|  |  |
| --- | --- |
|

|  |
| --- |
|  |

 |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **WORLD METEOROLOGICAL ORGANIZATION****COMMISSION FOR BASIC SYSTEMS** **DRAFT****Standardised Verification System (SVS) for Long-Range Forecasts (LRF)****Version  2.0 - 17 February 2000**Table of contents[1. Introduction](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#1.    Introduction)[2. Definitions](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.    Definitions)[2.1 Long-Range Forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.1    Long-Range Forecasts)[2.2 Deterministic Long-Range Forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.2    Deterministic Long-Range Forecasts)[2.3 Probabilistic Long-Range Forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.3    Probabilistic Long-Range Forecasts)[2.4 Terminology](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#2.4    Terminology)[3. SVS for Long-Range Forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.    SVS for Long-Range Forecasts)[3.1 Parameters to be verified](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.1    Parameters to be verified)[3.2 Verification areas](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.2    Verification areas)[3.3 Verification strategy](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.3    Verification strategy)[3.4 Verification scores](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.4    Verification scores)[3.4.1 RMSSS](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.4.1    RMSSS)[3.4.2 ROC](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.4.2    ROC)[3.4.2.1 Deterministic forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.4.2.1    Deterministic forecasts)[3.4.2.2 Probabilistic forecasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.4.2.2    Probabilistic forecasts)[3.5 Hindcasts](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#3.5    Hindcasts)[4. Verification data sets](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.    Verification data sets)[4.1 Data sets](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.1    Data sets)[4.2 Status of the verification data sets](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2    Status of the verification data sets)[4.2.1 ECMWF reanalysis data](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.1    ECMWF reanalysis data)[4.2.2 ECMWF operational analyses](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.2    ECMWF operational analyses)[4.2.3 NCEP reanalysis data](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.3    NCEP reanalysis data)[4.2.4 Xie-Arkin](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.4    Xie-Arkin)[4.2.5 GPCP](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.5    GPCP)[4.2.6 UKMO/CRU](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.6    UKMO/CRU)[4.2.7 UKMO/RS (HADRT)](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.7    UKMO/RS (HADRT))[4.2.8 UKMO/GMSLP](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.8    UKMO/GMSLP)[4.2.9 Reynolds OI](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.9    Reynolds OI)[4.2.10 GISST](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.10    GISST)[4.2.11 GCOS surface network (GSN)](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.11    GCOS surface network (GSN))[4.2.12 GCOS upper air network (GUAN)](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#4.2.12    GCOS upper air network (GUAN))[5. Reporting Templates](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#5.    Reporting Templates)[5.1 Template for LRF system description](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#5.1    Template for LRF system description)[5.2 Template for LRF verification exchange](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#5.2    Template for LRF verification exchange)[6. Exchange of verification scores](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#6.    Exchange of verification scores)[Annex 1](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#Annex 1)[Annex 2](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#Annex 2)[Annex 3](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#Annex 3) Standardised Verification System (SVS) for Long-Range Forecasts (LRF) **1.    Introduction**The Commission for Basic Systems (CBS) of the World Meteorological Organisation (WMO) noted that there has been considerable progress in the development of long-range forecasting activities but that no comprehensive documentation of skill levels measured according to a common standard was available. It was noted that assessments of the scientific quality of long-range forecasts were not generally made available to users, apart from simple measures of skill and warning provided along with Internet products from some issuing Centres/Institutes.Long-range forecasts are being issued from several Centres/Institutes and are being made available in the public domain. Forecasts for specific locations may differ substantially at times, due to the inherent limited skill of long-range forecast systems. The Commission acknowledged the scientific merit of those differences and encouraged the various approaches as a means to spur progress on the research front. However, concerns were raised that this situation tended to lead to confusion amongst users, and ultimately was reflecting back on the science behind long-range forecasts.There was agreement on the need to have a more coherent approach to verification of long-range forecasts. The Commission agreed that its role was to develop procedures for the exchange of verification results, with a particular focus on the practical details of producing and exchanging appropriate verification scores.This document presents the detailed specifications for the development of a Standardised Verification System (SVS) for Long-Range Forecasts (LRF) within the framework of a WMO exchange of verification scores. The SVS for LRF described herein constitutes the basis for long-range forecast evaluation and validation, and for exchange of verification scores. It will grow as more requirements are adopted. **2.    Definitions*****2.1    Long-Range Forecasts***LRF extend from thirty (30) days up to two (2) years and are defined in Table 1.**Table 1**: Definition of long-range forecasts.

|  |  |
| --- | --- |
| Monthly outlook: | Description of averaged weather parameters expressed as departures from climate values for that month. |
| Three-month or 90-day outlook: | Description of averaged weather parameters expressed as departures from climate values for that three-month or 90-day period. |
| Seasonal outlook: | Description of averaged weather parameters expressed as departures from climate values for that season. |

Seasons have been loosely defined in the Northern Hemisphere as December-January-February (DJF) for Winter (Summer in the Southern Hemisphere), March-April-May (MAM) for Spring (Fall in the Southern Hemisphere), June-July-August (JJA) for Summer (Winter in the Southern Hemisphere) and September-October-November (SON) for Fall (Spring in the Southern Hemisphere). In the Tropical areas, seasons may have different definitions. Outlooks over longer periods such as multi-seasonal outlooks or tropical rainy season outlooks may be provided.It is recognised that in some countries long-range forecasts are considered to be climate products. ***2.2    Deterministic Long-Range Forecasts***Deterministic LRF provide details of expected occurrences or non-occurrences of an event (categorical or non-categorical). Deterministic LRF can be produced from a single run of a Numerical Weather Prediction (NWP) model or a General Circulation Model (GCM), or can be produced from the grand mean of the members of an Ensemble Prediction System (EPS), or can be based on an empirical model.The forecasts are either objective numerical values such as departure from normal of a given parameter or expected occurrences (or non-occurrences) of events classified into categories (above/below normal or above/near/below normal for example). Although equi-probable categories is preferred for consistency, other classifications can be used in a similar fashion. ***2.3    Probabilistic Long-Range Forecasts***Probabilistic LRF provide probabilities of occurrences or non-occurrences of an event or a set of fully inclusive events. Probabilistic LRF can be generated from an empirical model, or produced from an Ensemble Prediction System (EPS).The events can be classified into categories (above/below normal or above/near/below normal for example). Although equi-probable categories is preferred for consistency, other classifications can be used in a similar fashion. ***2.4    Terminology***There is no universally accepted definition of forecast period and forecast lead time. However, the definition in Table 2 will be used in this document.**Table 2**: Definitions of forecast period and lead time.

|  |  |
| --- | --- |
| Forecast period: | Forecast period is the validity period of a forecast. For example, long-range forecasts may be valid for a 90-day period or a season. |
| Lead time: | Lead time refers to the period of time between the issue time of the forecast and the beginning of the forecast validity period. Long-range forecasts based on all data up to the beginning of the forecast validity period are said to be of lead zero. The period of time between the issue time and the beginning of the validity period will categorise the lead. For example, a Winter seasonal forecast issued at the end of the preceding Summer season is said to be of one season lead. A seasonal forecast issued one month before the beginning of the validity period is said to be of one month lead. |

Figure 1 presents the definitions of Table 2 in graphical format.**Figure 1:**  Definition of forecast period and lead time.Forecast range determines how far into the future LRF are provided. Forecast range is thus the summation of lead time and forecast period. **3.    SVS for Long-Range Forecasts*****3.1    Parameters to be verified***Table 3 gives the list of parameters to be verified.**Table 3**: List of parameters to be verified.

|  |  |
| --- | --- |
| 1. | Surface air temperature anomaly at screen level (T2m) |
| 2. | Precipitation anomaly |
| 3. | 500 hPa geopotential height anomaly |
| 4. | 850 hPa temperature anomaly |
| 5. | Mean Sea Level (MSL) pressure anomaly |
| 6. | Sea surface temperature (SST) anomaly |
| 7. | Southern Oscillation Index (SOI)SOI is defined as the normalised difference of the normalised averaged mean sea level pressure anomaly at Tahiti (149.6o W, 17.5o S) and the normalised averaged mean sea level pressure anomaly at Darwin, Australia (130.9o E, 12.4o S):https://www.wmo.int/pages/prog/www/DPS/Image210.gifwhere: https://www.wmo.int/pages/prog/www/DPS/Image211.gif averaged mean sea level pressure;https://www.wmo.int/pages/prog/www/DPS/Image212.gif climatological averaged mean sea level pressure;https://www.wmo.int/pages/prog/www/DPS/Image213.gif standard deviation of the averaged mean sea level pressure;https://www.wmo.int/pages/prog/www/DPS/Image214.gif standard deviation of the numerator over the verification sample. |

Both deterministic and probabilistic forecasts are verified if available. The areas where the 850 hPa temperature is below ground is masked out and not included in the overall verification. ***3.2    Verification areas***The parameters defined in section 3.1 are verified over areas defined in Table 4.**Table 4**: Verification areas for each of the parameters in Table 3.

|  |  |
| --- | --- |
| **Parameters** | **Verification areas** |
| Surface air temperature anomaly at screen level (T2m) | * Tropics: from 30o S to 30o N all inclusive
* Tropical Africa: from 10o S to 10o N and from 15o W to 45o E all inclusive
* Tropical South America: from 10o S to 10o N and from 80o W to 35o W all inclusive
* Tropical Southeast Asia: from 10o S to 10o N and from 95o E to 150o E all inclusive
* Nioo-3 region: from 150o W to 90o W and from 5o S to 5o N all inclusive
* Northern Extra-Tropics: from 30o N to 90o N, all inclusive
* whole area
* land portion
* oceanic portion
* Southern Extra-Tropics: from 30o S to 90o S, all inclusive
* whole area
* land portion
* oceanic portion
 |
| Precipitation anomaly | * Tropics: from 30o S to 30o N, all inclusive
* Tropical Africa: from 10o S to 10o N and from 15o W to 45o E all inclusive
* Tropical South America: from 10o S to 10o N and from 80o W to 35o W all inclusive
* Tropical Southeast Asia: from 10o S to 10o N and from 95o E to 150o E all inclusive
* Southern Asia: from 5o N to 30o N and from 70o E to 90o E all inclusive
* Nioo-3 region: from 150o W to 90o W and from 5o S to 5o N all inclusive
* Northern Extra-Tropics: from 30o N to 90o N, all inclusive
* whole area
* land portion
* oceanic portion
* Southern Extra-Tropics: from 30o S to 90o S, all inclusive
* whole area
* land portion
* oceanic portion
 |
| 500 hPa geopotential height anomaly | * Northern Extra-Tropics: from 30o N to 90o N, all inclusive
* Southern Extra-Tropics: from 30o S to 90o S, all inclusive
 |
| 850 hPa temperature anomaly | * Tropics: from 30o S to 30o N all inclusive
* Tropical Africa: from 10o S to 10o N and from 15o W to 45o E all inclusive
* Tropical South America: from 10o S to 10o N and from 80o W to 35o W all inclusive
* Tropical Southeast Asia: from 10o S to 10o N and from 95o E to 150o E all inclusive
* Nioo-3 region: from 150o W to 90o W and from 5o S to 5o N all inclusive
* Northern Extra-Tropics: from 30o N to 90o N, all inclusive
* whole area
* land portion
* oceanic portion
* Southern Extra-Tropics: from 30o S to 90o S, all inclusive
* whole area
* land portion
* oceanic portion
 |
| Mean Sea Level (MSL) pressure anomaly | * Northern Extra-Tropics: from 30o N to 90o N, all inclusive
* Southern Extra-Tropics: from 30o S to 90o S, all inclusive
 |
| Sea surface temperature (SST) anomaly | * Nioo-1+2 region: from 90o W to 80o W and from 10o S to 0o N all inclusive
* Nioo-3 region: from 150o W to 90o W and from 5o S to 5o N all inclusive
* Nioo-3.4 region: from 160o E to 90o W and from 5o S to 5o N all inclusive
* Nioo-4 region: from 160o E to 150o W and from 5o S to 5o N all inclusive
* Pacific warm pool: from 0o N to 4o N and from 130o E to 150o E all inclusive
* Tropical Indian Ocean: from 20o S to 20o N and from 45o E to 105o E all inclusive
* Tropical Atlantic ocean: from 20o S to 20o N and from 35o W to 15o W all inclusive
 |
| Southern Oscillation Index (SOI) | * not applicable
 |

Many LRF are produced that are applicable to limited local areas. It may not be possible to conduct verification over the areas recommended in Table 4. Appropriate verification areas should then be used with full details provided. ***3.3    Verification strategy***LRF verification should be done on a latitude/longitude grid, and at individual stations or groups of stations representing grid boxes or local areas as defined in section 3.2. Verification on a latitude/longitude grid is performed separately from the one done at stations.The verification latitude/longitude grid is recommended as being 2.5o by 2.5o , with origin at 0o N, 0o E. Both forecasts and the gridded verifying data sets are to be interpolated onto the same 2.5o by 2.5o grid.In order to handle spatial forecasts, predictions for each point within the verification grid should be treated as individual forecasts but with all results combined into the final outcome. The same approach is applied when verification is done at stations. Categorical forecasts can be treated for each category separately.Similarly, all forecasts are treated as independent and combined together into the final outcome, when verification is done over a long period of time (several years for example).Stratification of the verification data is based on forecast period, lead time and verification area. For example, seasonal forecast verification should be stratified according to season, meaning that verification results for different seasons should not be mixed. Forecasts with different lead times are similarly to be verified separately. ***3.4    Verification scores***The following verification scores are to be used: Root Mean Square Skill Score (RMSSS) and Relative Operating Characteristics (ROC). RMSSS is applicable to deterministic forecasts only, while ROC is applicable to both deterministic and probabilistic forecasts. RMSSS is applicable to non-categorical forecasts, while ROC is applicable to categorical forecasts. 3.4.1    RMSSSRMSSS is defined as:https://www.wmo.int/pages/prog/www/DPS/Image215.gifwhere: https://www.wmo.int/pages/prog/www/DPS/Image216.gif root mean square error of the forecasts.https://www.wmo.int/pages/prog/www/DPS/Image217.gif root mean square error of the standard used as forecast.Both persistence and climatology are used as standards. Persistence, for a given parameter, stands for the persisted anomaly from the forecast period immediately prior to the LRF period being verified (see Figure 2). For example, for seasonal forecasts, persistence is the seasonal anomaly from the season period prior to the season being verified. It is important to realise that only the anomaly of any given parameter can be persisted. The persisted anomaly is added to the background climatology to retrieve the persisted parameter. Climatology is equivalent to persisting a uniform anomaly of zero.**Figure 2:**Definition of persistence as applied in a forecast verification framework.  See Figure 1.RMSSS is computed at all grid points of a verification grid and/or at all stations.The root mean square error (RMS) is defined as:https://www.wmo.int/pages/prog/www/DPS/Image218.gifwhere: https://www.wmo.int/pages/prog/www/DPS/Image219.gif forecast anomaly value or value of the standard at grid point i or at station i.https://www.wmo.int/pages/prog/www/DPS/Image220.gif analysed anomaly value at grid point i or observed anomaly value at station i.https://www.wmo.int/pages/prog/www/DPS/Image221.gif for all stations, when verification is done at stations.https://www.wmo.int/pages/prog/www/DPS/Image222.gifat grid point i, when verification is done on a grid, with:https://www.wmo.int/pages/prog/www/DPS/Image223.gif the latitude at grid point i.https://www.wmo.int/pages/prog/www/DPS/Image224.gif total number of grid points or stations where verification is carried.RMSSS is given as a percentage, while all RMS scores are given in the same units as the verified LRF parameter.

|  |
| --- |
| RMSSS for deterministic forecasts with respect to persistence and climatology and RMS for the forecasts, persistence and climatology are included in the exchange of verification scores. |

 3.4.2    ROCVerification methodology using Relative Operating Characteristics (ROC), is derived from signal detection theory. This methodology is intended to provide information on the characteristics of systems upon which management decisions can be taken. In the case of weather/climate forecasts, the decision might relate to the most appropriate manner in which to use a forecast system for a given purpose. ROC is applicable to both deterministic and probabilistic categorical forecasts and is useful in contrasting characteristics of deterministic and probabilistic systems. The derivation of ROC is based on contingency tables giving the number of observed occurrences and non-occurrences of an event as a function of the forecast occurrences and non-occurrences of that event (deterministic or probabilistic). The events are defined as binary, which means that only two outcomes are possible, an occurrence or a non-occurrence.The binary event can be defined as the occurrence of one of two possible categories when the outcome of the LRF system is in two categories. When the outcome of the LRF system is in three (or more) categories, the binary event is defined in terms of occurrences of one category against the remaining ones. In those circumstances, ROC has to be calculated for each possible category. **3.4.2.1    Deterministic forecasts**Table 5 shows a general contingency table for deterministic forecasts. In Table 5, T is the grand sum of all the proper weights applied on each occurrence and non-occurrence of the events.When verification is done at stations, the weighting factor is one. Consequently, the number of occurrences and non-occurrences of the event are entered in the contingency table of Table 5.However, when verification is done on a grid, the weighting factor is cos( i), where  i is the latitude at grid point i. This approach is similar to the weighting factor used in the RMS calculation of section 3.4.1. Consequently, each number entered in the contingency table of Table 5, is, in fact, a summation of the weights properly assigned.**Table 5:**General contingency table for deterministic forecasts with definitions of the different parameters.

|  |  |
| --- | --- |
|   | **observations** |
| **forecasts** |   | **occurrences** | **non-occurrences** |   |
| **occurrences** | O1 | NO1 | O1+ NO1 |
| **non-occurrences** | O2 | NO2 | O2+ NO2 |
|   | O1+ O2 | NO1+ NO2 | T |
|   |
| where: | O1represents the correct forecasts or hits. | https://www.wmo.int/pages/prog/www/DPS/Image240.gif(OF) being 1 when the event occurrence is observed and forecast; 0 otherwise. The summation is over all grid points or stations. |
|   | O2represents the misses. | https://www.wmo.int/pages/prog/www/DPS/Image241.gif(ONF) being 1 when the event occurrence is observed but not forecast; 0 otherwise. The summation is over all grid points or stations. |
|   | NO1represents the false alarms. | https://www.wmo.int/pages/prog/www/DPS/Image242.gif(NOF) being 1 when the event occurrence is not observed but was forecast; 0 otherwise. The summation is over all grid points or stations. |
|   | NO2represents the correct rejections. | https://www.wmo.int/pages/prog/www/DPS/Image243.gif(NONF) being 1 when the event occurrence is not observed and not forecast; 0 otherwise. The summation is over all grid points or stations. |
|   | https://www.wmo.int/pages/prog/www/DPS/Image244.gif for all stations, when verification is done at stations.https://www.wmo.int/pages/prog/www/DPS/Image245.gifat grid point i, when verification is done on a grid.https://www.wmo.int/pages/prog/www/DPS/Image246.gif the latitude at grid point i. |

 Using stratification by observations (rather than by forecast), the Hit Rate (HR) is defined as (referring to Table 5):https://www.wmo.int/pages/prog/www/DPS/Image226.gifThe range of values for HR goes from 0 to 1, the latter value being desirable. An HR of one means that all occurrences of the event were correctly forecast.The False Alarm Rate (FAR) is defined as:https://www.wmo.int/pages/prog/www/DPS/Image227.gifThe range of values for FAR goes from 0 to 1, the former value being desirable. A FAR of zero means that in the verification sample, no non-occurrences of the event were forecast to occur.[Hanssen and Kuipers score](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#Hanseen and Kuipers) (1)   is calculated for deterministic forecasts. Hanssen and Kuipers score (KS) is defined as:https://www.wmo.int/pages/prog/www/DPS/Image228.gifThe range of KS goes from -1 to +1, the latter value corresponding to perfect forecasts (HR being 1 and FAR being 0). KS can be scaled so that the range of possible values goes from 0 to 1 (1 being for perfect forecasts):https://www.wmo.int/pages/prog/www/DPS/Image229.gifThe advantage of scaling KS is that it becomes comparable to the area under the ROC curve for probabilistic forecasts (see section 3.4.2.2) where a perfect forecast system has an area of one and a forecast system with no information has an area of 0.5 (HR being equal to FAR).

|  |
| --- |
| Contingency tables for deterministic categorical forecasts (such as in Table 5) are part of the exchange of LRF verification scores. The scaled Hanssen and Kuipers score for deterministic categorical forecasts is also included together with the contingency tables. One contingency table is filled in when the outcome of the LRF system is in two categories; however, one contingency table has to be filled in for each type of possible binary events, when the outcome of the LRF system is in three (or more) categories (for example, for LRF system whose forecasts are in three categories, three contingency tables are filled in, one for each category against the remaining two). When deterministic LRF are generated with an Ensemble Prediction System, the ensemble size should be specified. |

 **3.4.2.2    Probabilistic forecasts**Table 6 shows a contingency table (similar to Table 5) that can be built for probabilistic forecasts of binary events.When verification is done at stations, the weighting factor is one. Consequently, the summation of occurrences and non-occurrences of the event, stratified according to forecast probability intervals, are entered in the contingency table of Table 6.However, when verification is done on a grid, the weighting factor is cos( i), where  i is the latitude at grid point i. This approach is similar to the weighting factor used in the RMS calculation of section 3.4.1. Consequently, each number entered in the contingency table of Table 6, is, in fact, a summation of weights, properly assigned.**Table 6:**  General contingency table for probabilistic forecasts of binary events with definitions of the different parameters.

|  |  |
| --- | --- |
|   |   |
| **bin number** | **forecast probabilities** | **observed occurrences** | **observed non-occurrences** |   |
| 1 | 0-P2 (%) | O1 | NO1 |
| 2 | P2-P3 (%) | O2 | NO2 |
| 3 | P3-P4 (%) | O3 | NO3 |
| o o o | o o o | o o o | o o o |
| n | Pn-Pn+1 (%) | On | Non |
| o o o | o o o | o o o | o o o |
| N | PN-100 (%) | ON | NON |
|   |
| where: | n = number of the nth probability interval or bin n; n goes from 1 to N.Pn = lower probability limit for bin n.Pn+1 = upper probability limit for bin n.N = number of probability intervals or bins. |
| https://www.wmo.int/pages/prog/www/DPS/Image247.gif | (O) being 1 when an event corresponding to a forecast in bin n, is observed as an occurrence; 0 otherwise. The summation is over all forecasts in bin n, at all grid points or stations. |
| https://www.wmo.int/pages/prog/www/DPS/Image248.gif | (NO) being 1 when an event corresponding to a forecast in bin n, is not observed; 0 otherwise. The summation is over all forecasts in bin n, at all grid points i or stations i |
|   | https://www.wmo.int/pages/prog/www/DPS/Image249.gif for all stations, when verification is done at stations.https://www.wmo.int/pages/prog/www/DPS/Image250.gifat grid point i, when verification is done on a grid.https://www.wmo.int/pages/prog/www/DPS/Image251.gif the latitude at grid point i. |

 To build the contingency table in Table 6, probability forecasts of the binary event are grouped in categories or bins in ascending order, from 1 to N, with probabilities in bin n-1 lower than those in bin n (n goes from 1 to N). The lower probability limit for bin n is Pn-1 and the upper limit is Pn. The lower probability limit for bin 1 is 0%, while the upper limit in bin N is 100%. The summation of the weights on the observed occurrences and non-occurrences of the event corresponding to each forecast in a given probability interval (bin n for example) is entered in the contingency table.Hit rate and false alarm rate are calculated for each probability threshold Pn (see Table 6). The hit rate for probability threshold Pn (HRn) is defined as (referring to Table 6):https://www.wmo.int/pages/prog/www/DPS/Image231.gif and the false alarm rate (FARn) is defined as:https://www.wmo.int/pages/prog/www/DPS/Image232.gifwhere n goes from 1 to N. The range of values for HRn goes from 0 to 1, the latter value being desirable. The range of values for FARn goes from 0 to 1, zero being desirable. Frequent practice is for probability intervals of 10% (10 bins, or N=10) to be used. However the number of bins (N) should be consistent with the number of members in the ensemble prediction system (EPS) used to calculate the forecast probabilities. For example, intervals of 33% for a nine-member ensemble system could be more appropriate.Hit rate (HR) and false alarm rate (FAR) are calculated for each probability threshold Pn, giving N points on a graph of HR (vertical axis) against FAR (horizontal axis) to form the Relative Operating Characteristics (ROC) curve. This curve, by definition, must pass through the points (0,0) and (1,1) (for events being predicted only with 100% probabilities and for all probabilities exceeding 0% respectively). The further the curve lies towards the upper left-hand corner (where HR=1 and FAR=0) the better; no-skill forecasts are indicated by a diagonal line (where HR=FAR).The area under the ROC curve is a commonly used summary statistics representing the skill of the forecast system. The area is standardised against the total area of the figure such that a perfect forecast system has an area of one and a curve lying along the diagonal (no information) has an area of 0.5. The normalised ROC area has become known as the ROC score. Not only can the areas be used to contrast different curves, but they are also a basis for Monte Carlo significance tests. It is proposed that Monte Carlo testing should be done within the forecast data set itself. The area under the ROC curve can be calculated using the Trapezium rule. Although simple to apply, the Trapezium rule renders the ROC score dependent on the number of points on the ROC curve, and care should be taken in interpreting the results. Other techniques are available to calculate the [ROC score](https://www.wmo.int/pages/prog/www/DPS/SVS-for-LRF.html#ROC score) (2).

|  |
| --- |
| Contingency tables for probabilistic forecasts (such as in Table 6) are part of the exchange of LRF verification scores. The ROC score (area under the ROC curve, normalised to one) for probabilistic forecasts is also included together with the contingency tables. One contingency table is filled in when the outcome of the LRF system is in two categories; however, one contingency table has to be filled in for each type of possible binary events, when the outcome of the LRF system is in three (or more) categories (for example, for LRF system whose forecasts are in three categories, three contingency tables are filled in, one for each category against the remaining two). When LRF are generated with an Ensemble Prediction System, the ensemble size should be specified. |

 ***3.5    Hindcasts***In contrast to short- and medium-range dynamical Numerical Weather Prediction (NWP) forecasts, LRF are produced relatively few times a year (for example, one forecast for each season or one forecast for the following 90-day period, issued every month). Therefore the verification sampling for LRF may be limited, possibly to the point where the validity and significance of the verification results may be questionable. Providing verification for a few seasons, or even over a few years only may be misleading and may not give a fair assessment of the skill of any LRF system. LRF systems should be verified over as long a period as possible in hindcast mode. Although there are limitations on the availability of verification data sets and in spite of the fact that validating numerical forecast systems in hindcast mode requires large computer resources, the hindcast period should be as long as possible, at least 30 years representing the desirable immediate objective. Model validation in hindcast mode is one of the most important aspect of any LRF system.Verification in hindcast mode should be achieved in a form as close as possible to the real time operating mode in terms of resolution, ensemble size and parameters. In particular dynamical models must not make any use of future data. Validation of empirical models should be done in a cross-validation framework with models trained on the original data set after removing a few years including and following the year at which the models will be verified (ideally excluding a total of five years), and the procedure repeated every year over the entire hindcast period. The same restriction should apply to bias correction used by some dynamical models.

|  |
| --- |
| Verification results over the hindcast period are part of the exchange of LRF verification scores. |

 **4.    Verification data sets**The same data should be used to generate both climatology and verification data sets, although the forecasts issuing Centres/Institutes own analyses or ECMWF reanalyses and subsequent operational analyses may be used when other data are not available. Use of NCEP reanalysis data is also another option.Many LRF are produced that are applicable to limited or local areas. It may not be possible to use the data in either the recommended climatology or verification data sets for validation or verification purposes in these cases. Appropriate data sets should then be used with full details provided. ***4.1    Data sets***Table 7 gives the list of verification data sets that should be used as appropriate.**Table 7**: Verification data sets that should be used.

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Gridded verification data sets** | **Observation data sets** |
| 1.   Surface air temperature anomaly at screen level (T2m) | * ECMWF reanalysis
* ECMWF operational analysis
* NCEP reanalysis
* Centre/Institute own operational analysis
* UKMO/CRU
 | * GCOS surface network (GSN)
* local network
 |
| 2.  Precipitation anomaly | * Xie-Arkin
* GPCP
* ECMWF reanalysis
* NCEP reanalysis
* Centre/Institute own operational analysis
 | * GCOS surface network (GSN)
* local network
 |
| 3.  500 hPa geopotential height anomaly | * ECMWF reanalysis
* ECMWF operational analysis
* NCEP reanalysis
* Centre/Institute own operational analysis
 | * GCOS upper air network (GUAN)
 |
| 4.  850 hPa temperature anomaly | * ECMWF reanalysis
* ECMWF operational analysis
* NCEP reanalysis
* Centre/Institute own operational analysis
* UKMO/RS
 | * GCOS upper air network (GUAN)
 |
| 5.  Mean Sea Level (MSL) pressure anomaly | * ECMWF reanalysis
* ECMWF operational analysis
* NECEP reanalysis
* Centre/Institute own operational analysis
* UKMO/GMSLP
 | * GCOS surface network (GSN)
 |
| 6.  Sea surface temperature (SST) anomaly | * Reynolds OI with option for additional use of GISST
 | * not applicable
 |
| 7.  Southern Oscillation Index (SOI) | * not applicable
 | * Tahiti and Darwin observations
 |

 ***4.2    Status of the verification data sets***The following paragraphs give the status of the various proposed verification data sets, as of January 2000:4.2.1    ECMWF reanalysis data

|  |  |
| --- | --- |
| Availability: | * Available from ECMWF.
 |
| Period: | * 1979-1993.
 |
| Type: | * Monthly analyses of various meteorological fields.
 |
| Grid: | * 2.5o by 2.5o
 |
| Update frequency: | * None.
 |
| Climatology: | * None.
 |
| Reference: | * http://www.ecmwf.int/research/era
* http://www.badc.rl.ac.uk/data/ecmwf-grid/
 |
| Web site: | * http://www.badc.rl.ac.uk/data/ecmwf-era/
 |

 4.2.2    ECMWF operational analyses

|  |  |
| --- | --- |
| Availability: | * ECMWF
* Not available without an account.
 |
| Period: | * 1985 to present.
 |
| Type: | * Twice daily analyses of various meteorological fields (00 and 12 UTC).
 |
| Grid: | * 2.5o by 2.5o
 |
| Update frequency: | * Daily.
 |
| Climatology: | * None.
 |
| Reference: | * http://www.badc.rl.ac.uk/data/ecmwf-grid/
 |
| Web site: | * http://www.badc.rl.ac.uk/data/ecmwf-op/
 |

 4.2.3    NCEP reanalysis data

|  |  |
| --- | --- |
| Availability: | * Available from NCEP.
 |
| Period: | * 1957-1999.
 |
| Type: | * Monthly and daily analyses of various meteorological fields.
 |
| Grid: | * 2.5o by 2.5o
 |
| Update frequency: | * Monthly
 |
| Climatology: | * 1958-1997
 |
| Reference: | Kalnay E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds (NCEP Environmental Modeling Center), M. Chelliah, W. Ebisuzaki, W.Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang (NCEP Climate Prediction Center) Roy Jenne, Dennis Joseph (NCAR), 1996: The NCEP/NCAR 40-Year Reanalysis Project, Bull. American Met. Soc. (BAMS), |
| Web site: | * http://wesley.wwb.noaa.gov/reanalysis.html
* http://wesley.wwb.noaa.gov/data.html
 |

 4.2.4    Xie-Arkin

|  |  |
| --- | --- |
| Availability: | * NOAA
 |
| Period: | * 1979-1998.
 |
| Type: | * Rain gauges, satellites and model precipitation amount values.
* Choice of grids with missing values in the polar regions or completed with model data.
* Monthly means.
 |
| Grid: | * 2.5o by 2.5o
 |
| Update frequency: | * Every 3 to 6 months.
 |
| Climatology: | * None.
 |
| Reference: | * Xie, Pingping, Phillip A. Arkin, 1997: Global Precipitation: A 17-Year Monthly Analysis Based on Gauge Observations, Satellite Estimates, and Numerical Model Outputs. Bulletin of the American Meteorological Society: Vol. 78, No. 11, 2539–2558.
 |
| Web site: | * http://www.cdc.noaa.gov/cdc/data.cmap.html
 |

4.2.5    GPCP

|  |  |
| --- | --- |
| Availability: | * NASA
 |
| Period: | * 1987-1999.
 |
| Type: | * Similar to Xie-Arkin data.
 |
| Grid: | * 2.5o by 2.5o
 |
| Update frequency: | * Unknown.
 |
| Climatology: | * None.
 |
| Reference: | * Huffman, George J., Robert F. Adler, Philip Arkin, Alfred Chang, Ralph Ferraro, Arnold Gruber, John Janowiak, Alan McNab, Bruno Rudolf, Udo Schneider, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. Bulletin of the American Meteorological Society: Vol. 78, No. 1, 5–20.
 |
| Web site: | * http://daac.gsfc.nasa.gov/CAMPAIGN\_DOCS/FTP\_SITE/INT\_DIS/readmes/gpcp\_global\_precip.html
 |

4.2.6    UKMO/CRU

|  |  |
| --- | --- |
| Availability: | * UKMO/Hadley Centre
 |
| Period: | * 1851-1998.
 |
| Type: | * Monthly surface air temperature (T2m) anomalies from 1961-1990 climate.
 |
| Grid: | * 5o by 5o
 |
| Update frequency: | * Monthly.
 |
| Climatology | * 1961-1990.
 |
| Reference: | * Jones, P. D., M. New, D. E. Parker, S. Martin and I. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. Rev. Geophys., 37, 173-199.
 |
| Web site: | * http://www.cru.uea.ac.uk/cru/data/temperat.htm
 |

4.2.7    UKMO/RS (HADRT)

|  |  |
| --- | --- |
| Availability: | * UKMO/Hadley Centre
 |
| Period: | * 1958 to present.
 |
| Type: | * Monthly and seasonal means. See text below for the type of parameters.
 |
| Grid: | * various grids (see below)
 |
| Update frequency: | * Monthly.
 |
| Climatology: | * 1971-1990.
 |
| Reference: | * Parker, D.E., Gordon, M., Brown, S.J., and O'Donnell, M. 1998: The New Monthly Gridded Global Upper-Air Temperature Data Sets (HADRT2. Hadley Centre Internal note no. 84
* Parker, D.E., Gordon, M., Cullum, D.P.N, Sexton, D.M.H, Folland, C.K., and Rayner, N. 1997: A New Gridded Radiosonde Temperature Data Base and Recent Temperature Trends. Geophys. Res. Letters, 24, 1499-1502
 |
| Web site: | * http://www.badc.rl.ac.uk/data/hadrt/
 |

The HADRT data sets consist of monthly or seasonal temperature anomalies from the 1971-1990 climate normal on a global grid, computed from radiosonde station data from 1958 to present. Anomalies are available for 9 standard levels as well as tropospheric (850 - 300hPa) and stratospheric (150 - 30hPa) averages. In some versions bias corrections linked to instrumental or operational discontinuities have been applied to data. The current versions are as follows:HADRT2.0Contains monthly data from 1958 - present, on a 5 degree latitude by 10 degree longitude grid. No bias corrections are applied to the station data. Anomalies are with respect to 1971 - 1990 and available for the following standard levels, 850, 700, 500, 300, 200, 150, 100, 50, 30 hPa.HADRT2.1As HADRT2.0 but with bias corrections made to many station time series worldwide. The adjustments were calculated by reference to MSU2R version 'c' in the troposphere (850 - 300hPa), and MSU4 in the stratosphere (150 - 30hPa), but only for known changes in instrumental or operational procedures for the period after 1979. Available for all HADRT2.0 levels except 30hPa where data were too sparse. (70% data availability required for reconverting anomalies after MSU comparisons)HADRT2.1sThis is a combination of the above data sets, made to remove the influence of MSU2R in the troposphere. HADRT2.0 is used up to and including 200hPa, and HADRT2.1 is used above 200hPa.HADRT2.2This is an eigenvector reconstructed grid data set from 1958 - present, on a 10 degree latitude by 20 degree longitude grid, created from HADRT2.1. Values are stored as seasonal or annual anomalies for all levels except 30hPa. The eigenvector reconstruction was used to fill in missing seasons or years in boxes with 70% of seasonal or annual data available Parker et. al(1997).HADRT2.2uThis is an eigenvector reconstructed grid data set as above, but created from HADRT2.0.HADRT2.3This is a globally complete data set based on HADRT2.1 but with gaps filled in by reference to the second derivative of the corresponding NCEP reanalysis temperature fields, Parker et. al(1998), using the Laplacian technique of Reynolds(1988)HADRT2.3sAs above but HADRT2.1s is used as the base data set.These data sets are available for use in scientific research upon the signing of a short license agreement.HADRT2.3 and HADRT2.3s data sets are recommended. 4.2.8    UKMO/GMSLP

|  |  |
| --- | --- |
| Availability: | * UKMO/Hadley Centre
 |
| Period: | * 1949-1994.
 |
| Type: | * Monthly mean sea level pressure averages.
 |
| Grid: | * 5o by 5o
 |
| Update frequency: | * None.
 |
| Climatology: | * None.
 |
| Reference: | * Basnett and Parker (1997) UKMO/Hadley Centre Climate Research note 79.
 |
| Web site: | * http://www.meto.govt.uk/sec5/CR\_div/climate\_index/hadley\_gmslp.html
 |

4.2.9    Reynolds OI

|  |  |
| --- | --- |
| Availability: | * NOAA/CDC
 |
| Period: | * 1981-1998.
 |
| Type: | * Weekly or monthly sea surface temperature (SST) means.
 |
| Grid: | * 1o by 1o
* 2o by 2o
 |
| Update frequency: | * 2-4 times a year.
 |
| Climatology: | * None.
 |
| Reference: | * Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. J. Climate, 7, 929-948.
 |
| Web site: | * http://www.cdc.noaa.gov/cdc/data.reynolds\_sst.html
 |

The SST data products are derived from ship, satellite, and sea ice limit data. There are two main categories of data:1. OI weekly and monthly composite analyses, November 1981 to present: analyses that combine ship observations satellite data and realistic sea-ice on a 1o by 1o weekly and monthly grids;
2. reconstructed historical monthly analyses, from 1950 to 1992 using EOF interpolation as basis functions, to create a 2o by 2o monthly grid. Analyses are limited from 69o N to 25o S. OI climatology is used to fill the regions outside the analyses range.

 4.2.10    GISST

|  |  |
| --- | --- |
| Availability: | * UKMO/Hadley Centre and British Atmospheric Data Centre (BADC).
 |
| Period: | * 1900-1994 (version 2.2) and 1871-1999 (version 2.3b).
 |
| Type: | * Global sea surface temperature (SST) fields were created using a variety of techniques including EOF reconstruction. These data have been shown to contain good ENSO variability back to the late nineteenth century and benefit from a consistent variation of SST with sea-ice concentration within the marginal ice zone.
* Monthly means.
 |
| Grid: | * 1o by 1o
 |
| Update frequency: | * Monthly (for clients of British Atmospheric Data Centre, BADC).
 |
| Climatology: | * 1961-1990.
 |
| Reference: | * Parker et al. (1995) Hadley Centre/UKMO, Climate Research note no.63.
 |
| Web site: | * WEB: http://www.meto.govt.uk/sec5/CR\_div/climate\_index/hadley\_gisst.html
 |

4.2.11    GCOS surface network (GSN)The climatology for the GSN stations is not available. Information on GCOS data can be found at:http://193.135.216.2/web/gcos/gcoshome.htmlThe UKMO/CRU and Global Historical Climatology Network (GHCN) data sets include approximately 96% of the GSN stations. GHCN monthly surface air temperature (T2m) averages are available at stations or on a 5o by 5o grid (similar to UKMO/CRU data set). The data set covers the period from 1851 to 1995. The data set is available at: ftp://www.ncdc.noaa.gov/pub/data/ghcn/v2/ghcnftp.htmlInformation on GHCN data set is available at:http://www.ncdc.noaa.gov/ol/climate/research/ghcn/ghcnoverview.htmlThe GHCN and UKMO/CRU data sets could be an alternative to GCOS/GSN. 4.2.12    GCOS upper air network (GUAN)The climatology for the GUAN stations is not available. Information on the GCOS/GUAN data can be found at: http://193.135.216.2/web/gcos/guan.html **5.    Reporting Templates**Two types of templates are to be filled in. The first one is related to the description of the LRF system. The second type of template is for the exchange of verification results. There is a different template for each of deterministic and probabilistic forecast verifications. ***5.1    Template for LRF system description***A description of the template for a LRF system is given in Annex 1, together with instructions on how to enter the data.

|  |
| --- |
| The template for LRF system description should be updated when ever there are changes in the LRF system. |

 ***5.2    Template for LRF verification exchange***Two templates for exchange of LRF verification results are presented in Annex 2 and Annex 3: one for deterministic forecasts (Annex 2) and one for probabilistic forecasts (Annex 3). These templates should be filled in as appropriate once a year.

|  |
| --- |
| The templates for deterministic forecasts and/or for probabilistic forecasts should be filled in as appropriate for exchange of verification scores. These templates should be filled in separately for all combinations of verified parameters, forecast periods, forecast lead times, and verification areas. Verification results over the hindcast period need to be provided once, and should be updated when ever a new LRF system is implemented. Verification results for current or recent forecasts should be provided once a year (generally at the beginning of the following year). |

**6.    Exchange of verification scores**HTML version of the templates in Annexes 1 to 3 will be posted on a central Web site. Each participating Organisation in the WMO exchange of LRF verification scores is urged to obtain copies of the templates and fill them in as appropriate and send them back to be posted on the central Web site.The template in Annex 1 (LRF system description) needs to be updated as and when required. The templates in Annex 2 and Annex 3 are posted once a year or when a LRF system undergoes an upgrade. The verification results pertaining to hindcasts, should be updated as required.The address of the central Web site will be provided at a later stage. Notes:(1)   See:  Hanseen A.J. and W.J. Kuipers, 1965:  On the relationship between the frequency of rain and various meteorological parameters.   Koninklijk Nederlands Meteorologist Institua Meded. Verhand, 81-2-15.See also:  Stanski H.R., L.J. Wilson and W.R. Burrows, 1989: Survey of common verification methods in meteorology.  World Weather Watch Technical Report No. 8, WMO/TD 358, 114pp.(2)  See for example: Mason I., 1987:  A model for assessment of weather forecast.  Australian Met. Magazine, 30, 291-303.**Annex 1**Template used for LRF system description:

|  |
| --- |
| **Identification** |
| **Country:** | **1** |
| **Meteorological Centre:** | **2** |
| **LRF system identification:** | **3** |
|   |
| **Description of Long-Range Forecast (LRF) System** |
| **Status of LRF system:** | **information:** | **4** |
| **dissemination:** | **5** |
| **guidance:** | **6** |
| **Type of LRF system:** | **numerical:** | **7** | **empirical:** | **8** |
| **hybrid:** | **9** | **coupled:** | **10** |
| **statistics:** | **11** |
| **Type of forecasts:** | **deterministic:** | **12** | **probabilistic:** | **13** |
| **LRF output products:** | **parameter:** | **14** |
| **categories:** | **15** |
| **frequency:** | **16** |
| **forecast period:** | **17** |
| **lead time:** | **18** |
| **forecast area:** | **19** |
| **Model description:** | **20** |
| **Model resolution:** | **horizontal:** | **21** | **vertical:** | **22** |
| **Bias correction:** | **23** |
| **Ensemble forecasting:** | **ensemble size:** | **24** |
|   | **initialisation:** | **25** |
| **SST specification:** | **26** |
| **Sea-Ice specification:** | **27** |
| **Snow specification:** | **28** |
| **Soil temperature:** | **29** |
| **Soil moisture:** | **30** |
| **Hindcast evaluation:** | **hindcast period:** | **31** |
| **cross-validation:** | **32** |
| **particularities:** | **33** |

Although the template above shows only one line per entry, it is possible to enter new lines to have enough space to fill in the required information properly.1. Enter the name of the Country of the Meteorological Centre or Institute responsible for the LRF system.
2. Enter the name of the Meteorological Centre or Institute responsible for the LRF system.
3. Enter an identification name for the LRF system described in this template. If there are more than one LRF system, a LRF system description template should be filled in for each one of them.
4. Enter "Yes" if information based on the LRF system is made accessible to users.
5. If "Yes" is entered in box 3, describe how LRF information is made available to users or how users can access LRF information.
6. If "Yes" is entered in box 3, describe the interpretation guidance material that is provided to users, if any.
7. Enter "Yes" if the LRF system is based on numerical models, either NWP or GCM or both.
8. Enter "Yes" if the LRF system is based on statistical or empirical models.
9. Enter "Yes" if the LRF system is based on a blend of dynamical and empirical models.
10. Enter "Yes" if the LRF system is based on an atmospheric model (either dynamical or empirical) coupled with an oceanic model (either dynamical or empirical).
11. Describe, if any, the statistical adaptation methods applied to model outputs. For example, statistical adaptation system could be based on Model Output Statistics (MOS) or on Perfect Prog (PP).
12. Enter "Yes" if the forecasts are deterministic.
13. Enter "Yes" if the forecasts are probabilistic.
14. Boxes 13 to 18 inclusive should be used to describe the list of LRF output products and should be repeated for each one of them. Box 13 should be used to enter one output parameter.
15. If forecast parameter in box 13 is categorised, such as "below normal", "normal" or "above normal", indicate the definition of the categories. If no categorisation is applied, enter "objective" or leave box 14 empty.
16. Indicate the frequency of issue of the output parameter described in box 13. For example, seasonal forecasts may be issued every month or every season.
17. Indicate the valid period of the forecasts of the output parameter described in box 13, according to the definition in section 2.4.
18. Indicate the forecast lead time of the forecasts of the output parameter described in box 13, according to the definition in section 2.5.
19. Indicate the areas over which the forecasts of the output parameter described in box 13 are valid. For example, forecasts are global or hemispheric, or valid over a particular country. Boxes 13 to 18 inclusive should be repeated for each LRF output parameter.
20. Provide a short narrative description of the dynamical models used in the LRF system if applicable. If the LRF system include an empirical model, provide a list of predictors used.
21. Indicate the horizontal resolution of the dynamical models used in the LRF system if applicable.
22. Indicate the vertical resolution of the dynamical models used in the LRF system if applicable.
23. Describe the bias correction applied, if any.
24. If the LRF system is based on ensemble predictions, indicate the size of the ensemble.
25. If the LRF system is based on ensemble predictions, describe the initialisation method to generate the different members of the ensemble.
26. Indicate how the sea surface temperature (SST) is prescribed at initial condition and how it is prescribed throughout the integration of the model.
27. Indicate how sea-ice is prescribed at initial condition and how it is prescribed throughout the integration of the model.
28. Indicate how snow is prescribed at initial condition and how it is prescribed throughout the integration of the model.
29. Indicate how soil temperature is prescribed at initial condition and how it is prescribed throughout the integration of the model.
30. Indicate how soil moisture is prescribed at initial condition and how it is prescribed throughout the integration of the model.
31. If the LRF system has been evaluated in hindcast mode, indicate the length of the hindcast period. If there is an entry in box 30, reporting template 2 and/or 3 should also be filled in.
32. If the LRF system has been cross-validated in hindcast mode, indicate the number of years that have been removed from the data set.
33. If the LRF system has been evaluated in hindcast mode, indicate the differences that may exist between the hindcast version and the real-time one.

 **Annex 2**Template used for exchange of verification scores for deterministic forecasts:

|  |
| --- |
| **Identification** |
| **Country:** | **1** |
| **Meteorological Centre:** | **2** |
| **LRF system identification:** | **3** |
|   |
| **Long-Range Forecast (LRF) verification results - deterministic forecasts** |
| **Verified parameter:** | **4** |
| **Forecast period:** | **5** |
| **Forecast lead time:** | **6** |
| **Verification area:** | **7** |
| **Verification period:** | **8** |
| **Verification data set:** | **9** |
| **Climatology data set:** | **10** |
| **Persistence:** | **11** |
|   |
| **RMS-forecast** | **RMS-persistence** | **RMS-climatology** | **RMSSS-persistence** | **RMSSS-climatology** |
| **12** | **13** | **14** | **15** | **16** |
|   |
| **Binary event:** | **17** |
| **Contingency Table:** |   | **observations** |
| **forecasts** | **occurrences** | **non-occurrences** |
| **occurrences** | **18** | **19** |
| **non-occurrences** | **20** | **21** |
| **Kuipers Score:** | **22** |   |   |

Although the template above shows only one line per entry, it is possible to enter new lines to have enough space to fill in the required information properly.1. Enter the name of the Country of the Meteorological Centre responsible for the LRF system.
2. Enter the name of the Meteorological Centre responsible for the LRF system.
3. Enter an identification name for the LRF system upon which the verified parameter in this template is based. Refer to box 3 in annex 1.
4. Indicate the meteorological parameter, with units, for which verification results are entered in this template. Each verified parameter requires it own reporting template.
5. Indicate the valid period of the forecasts, according to the definition in section 2.4. For example, monthly forecast or seasonal forecasts.
6. Indicate the forecast lead time, according to the definition in section 2.5. For example, seasonal forecasts issued for the next year.
7. Indicate over which area the verification is performed. A list of possible verification areas is provided in section 3.2.
8. Indicate the period over which the verification has been done. For example, this period may be an entire hindcast period, or a season in a particular year.
9. Indicate the verification data set used. A list of possible verification data sets is provided in section 3.5.
10. Indicate the climatology data set used.
11. Describe what has been used for persistence if applicable. If persistence is not used, indicate the reasons why.
12. Enter the Root Mean Square (RMS) error of the forecast. A definition of RMS is given in section 3.4.1.
13. Enter the Root Mean Square (RMS) error of persistence. A definition of RMS is given in section 3.4.1.
14. Enter the Root Mean Square (RMS) error of climatology. A definition of RMS is given in section 3.4.1.
15. Enter the Root Mean Square error Skill Score (RMSSS), based on persistence as a standard. A definition of RMSSS is given in section 3.4.1.
16. Enter the Root Mean Square error Skill Score (RMSSS), based on climatology as a standard. A definition of RMSSS is given in section 3.4.1.
17. Give the definition of the binary event used in the contingency table.
18. Enter the number of hits (cases of observed occurrences of the binary event of the verified parameter that were forecast as occurrences).
19. Enter the number of false alarms (cases of observed non-occurrences of the binary event of the verified parameter that were forecast as occurrences).
20. Enter the number of misses (cases of observed occurrences of the binary event of the verified parameter that were forecast as non-occurrences).
21. Enter the number of correct rejections (cases of observed non-occurrences of the binary event of the verified parameter that were forecast as non-occurrences).
22. Enter the value of the scaled Hanssen and Kuipers score. A definition of the scaled Hanssen and Kuipers score is provided in section 3.4.2.1.

 **Annex 3**Template used for exchange of verification scores for probabilistic forecasts:

|  |
| --- |
| **Identification** |
| **Country:** | **1** |
| **Meteorological Centre:** | **2** |
| **LRF system identification:** | **3** |
|   |
| **Long-Range Forecast (LRF) verification results - probabilistic forecasts** |
| **Verified parameter:** | **4** |
| **Forecast period:** | **5** |
| **Forecast lead time:** | **6** |
| **Verification area:** | **7** |
| **Verification period:** | **8** |
| **Verification data set:** | **9** |
| **Climatology data set:** | **10** |
|   |
| **Binary event:** | **11** |
| **Contingency Table:** |   | **observations** |
| **probability intervals** | **occurrences** | **non-occurrences** |
| **12** | **13** | **14** |
| **15** | **16** | **17** |
| **18** | **19** | **20** |
| **21** | **22** | **23** |
| **24** | **25** | **26** |
| **27** | **28** | **29** |
| **30** | **31** | **32** |
| **33** | **34** | **35** |
| **36** | **37** | **38** |
| **39** | **40** | **41** |
| **ROC score:** | **42** |   |

Although the template above shows only one line per entry, it is possible to enter new lines to have enough space to fill in the required information properly.1. Enter the name of the Country of the Meteorological Centre responsible for the LRF system.
2. Enter the name of the Meteorological Centre responsible for the LRF system.
3. Enter an identification name for the LRF system upon which the verified parameter in this template is based. Refer to box 3 in annex 1.
4. Indicate the meteorological parameter, with units, for which verification results are entered in this template. Each verified parameter requires its own reporting template.
5. Indicate the valid period of the forecasts, according to the definition in section 2.4. For example, monthly forecast or seasonal forecasts.
6. Indicate the forecast lead time, according to the definition in section 2.5. For example, seasonal forecasts issued for the next year.
7. Indicate over which area the verification is performed. A list of possible verification areas is provided in section 3.2.
8. Indicate the period over which the verification has been done. For example, this period may be en entire hindcast period, or a season in a particular year.
9. Indicate the verification data set used. A list of possible verification data sets is provided in section 3.5.
10. Indicate the climatology data set used.
11. Give the definition of the binary event used in the contingency table.
12. Enter the lower and upper limit of the first probability interval. See note below.
13. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 14.
14. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 14.
15. Enter the lower and upper limit of the second probability interval.
16. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 17.
17. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 17.
18. Enter the lower and upper limit of the third probability interval.
19. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 20.
20. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 20.
21. Enter the lower and upper limit of the fourth the probability interval.
22. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 23.
23. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 23.
24. Enter the lower and upper limit of the fifth the probability interval.
25. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 26.
26. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 26.
27. Enter the lower and upper limit of the sixth the probability interval.
28. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 29.
29. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 29.
30. Enter the lower and upper limit of the seventh the probability interval.
31. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 32.
32. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 32.
33. Enter the lower and upper limit of the eighth probability interval.
34. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 35.
35. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 35.
36. Enter the lower and upper limit of the ninth probability interval.
37. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 38.
38. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 38.
39. Enter the lower and upper limit of the tenth probability interval.
40. Enter the number of observed occurrences of the binary event of the verified parameter corresponding to the probability interval in box 41.
41. Enter the number of observed non-occurrences of the binary event of the verified parameter corresponding to the probability interval in box 41.
42. Enter the area under the ROC curve (the area being normalised to one). Explanation is provided in section 3.4.2.2.

**Note**: Boxes 12 to 42 are filled in according to the number of probability intervals. See section 3.4.2.2 for more details. The above template provides space for a maximum of ten probability intervals. In order to have significant ROC statistics for probability forecasts, there must be a minimum of two probability intervals. If the number of probability intervals is less than 10, unused boxes in the template are left blank. The lower limit of the first probability interval must be 0%, while the upper limit of the last probability interval must be 100%.  |

|  |  |
| --- | --- |
|

|  |
| --- |
| **[WMO Front](http://www.wmo.ch/%22%20%5Ct%20%22_parent), [About WMO](http://www.wmo.ch/web-en/wmofact.html%22%20%5Ct%20%22_self), [WWW Front](http://www.wmo.ch/web/www/www.html%22%20%5Ct%20%22_self), [Library](http://www.wmo.ch/web/arep/lib1/homepage.html%22%20%5Ct%20%22_self), [International Weather](http://www.wmo.ch/web-en/member.html%22%20%5Ct%20%22_self)** |

  |