data base requirements of meteorological services tended to be unique, and given the small size of the sector, poorly served by the commercial sector. Over the past few years however, commercial databases have grown in both size and stature to the point where they can now be considered for use in some meteorological environments. The factors which facilitate this are the availability, relatively cheaply, of large scale computing environments and the adoption of, and adherence to, rigorous standards of both the database itself and the interface to the database through Structured Query Language (SQL).

7.2 Database Technology

A database is generally accepted to be a set of data controlled by a Database Management System (DBMS), which supplies defined standard mechanisms for the storage and retrieval of the data. These mechanisms allow logical associations to be defined and for associated data to be retrieved together, and for details of any physical structure, as opposed to logical structure, of the database to be hidden from the users. The DBMS also has procedures hidden from the user for backing up, archiving and restoration in the event of failure. A single DBMS may control a number of apparently separate databases.

As these features are programmed into the DBMS once, the applications using the database can be made simpler, and hence easier to maintain. If the physical format of the data needs to be altered to improve efficiency, the applications need not be aware of this change. Similarly if new data types are added, the applications only need altering if they need the new data.

If application software makes use of explicit relationships between data, it can be much more efficient, in computer resources, than using a general database. However, as the range and complexity of data relationships

increases, the increasing burden of application software maintenance favours the use of a database, especially as hardware continues to decrease in cost relative to software.

7.2.1 Indexed Sequential Data Bases

The simplest databases store data in records, and complete records have to be retrieved to inspect any data within the record. The records often have an Indexed Sequential structure. Many cheap personal computer databases are based on this technique. The WMO Volume A catalogue of stations has this kind of structure. Some items of information may be duplicated numerous times in different records. The physical structure of the data is not usually hidden from the applications retrieving the data.

7.2.2 Hierarchical Data Bases

Hierarchical databases define the data in terms of 'trees', so that duplication is minimised. Such databases allow efficient retrieval of data, providing it is required according to the pre-defined tree structure. These databases then evolved into network databases, where more flexibility in different ways of accessing the data was allowed. Such databases are called CODASYL after the committee that originally defined them. The networks have to be defined in advance when the database is constructed. Data of the same kind is often linked by pointers, and applications 'navigate' around the data base, retrieving linked data in successive retrievals.

7.2.3 Relational Data Bases

The relational data base is based on the mathematical theory of relations (set theory) and as such has a sound underlying basis. It was developed in the early 70s at IBM by Ted Codd and his associates. The basic data structure of the relational database is the table which is used to represent a relation. Each row of a table represents a different record and each column of a table represents a different field. The order of rows or columns is insignificant. One of the main tenets of relational database theory is that one fact is stored in one place in the database.

There is a separate theory called "normalisation" that is concerned with arranging the structure of the tables in a database (which columns appear in which tables) to ensure this. For each table, a group of columns is used to uniquely identify each row. That group of columns is known as the primary key. A foreign key is one or more columns in one table that serve as the primary key of another table. This allows joining together of information across tables. Computer science has developed, amongst other languages, the relational calculus for manipulating data in the various tables of a database. It is non-procedural in that it specifies what data to use rather than how to access that data. Relational calculus has the property of closure in that the result of applying a relational calculus expression to a table is another table. SQL is based on the relational calculus. It also includes a means of defining the tables, and creating and deleting them as well as formatting and security features. For a system to be relational, it must at least:

Represent all information as values in tables;

Have no user visible links between tables;

Support the relational calculus (or an equivalent).

7.2.4 Semantic Data Bases

If the relationships between the data changes faster than the data, it is more appropriate to store the relationships (links) between data items explicitly, rather than as tables. Such a database is known as a semantic database. These are not widely available commercially, and consume even more computing power.

7.2.5 Object Oriented Data Bases

Object-oriented databases could be considered an extension of semantic databases, where items and their behaviour are stored. For example, a SYNOP message could be stored, and the database would link it to the software that knew how to plot a surface synoptic plot model, as opposed to a contouring package. A GRIB or GRID message would be linked automatically to the contouring package but not the plotting software. Object-oriented databases are not widely available and are not well understood, and will undoubtedly be expensive! They do not necessarily impose a well-defined structure on the data.

7.3 Query Languages

Most commercially available databases have a query language for querying the contents. Such languages may be interactive, or embedded in application programs. The latest are all non-procedural in that they specify logically what data is required, rather than 'how', i.e. specifying the procedure that the database must follow to extract the data.

The Standard Query Language (SQL) is such a language that has been standardised internationally (ISO 9075:1987) and is supported by most commercially available databases. SQL is based on the relational data model, though it does not necessarily have to be used with a relational database.

7.4 Data Base Technologies Employed by Meteorological Services

Modern data base concepts and technologies are very attractive because they provide users of data, and information derived from that data, with the tools to maximize the return from their investment in that data. Were meteorological services commencing operations from the present time, given adequate resources it would be relatively easy to adopt the new technologies. Most weather services however have been providing services for a considerable time and have a legacy of data accumulated under traditional data management practices.

Traditionally, applications built and used by the meteorological community were considered in relative isolation. Data that should have been together was not. The potential for flexible enquiries and reporting was thus limited. Additionally, the same data was held in several places on several different computer files. This again lead to applications in relative isolation and the fear of interfering with the data files of an existing application as well as update synchronisation problems if copies of the files were kept at several sites. From these experiences many lessons have been learnt concerning the construction of more flexible, and efficient meteorological data bases.

In particular, the storage formats, and retrieval formats, should be as similar as possible for different databases. Then retrieval and storage programs could be reused and software simplified. Ideally, the retrieval formats should be portable, preferably across networks, to allow the distribution of applications where appropriate.

A logical retrieval interface for observational data (i.e. retrieval by element, time, location, level, etc) is applicable also to the retrieval of field data from databases. Because an entity in such a database is generally two or three magnitudes bigger in size than an observation, and two orders of magnitude fewer in numbers, it may not be appropriate to store these data in the same database as observations unless efficiency is sacrificed. (An observation is typically 100 Bytes, but a radar image may be 64KB, and a model field is similar.) Physically separate databases could, and where possible should be implemented with the same interface.

7.4.1 Relational Databases and Meteorology

As noted above, the three broad categories of meteorological data are observations, model output in the form of gridded fields, and satellite and radar data in the form of images. Unfortunately, two of these - grids and images - do not conveniently fall into any type of relational database type of field - these basically being numbers and characters. However, with the fax and image revolution, a lot of vendors are now supporting another type of field in the form of variable length bit-strings called BLOBS (Binary Large OBjects). These can be used for the storage and retrieval of gridded fields and image data. However, the database knows nothing about these fields other than their length. The concept of differencing two BLOBS, for example, (when it makes sense to do so) may not be supported by the database whereas differencing two numeric fields obviously is. The database may support users writing code to perform operations like these - called BLOB filters - and store this code centrally

in the database but this facility is not universal. Observational data may be stored in the database and the database interrogated and data retrieved quite easily. The inhibiting factor here is volume.

Some database designs have used the communications standard BUFR and GRIB for storage of meteorological data. BUFR and GRIB cut down the size of data being stored and thus reduce access and/or storage time. This may be particularly relevant for NWP modelers. The packing or unpacking of data can be overlapped with storing or fetching of the next record for efficient processing. This design works especially well with GRIB where there is little variation in how the data is to be viewed (usually, gridded data is needed in its entirety for processing) but again, translating the observations into BUFR code means that the database needs to be taught the structure of the resultant bit-stream, if that is possible, to enable queries to be posed.

7.4.2 Relational Database Products and Costs

There are many commercially available products with many different features, it is beyond the scope of this Guide to review this ever changing set of products.

The prices for commercial, relational databases differ from vendor to vendor and depend on the hardware platform on which the software is being run. A rule of thumb is that costs are split 50%-50% for software-hardware for systems totalling up to US $500,000 (1992 dollars) after which the hardware tends to dominate the costs. Of course, for the definitive costs, vendors are more than happy to provide price lists.

Advantages

A common interface to the data through SQL means that both users and programmers do not have to learn more than one access method no matter how many systems they work with.

It has been suggested that productivity using relational databases is improved 5-20 times over other approaches.

The other features including security and backup that come with relational databases means programmers can concentrate on building applications.

Sites have the ability to change vendors of their relational database software with minimal impact on their systems if they adhere to standards. Standards are now being set to allow multivendor database interoperability.

Establishment of better communications standards allow wide area networking of databases as well as full binary data support rather than being restricted to character codes.

Disadvantages

Installation costs can be substantial, and as such, every last feature provided must be used to give as cost efficient as possible implementation.

Planning and design work necessary for the introduction of a new database leads to significant new overheads. Furthermore existing applications must be converted to interact with the new data base and staff must be trained to use the new technology.

There are overheads in accessing data held in the new data base structures; generally these overheads take the form of additional computer hardware needed to host the operations. Additionally, support and maintenance becomes more complex with the more complex data base environment.

Users and management must understand the benefits of relational database technology. In particular, users may no longer feel that they individually own data they routinely use. It becomes a corporate entity in the new environment. Some users may resent this.

Meteorological applications tend to be resource hungry making data access times in the new data base environments appear slower than for a traditional file structure built to service a single application. This aspect

is accentuated if hardware support for the data base is marginal and data volumes are increased substantially; for example, by doubling the resolution of a numerical model which feeds the data base. Thus, meteorology will always be tending to test both the limits and robustness of both the relational model and more importantly vendors software.

A quite sophisticated infrastructure of personnel needs to be established to support all of this new technology.

7.5 Storage of Quality Control Information in Meteorological Data Bases

In general it must be expected that meteorological data processing systems will generate Quality Control (Q/C) information, and require that the data base is designed so as to allow for it to be stored in such a fashion that it is logically associated with the original data. The two major options for meeting this requirement are either to store the Q/C information in the same database, or to store it separately, but with keys to associate it with the original data.

There is a continuum of Q/C of observational data: internal consistency checking of reports, checks against climatological values, checks against forecast (background) fields, checks against surrounding observations, of the same or differing types. Checking may generate substituted values, or even completely new (bogus) observations.

If meteorological processing of the data is performed within a database, the processing software should be written in a standard portable language, such as FORTRAN 77. Alternatively, Q/C could be performed on observational data by external processes and then allow the Q/C information to be stored within the database.

The application of Q/C by the climatologists may extend over a relatively long period (some data arrives at least a month late), and therefore, some data could be entered into the database at low priority, such as at weekends or overnight. Adding this data should not seriously affect the rapid availability of data, without comprehensive Q/C information, for forecasting purposes.

The Q/C applied by NWP systems is much more time critical - the bulk of it must be done within the context of a forecast suite (i.e. within 12, 6 or even 3 hours, depending on the numerical model). The Q/C information is typically 3 to 4 times greater in volume than that produced for conventional data. The NWP suites also use satellite data, in greatly increasing quantities.

The amount of satellite data is expected to grow drastically as various new satellites transfer from research to operational status. It seems likely that volumes could easily reach 300MB per day by the mid 1990s, even when using efficient modern binary codes for exchange and storage.

It may not be appropriate to store Q/C information in a separate database as the logical links might be difficult. Applications would need to interface to two databases, increasing the cost of software and its maintenance. The approach of automatically updating a second database regularly from the first, and have the second database contain both the original data and the Q/C information, would be considerably easier and has, in effect, been followed in existing systems even though this duplicates the storage of data.

Adding Q/C information can be construed as a risk to the integrity of the GTS data. Therefore extra safeguards, such as passwords and physically separate storage for this information, should be considered. As the requirements for security and integrity features increase, the complexity of implementing them increasingly favours the use of a proprietary DBMS.

Facilities for the monitoring of the flow of the observational and associated data should be incorporated into the database. Such statistics would include percentage failure of the Q/C checks, timeliness of arrival in the database, as well as computing resources consumed. Such statistics are essential for real-time usage and are also useful for climatological data. Such meteorologically orientated statistics are unlikely to be available in commercial databases.

CHAPTER 8

THE DISTRIBUTED DATA BASES CONCEPT

8.1 Introduction

At CBS IX the "Distributed Data Bases (DDBs) Concept" was endorsed and the Working group on Data Management tasked to further develop it, and to suggest an implementation plan, in order that emerging data management problems (felt to be caused by a lack of integration between the GTS, GDPS and GOS, and by the pressure of newly emerging technologies and user groups) could be tackled and solved. This plan is to be brought before CBS for further consideration. This chapter describes the evolution of the DDBs concept to the present time.

8.2 The Relationship of the DDBs to the existing WWW infra-structure

At the present time it is believed that the most efficient delivery of basic data (observations and products) in the WWW can best be supported by an improvement and enhancement of the data management functions within the GTS. These, together with the necessary upgrade of the transmission capacity of the GTS, will lead to improving the routine data distribution to meet the future basic WWW data exchange requirements.

Beyond the exchange of the basic data there is also a demand for the provision and exchange of new data to be facilitated by the WWW. On the one hand, it is expected that the WWW can play an important role in the provision of support services to other WMO Programmes, (for example, GCOS and GOOS) on the other hand, a vast amount of auxiliary data and reference data exists for which specific data sets are needed in irregular intervals for the direct support of WWW operations. Both requirements suggest that it is timely to consider the development of a flexible, modern DDBs infra-structure. The DDBs concept should meet the requirements for a data handling system for data needed in the WWW system but not routinely exchanged on the GTS, and support the new and emerging needs for data, including those which originate outside of the WWW system.

One way of meeting these needs is to make information available concerning the nature and location of a variety of meteorological and environmental data sets and develop specific purpose databases. Such databases would help centres to better respond to specific tasks within the WWW systems and also assist centres in tasks that are not necessarily central to the traditional WWW operations. The information should be made known to the meteorological community through the provision of the corresponding metadata (data about data), without necessarily the need to make these data sets themselves available on the same network (e.g., GTS) on which the metadata are obtainable. Depending on the nature of the data requirements, different means of communication and/or telecommunications are possible for sending the requests and replies.

Because the GTS is not well suited for the implementation of an any node-to-any node request/reply system by virtue of its predominant role to meet routine requirements, and because of its prevailing architecture, it is recommended that non-routine access to the DDBs be primarily via public networks (e.g., datalinks over the telephone network, public packet switched data networks or ISDN) so that the performance of the GTS is not degraded. This introduces a "user pays" mechanism which will also serve to protect the DDBs request/reply system itself from being monopolised by extravagant users.

Should the frequency and importance of obtaining certain data reach a point where it is economically justified and internationally agreed to use the GTS, then the communications links can be upgraded accordingly in a planned manner and such data transfers moved to the GTS.

Use of an open systems architecture and of public data network facilities, whilst providing greater flexibility and a higher level of availability than fixed dedicated circuits of the present closed GTS system, will make the system more vulnerable to unauthorised access than has been the case to date. It will be necessary to mandate minimum security measures to be implemented by participating Members. In addition to these "standard" minimum measures, Members may feel the need to implement additional security features. The implementation of security measures will not be costless.

8.3 Implementation of the DDBs

Figure 8.1 provides a schematic view of the structure of the DDBs and their relationship to the GTS. This logical description has the following features:

An upper "DDB plane", in which all databases are logically inter-connected;

On the upper plane there is any node-to-any node connectivity, with bulk data transfer arrangements in place, but whose operation would not diminish the efficiency of the "basic" GTS;

The upper plane could possibly be a "managed network";

Standardisation of upper plane database structures (possibly using the relational model);

There is a gateway between the DDB and RTH and NMC;

A lower layer which is the "basic" GTS;

The basic GTS continues to store and forward essential observational data and meteorological products between adjacent nodes, with improved communications, automation of routeing tables, streamlined headers, etc.

The need for a catalogue system in any new database is fundamental. Such a catalogue system will allow that participating databases can "know" what data is held by other databases. A database catalogue contains metadata. For a catalogue to be useful there must be ways of "browsing" a database's catalogue from a remote system.

Modern, relational databases are built using a series of tables which contain an in-built system of defining the relationships between tables. This table structure lends itself to building a catalogue, or table of metadata, to meet the DDB requirements.

8.4 A Possible Implementation Strategy

A strategy is emerging for achieving the implementation of the DDBs. In essence, the strategy is to identify those elements which must be present in the DDBs and commence to build them. It is currently anticipated that prototypes will be built first to be followed by limited implementations between willing volunteers after this, and finally after consideration by CBS, adoption of the successful techniques and technologies as WMO standards. Application of such a strategy as this requires careful selection of the first demonstration implementations. Later, the operational implementation of the DDBs could be done following one of two basic models:

(a) By further development of the database systems within the existing communications and computing infra-structure of Members, to agreed standards, to form a loosely coupled (all peer) set of WWW DDB;

(b) By implementing a separate database system utilising a commercial database management system, possibly running on a standardised hardware platform (e.g., a moderate sized UNIX machine and/or application of Structured Query Language (SQL), as part of a managed network.

The second option could offer considerable savings in software development costs and may harmonise well with the plans of some centres for the upgrading or further development of their national systems. Furthermore, modern database systems are available that can handle large binary objects and are, therefore, potentially suitable for WMO DDB applications, particularly if in the longer term the DDB are to provide back-up facilities for major centres.

8.4.1 A Possible Prototype Element of the Set of DDBs

It is generally assumed that any specific DDB consists of data and/or metadata residing at a specific location responsible for this data and of the pertinent mechanisms for accessing and retrieving the data from outside of this specific DDB location. The first prototypes will ideally use similar database management systems and employ the same access methods via the same type of communication system. If possible, there will be some complementarity of functionality so that Members gain more from the coexistence of a number of databases than if each database were a separate, distinct entity.

Possible prototype DDBs can be grouped into the following three categories:

Category I Data and/or metadata germane to routine WWW operations; requests should be possible for the metadata, the data itself, or both and could be responded to via the GTS or by other communications or telecommunications means, as appropriate.

Category II Auxiliary data and related metadata needed for supporting specific tasks within the WWW system and beyond; this type of data, and corresponding metadata may reside in or outside of a WWW center; requests should be possible for the metadata, the data itself, or both, and could be responded to via the GTS or by other means, as appropriate.

Category III Metadata (only) describing data that exist in WWW centres or in other centres which are of importance to the meteorological community in general; requests should be made mainly outside of the GTS (i.e., using non-GTS means); the exchange of the metadata should also not use the GTS, but other means including, for example, common mail services.

Table 8.1 contains examples of data sets and/or metadata which would typically fit into the DDB concept. Some of these sets are already in existence in a suitable form, others will have to be generated or adapted.

Table 8.1 - Examples of data sets to be considered for ad hoc exchange using DDBs: (NOT AVAILABLE)

Category Data sets

I WWW data and/or metadata:

- Catalogues on data and products existing at centres
- WMO Publication No. 9, Vols. A, B, C
- Information on the operational status of centres
- Data sets containing satellite remote sensing data
- Metadata relating to numerical products, e.g., skillscores
- International and national BUFR tables
- Information on changes to forecast models

II Auxiliary data sets

- hydrological
- climatological
- agro-meteorological
- oceanographical information, observations products
- reference data, e.g., information on droughts, locusts

III Metadata

- climate data sets
- environmental data sets
- image data of earth observing satellites
- TOGA data sets
- Long-term re-analysis data
- Information on available meteorological application software
- Experimental data sets (e.g., FGGE, GARP/GATE,Wetnet, AMEX)

Any DDB constructed should have the potential to serve both the WWW operations and also users outside of the WWW system, in a similar way as the WWW system itself provides basic services to other WMO Programmes. Users in meteorological centres would benefit from gaining access to information of the types listed in Table 8.1, which could be provided in DDB through their centres' communications system, and which would not be conveniently available through other means. Centres participating in the DDB services would have to dedicate resources towards the implementation of such services, e.g., creation/adaptation of locally maintained data sets, access mechanisms to data sets residing outside, and equipment and services that enable local users to access the data sets. Before any resources could be committed, it is important to establish a full understanding of the ultimate requirements of the users and the benefits to be gained. The Working Group on Data Management is working to further identify requirements and volunteers to establish specified data bases.

8.5 The Role of the DDBs in Data Collection

The DDBs concept has been developed to provide improved data access to users who have the following characteristics:

They do not necessarily need the data in real-time;

They have access to the modern communications and computer systems necessary to handle large volumes of data quickly;

They do not necessarily need the data (or metadata) routinely, but possibly on an ad-hoc basis.

It is believed that the concept also provides potential advantages to those undertaking data collection associated with special environmental monitoring projects which are at least regional and possibly global in scale.

Currently data can be ingested into the WWW real time system at any site where there is a telex, or some computer system able to communicate with a Regional Telecommunications Hub (RTH). This data insertion can be done through slow speed lines (e.g. 50 baud) or by high speed data links. Traditionally observations were transmitted to RTHs at low speed (via telex) and collected into bulletins for exchange on the GTS.

More recently satellite collection of data (from DCPs) has led to the substantial accumulation of bulletins of data at the sites of high powered computer systems. Clearly it would be possible to establish elements of the DDBs at these sites as well as undertake the more traditional insertion of these data into the GTS as collectives.

Figure 8.1 - A Schematic View of the Structure of the DDBs and their Relationship to the Existing GTS Infra-structure (Not Available)

CHAPTER 9

DATA MONITORING

9.1 Introduction

For a complex system, such as the WWW, the achievement and maintenance of a high level of operational efficiency and effectiveness will only occur if there is an ongoing program of system monitoring, and if the results from this monitoring are taken into consideration by those responsible for the various components of the system. Within the WWW there are two distinct aspects of data monitoring - data quantity monitoring and data quality monitoring. Data quantity monitoring is concerned with the timely collection and receipt of observations and products, while data quality monitoring is concerned with the accuracy of the observations and products. Because the WWW is fundamentally reliant on the quality and quantity of data available within the system, data monitoring is a crucial activity.

To achieve success, the WWW needs to:

Provide sufficient observational data of acceptably high quality, collected in a timely manner;

Provide data transport facilities, within appropriate time frames, to enable observational data to be made available both to direct users and to centres responsible for the generation of products;

Generate products (analyses, forecasts, climatology, archives, etc.), within appropriate time frames, sufficient for the needs of the WWW;

Provide data transport facilities, within appropriate time frames, for the dissemination of the generated products to the users;

Generate, accumulate, and where appropriate, exchange metadata concerning the accuracy, timeliness, efficiency, and completeness of the critical components of the system listed above.

The performance of the WWW depends on the quality of the observational data, the efficiency and completeness of the data generation, collection and transport, and the quality of the generated products. Data monitoring is a means of assessing these factors of performance; the design of data monitoring procedures should aim both to identify possible deficiencies with respect to quantity and quality, and to indicate how those deficiencies might be rectified.

The role of Data Management with respect to data monitoring is that of an integrating function. Certain aspects of data monitoring need to be defined, and/or performed within the context of the GOS, the GDPS, and the GTS. All aspects of data management need to be co-ordinated within the context of the over-all WWW system. It is the function of DM to ensure that such co- ordination is achieved. This requires that:

The appropriate metadata be identified;

The need for data monitoring procedures are sufficiently defined for the working groups responsible to act;

The procedures defined by the working groups are consistent and sufficient;

That requirements of working groups on other working groups are sufficiently defined and can reasonably be met;

Data representation forms suitable for the transport and storage of monitoring information are developed;

Sufficient arrangements are made for the post-processing of monitoring information to ensure that the WWW systems efficiency is understood, areas of weakness identified, and remedial action undertaken where feasible.

The actual definition of procedures, and their implementation is primarily the responsibility of the working group concerned, taking into account advice and suggestions from the Working Group on Data Management, using the appropriate representation forms to represent monitoring data produced, and conveying to the Working Group on Data Management what co-operation and assistance are necessary from the other working groups.

9.2 The Identification of Errors and Deficiencies within the WWW

Table 9.1 illustrates some of the major WWW functions performed by components of the GOS, the GDPS, and the GTS. Data deficiencies may result either from data of poor quality, and/or from the unavailability of potentially useful data. The study of these two forms of deficiency are referred to as "data quality monitoring", and "data quantity monitoring" respectively.

Some likely causes of deficiencies with respect to data quality and quantity resulting from the major WWW functions are listed in Table 9.1 below.

It is potentially possible to identify and rectify many of these deficiencies. The monitoring required to achieve this entails the expenditure of resources to examine data, compile various statistics and system performance information (trace information, quantitative difference statistics, bias assessment, quality indicators, etc.). Some assessment can and should be performed in near-real time, whereas other evaluations can only be accomplished after the gathering of sufficient data over a long period. The aim must be continually to endeavour to develop the most cost effective means of monitoring data as near as possible to the beginning of their life cycle without the resulting processes causing excessive delay to their dissemination within acceptable time frames.

Table 9.1 - Possible source of WWW data deficiencies, as they occur in the various WWW components observing, collecting, etc.

OBSERVING	
Quality	Quantity
- instrumentation error	- observation not made - not made at right time
- encoding error	- measuring error
	- made but unable to send
COLLECTING	
Quality	Quantity
- formatting error	- not received
- detectable validation error	- received late
- not transmitted	
- transmitted late	
TRANSMISSION	
Quality	Quantity

- formatting error	- not received
- received late	
- not transmitted	
- transmitted late	
- not correctly routed	

RECEPTION

Quality	Quantity
- formatting error	- not received
- transmission error	- received late

PRE-PROCESSING

Quality	Quantity
- formatting error	- not received
- detectable validation error	- received late

EVALUATION

Quality

- detectable validation error
- mutual inconsistency
- temporal inconsistency
- bias
- forecast error

9.3 The Role of the GOS

The maintenance and support of an observational system requires considerable investment. The objective of data monitoring within the GOS is to endeavour to ensure that such investment leads to a maximum return in terms of the value of the observations produced. Observational data are reduced in value:

- When they are late;
- When they are subject to error;
- When not delivered to potential users.

The identification and rectification of such deficiencies, where possible, as part of the observing process is particularly effective, as it may avoid the need for remedial action at a later stage.

9.3.1 Quality of Data

It is a function of the GOS to define:

- Acceptable standards of observational practice;
- Quality control procedures to be applied during the observational process;
- Information on expected accuracy to be associated with (and possibly included within) observations from particular observing systems;
- Checking of the formatting of observational data prior to general distribution.

Guidance material, drawn up by the GOS, concerning the above should also reflect and take account of any expressed needs of the users of the data.

9.3.2 Quantity of Data

The collection of observational data, checking, monitoring of the completeness of reception, and subsequent assembly into suitable messages for further distribution form a time critical path within the WWW data flow. The GOS responsibility is to define:

Acceptable procedures for the timely completion of these processes;

The compilation, updating, collection and, as appropriate, exchange of metadata recording the progress of these processes;

Appropriate means to take action when non-receipt of data is detected.

It is not possible to detect and address deficiencies in the timely distribution of data by monitoring the data transport performance of the GTS alone; some delays may occur in the collection process before data are passed to the GTS for onward transmission.

9.4 The Role of the GTS

Efficient data transport is essential for an effective WWW. The first priority of the GTS is to disseminate the observational data generated within the GOS, together with such vital information as warnings of severe weather. A further important function is the dissemination of appropriate products generated within the GDPS.

Within the GTS, data are handled as collective entities termed "meteorological messages", conforming to GTS defined standards. It is not a function of the GTS to be concerned with the data within such messages or bulletins, excepting such information as may be necessary for message recognition. In consequence, monitoring within the GTS must be in terms of messages rather than in terms of the individual data items contained therein.

9.4.1 Quality of Data

It is a function of the GTS to ensure that data disseminated are correctly formatted, both before and after transmission, and to develop procedures to ensure that this is done.

9.4.2 Quantity of Data

Messages passed to the GTS for dissemination need to be routed to appropriate recipients, and delivered as soon as possible; some messages are of more importance than others, and this order of importance is reflected in the GTS system of allocation of priorities.

To ensure that deficiencies can be identified and remedied, the GTS needs to develop:

Appropriate systems to monitor performance;

The collection and exchange of performance metadata;

Methods to ensure that routing information is kept up-to-date, and regularly checked.

9.5 The Role of the GDPS

The GDPS depends on the timely delivery of sufficient observational data of an acceptable quality as input to those processes which generate analysis and forecast products. The processing of the observational data results in quality related information which may often be significant in the early detection of problems not easily detected as part of the functions of the GOS. Other quality information is significant in the longer term, indicating the existence of systematic bias in either observational data or products.

9.5.1 Quality of Data

The GDPS in conjunction with CBS has developed a system of quality monitoring for observational data based on monitoring statistics generated or co-ordinated by lead centres for various types of data.

It is the function of the GDPS to define how these procedures can be further developed and improved. Feedback of data monitoring results to the data producer is considered to be of prime importance. In particular, it is believed that, whereas currently feedback to originators in delayed mode has proved to be effective, there might be a case for some feedback to take place in near-real-time.

A further function of the GDPS is to monitor the quality of the products they produce, together with the quality of products received from other centres.

9.5.2 Quantity of Data

Within the data transport mechanism of the GTS data are handled in terms of messages. Messages may be delivered, but they may contain many observations which are "nil". Thus, it is only at the level of the GDPS that the observational contents of messages can be fully monitored.

It should be a function of the GDPS to compile, assemble, collect and, where appropriate, exchange suitable metadata to enable monitoring at the level of the individual observations and products to be achieved. To enable such monitoring to be effective it may be necessary to retain within the metadata information concerning the message in which the observation or product was found, and the time of receipt of that message.

9.6 The Role of Data Management

As an integrating component, the first priority for Data Management is to monitor the monitoring procedures developed within the other CBS Working Groups. Continual liaison is necessary to ensure that the Working Group on Data Management is aware of the needs, developments, and potential for further development of monitoring procedures within the other working groups, and is able to suggest suitable data management techniques to assist in the development of sound monitoring procedures.

Data Management is also responsible for the development of data representation forms, and must respond to the need to represent data appropriate to monitoring techniques.

Since some of the monitoring information cannot be assessed purely within the scope of any single working group, there is a case for the Working Group on Data Management, together with the WMO Secretariat, to be directly involved in the evaluation of monitoring results.

CHAPTER 10

COMPUTER SOFTWARE EXCHANGE

10.1 Introduction

The improvement of data management techniques in the WWW is largely built around the creation and implementation of computer based systems. These systems provide greater functionality than was previously possible, provide access to data and information not available through manual means, and at the same time ensure that the integrity of these data and information is preserved.

The types of data and products that can only be generated, handled, pre-processed, or displayed by means of computers include:

1. Handling of WMO binary codes (GRIB and BUFR);
2. Performing conversions between binary formats and character codes or between different representations of graphical data (e.g., T4-coded facsimile, CGM, raster graphics, vector graphics, etc.);
3. Receiving and using satellite images;
4. Participating in satellite based point-to-multipoint broadcast services;
5. Accessing and using data generated by automated observing systems;
6. Receiving and using numerical weather prediction products and generating value added products;
7. Generating forecast model products.

In order to meet the data management needs associated with these data and information WWW DM maintains a registry of software Members are willing to make generally available to the meteorological community. To ensure that the software offered are of maximum utility to Members WWW DM must also work to advise on the most appropriate standards to which software should be developed and the techniques for doing so.

10.2 Software Exchange

10.2.1 The Concept

In order to improve the computerised data management capabilities of Members WMO has undertaken a number of initiatives aimed at providing various kinds of assistance in form of donations of computer hardware and applications packages, such as the MSS (Meteorological Messages Switching), graphical display software and dedicated training under the SHARE programme (Software Help, Applications, Research and Education Programme). These activities are being carried out as VCP/UNDP projects. Support of this type is rather demanding in terms of organizational requirements, engineering support and, above all, financial resources; accordingly only a relatively small number of countries can benefit from these projects in any given financial planning cycle of WMO.

Countries that are in the process of automating their meteorological services have a steadily growing requirement for meteorological applications software to meet newly emerging requests for products in their own country, and to stay abreast with procedural changes agreed by WMO (such as changes in codes or amendments in telecommunication procedures). Some have recently acquired the capability of handling the X.25 telecommunication protocol for data exchange on the GTS and wish to receive data formatted in GRIB and BUFR. It is difficult for these centers to obtain the necessary data handling software for processing binary-formatted data.

It is, therefore, important that WMO activates other resources to help Members in acquiring software and related technical assistance. CBS felt that it may be a promising and cost-effective way to encourage Members to exchange software that is already available in meteorological centers. This approach raises a number of questions pertaining to software compatibility and portability.

Three trends in computer applications will make the porting of software between different computer centers simpler. Firstly, standardization in the meteorological community is leading to the introduction of common data formats, data handling and telecommunications procedures. Many basic functions performed by computer programs in meteorological centers are essentially very similar or even identical. This is particularly obvious for real-time functions such as handling of WMO-formatted messages, handling of WMO-coded reports, plotting of station-model and contour charts and for a wide range of non-real-time data management functions. Suitable programs are available in meteorological centers of most developed countries. Secondly, on-going efforts in WMO in developing WMO-agreed standards for software designing, programming techniques and software documentation will gradually alleviate the level of incompatibility in meteorological computer solutions. Thirdly, international and industry standards spread aggressively and are now more readily accepted in the software laboratories of meteorological services than some years ago.

In addition, many computer vendors offer in their line of products conversion programs for converting a program source code from one industry-specific format to another, such as from an IBM FORTRAN dialect to a Digital Equipment FORTRAN dialect. These type of tools reduce the amount of work required to adapt computer programs from other sources.

All this lessens hardware-induced incompatibilities of computer programs and make it possible for Members to contemplate adapting other centers' operational software for use in their own operations vis-a-vis "in-house" development or commercial procurement.

10.2.2 The Objectives of WWW Software Exchange

After technical surveys by its working groups, CBS-Ext.(90) requested that the WMO Secretariat organize an exchange of computer software among WMO Members.

The objectives of the CBS Software Exchange are to:

a. Strengthen the WWW system by making more readily available a wider range of suitable meteorological software to Members for their own application;

b. Improve the self-sufficiency of evolving computer centers of developing countries by encouraging such Members to participate actively in the exchange of application software;

c. Assist in spreading well-proven software packages and standardized software techniques in the meteorological community;

d. Provide an overview of meteorological applications software offered and requested by WMO Members;

e. Assist in planning WMO co-ordinated computer projects by connoting the most prominent and most common requests for software and/or computer support.

10.2.3 The Software Registry

The WMO Secretariat has established a catalogue containing information on offered and requested computer software collected from Members, and will publish an updated edition on a yearly basis. The first edition of this catalogue, called CBS Software Registry, was distributed to all Members in January 1991.

In order to provide a manageable structure for the exchange a framework of categories has been established under which the individual computer programs are grouped. To this end, both the range of computer hardware and the various areas of meteorological software application are cataloged in simplified categories. These sets of categories serve as coarse guidelines for basic comparability and classification of the computer programs.

The computer hardware has been grouped in five categories:

1. Personal Computer;
2. Advanced Graphical Workstation;
3. Large Minicomputer;
4. Large Mainframe Computer;
5. Supercomputer.

The computer programs have been grouped in ten categories. They are:

1. Process Control Programs (automatic scheduler programs to control routine production depending on external parameters such as date, time, data availability, etc.);
2. Message Switching (programs for meteorological message switching for WMO-formatted messages or AFTN messages, dis-assembly and assembly of bulletins, serving teletype and data lines);
3. Pre-processing and Data Handling (e.g., programs for decoding, code transformation binary <-> character codes, quality control, monitoring of data availability);
4. Post-processing (e.g., programs for graphical representation/ visualization of data and products on printer, plotter, hard copy, graphical VDUs);
5. Numerical Analysis Programs (e.g., programs for the generation of a gridded field of a parameter reported at irregularly distributed locations);
6. Numerical Forecast Models (e.g., hemispherical models, window models, mesoscale and sub-mesoscale models, or models for specific forecasting tasks, mainly for short-range forecasting);
7. Other Objective Techniques (e.g., statistical evaluation of model output);
8. Long-term Archiving of Numerical Results (e.g., programs for packing/unpacking and storage/retrieval of numerical fields, perhaps using GRIB as storage format);
9. Expert Systems (Computer programs that use stored information to draw conclusions about a particular case);
10. Long-term Archiving of Observations e.g., programs for packing/unpacking and storage/retrieval of observational data, perhaps using BUFR as storage format, or SQL).

CBS decided that any exchange of software or provision of related support in the framework of this project, should be arranged as bilateral (multilateral) co-operation between donor country and recipient countries. The activities carried out in the framework of the CBS Software Exchange will be monitored by the CBS Working Group on Data Management, which will develop recommendations for its improvement for consideration by CBS, as appropriate.

The degree to which members have not used software offered in the registry indicates the strength of the forces working against standardisation. In essence these forces are:

While the weather services of Members participate in the WWW, they also meet unique national requirements which call for specialised data management and processing systems;

If the resources are available there are significant long-term advantages in the preparation of software within the centre using it - specifically, program faults can be more rapidly rectified and additions/alterations to meet changing requirements can be quickly accommodated;

As noted previously, much software is hardware specific, and members are required to employ hardware which can be supported locally.

Clearly if Members offer software which is:-

Modular in nature;

Written to internationally accepted, software development standards;

Well documented and supported by training seminars to develop user skills;

Written in such a fashion as to the "vendor" independent, and;

Meets well defined, widespread needs.

Then it is expected that the take-up rate of generally available software packages will increase. The issue to be addressed by WWW DM is; how to bring about this situation.

10.3 Software Standards

The exchange of software is assisted if the computer programs have been developed to widely accepted programming standards. At the present time a set of such standards is not used widely within the meteorological community. Gibson [1] has provided an example of such a system, named the DOCTOR programming system. It is worthwhile reviewing the aims of this system as it provides an indication of the advantages of adoption of programming standards.

DOCTOR attempts to:

Provide well presented code;

Produce source code following a standard structure;

Set up points of reference for external documentation;

Enable the inclusion within the source code of documentation which can be extracted mechanically;

Allow maximum communication between routines by storing universal variables in structured pools or common blocks;

Facilitate the recognition of variable types, and the differentiation between local variables, variables in common blocks, and dummy arguments to routines;

Provide a set of utility routines for copying vectors, resetting arrays, etc.

There are two elements to the production of computer software that can be widely used outside the laboratory in which it was generated; good documentation and the adherence to accepted coding conventions.

Gibson notes that documentation is a means of retaining the ability of code to be understood. The production of documentation is a skill at least as important as the skill required to design and generate the code itself. Good documentation increases the value of code - it assists maintenance, aids understanding, and can be invaluable if language to language re-coding should ever be necessary.

With respect to coding conventions Gibson notes that code should follow a modular structure, each module or routine fulfilling a stated purpose. Modules should be divided into numbered sections and sub-sections. Communication between modules should not involve long parameter lists - shared data should be made available through shared data pools, common blocks, etc.

Despite the apparent desirability of adopting programming standards, acceptance of such standards as advocated by Gibson has not been widespread. It is difficult to specify precisely the reasons for this, but possibly the adherence to a general set of programming guidelines is felt to limit a programmers ability to optimise code for specific environments. Perhaps also programmers do not see sufficient return for the "up-front" investment of mastering, and then systematically applying a set of rules. Whatever the reason for non-adherence to standards, there is an emerging alternative in the form of Computer-Aided Software Engineering (CASE).

10.3.1 Computer-Aided Software Engineering (CASE)

The problems faced by the meteorological community in being able to prepare and maintain the software base for a rapibly expanding, worldwide computer-based infrastructure are not unique. One possible emerging solution is CASE. There are now available a number of CASE products which offer integrated tools for the planning, analysis, design and generation of computer code for major software systems. These so-called CASE tools potentially offer increased programmer productivity given that a substantial initial investment has been made in purchasing, and mastering, the tools.

It is relatively early days in the development and implementation of CASE tools. Their usage is being explored in a number of weather services, and should they become become proven, then it is to be expected that the skill base of programmers trained to use them will become broader, the tools more cost effective and powerful and as a consequence, their use more widespread.

CHAPTER 11

ENDNOTE

1. Gibson J.K., The DOCTOR System - A DOCumenTary ORiented Programming System, ECMWF Technical Memorandum No. 52

LIST OF ACRONYMS

AIREP	Colloquial name for reports from aircraft
AMDAR	Colloquial name for code form FM 42 - IX Ext.
ANSI	American National Standards Institute
ARFOR	Colloquial name for code form FM 53 - IX
ASAP	Automatic Shipboard Aerological Programme
ASDAR	Aircraft-to-Satellite Data Relay
BTAB	BUFR Tabular Format (under development)
BUFR	FM 94-IX-BUFR Binary Universal Form for the Representation of meteorological data
CAD/CAM	Computer Aided Design/Computer Aided Manufacture
CAeM	Commission for Aeronautical Meteorology
CAVOK	Ceiling And Visibility OK
CBS	Commission for Basic Systems
CBS-X	Ninth session of the Commission for Basic Systems
CCITT	Consultative Committee for International Telephony and Telegraphy
Cg-X	Tenth session of WMO Congress
Cg-XI	Eleventh session of WMO Congress
CGI	Computer Graphics Interface
CGM	Computer Graphics Metafile
CLIMAT	Colloquial name for code form FM 71 - VI
CSM	Commission for Synoptic Meteorology

DDB	Distributed Databases
DIS	Draft International Standard
DM	Data Management
DP	Draft Proposed International Standard
FTAM	File Transfer and Management
GDPS	Global Data-processing System
GKS	Graphics Kernel System
GKSM	Graphics Kernel System Metafile
GFAF	Colloquial name for code form FM 49 - IX Ext.
GOS	Global Observing System
GRID	Colloquial name for code form FM 47 - IX Ext.
GRIB	FM 92-VII Ext-GRIB (Gridded Binary) processed data in the form of grid-point values expressed in binary form
GTS	Global Telecommunication System
IAC	Colloquial name for code form FM 54 - IV
ICAO	International Civil Aviation Organization
IS	International Standard
ISO	International Standards Organisation
ITU	International Telecommunications Union
Lidar	Light induced detection and ranging system
McIDAS	Man-computer Interactive Data Access System
MAGICS	Meteorological Applications Graphics Integrated Colour System
METAR	Colloquial name for code form FM 15 - IX Ext.
MMS	Marine Meteorological Services
MTN	Main Telecommunication Network
NCAR	National Centre for Atmospheric Research
NFS	Network File System
NMC	National Meteorological Centre
NURBS	Non-Uniform Rational Beta Splines
NWP	Numerical Weather Prediction
OSF	Open Software Foundation
OSI	Open Systems Interconnection
PEX	PHIGS Extensions to X-Windows
PHIGS	Programmer's Hierarchical Interactive Graphics System
RADOF	Colloquial name for code form FM 57 - IX Ext.
RADREP	Colloquial name for code form FM 22 - IX Ext.
ROFOR	Colloquial name for code form FM 54 - IX Ext.
RSMC	Regional Specialized Meteorological Centre
SARAD	Colloquial name for code form FM 87 - VIII Ext.
SATEM	Colloquial name for code form FM 86 - VIII Ext.
SATOB	Colloquial name for code form FM 88 - VI Ext.
SFAZI	Colloquial name for code form FM 81 - I
SLTP	WMO Second Long-term Plan
Sodar	Sound detection and ranging system
SPECI	Colloquial name for code form FM 16 - IX Ext.
SYNOP	Colloquial name for code form FM 12 - IX
TAF	Colloquial name for code form FM 51 - IX
TEMP	Colloquial name for code form FM 35 - IX
TLTP	WMO Third Long-term Plan

UI	Unix International
VDI	Virtual Device Interface
VDM	Virtual Device Metafile
WINTEM	Colloquial name for code form FM 50 - VIII Ext.
WGDM	CBS Working Group on Data Management
WGDM/SGC	CBS Working Group on Data Management/Sub-group on Codes
WGDM/SGDR	CBS Working Group on Data Management/Sub-group on Data Representation
WMC	World Meteorological Centre
WMO	World Meteorological Organisation
WWW	World Weather Watch
WWWDM	World Weather Watch Data Management
X	X-Windows