Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 569-573.

ASSESSING UNCERTAINTY OF SEASONAL PROBABILISTIC FORECASTS: DISTRIBUTION-FREE CONFIDENCE LIMITS

A. H. N. Maia [1] and H. Meinke [2]

[1] Embrapa Meio Ambiente, PO Box 69, Jaguariúna, SP, Brazil; ahmaia@cnpma.embrapa.br

[2] Department of Primary Industries and Fisheries, PO Box 102, Toowoomba, Qld 4350, Australia, holger.meinke@dpi.qld.gov.au

Introduction

Probabilistic climate information, including climate forecasts, often rely on time series data of prognostic variables (Y, eg. rainfall or yield), represented as cumulative distribution probabilities functions (CDFs) or their complement, probability of exceeding functions (POEs). Useful information for decision-making is then derived from such distributions and expressed as Y percentiles or the probability of Y exceeding a certain threshold *c* (*Pr[Y> c*]). Such estimates are frequently reported without any measure of uncertainty. The degree of uncertainty associated with such estimates depends on the length of the time series and their internal variability. Lack of uncertainty assessments can lead to misguided beliefs about the true performance of the forecast systems, possibly resulting in inappropriate actions by the decision maker (Potts et al. 1996; Jolliffe 2004; Maia et al. 2006).

However, even when uncertainty estimates are provided, these are often based on methods that rely on assumptions of data being normally distributed. This is in spite of the well-known fact that distributions of important climate variables, such as rainfall, are notoriously skewed, particularly in areas with strong seasonality (eg. high frequencies of 'zero' rainfall amounts). As an alternative for Normal-based procedures, we therefore propose the use of distribution free methods for constructing percentile and POE confidence limits as described in Hahn and Meeker (1991) and implemented into "The Capability Procedure'' of the SAS® System. Such distribution-free tools are particularly useful for spatial uncertainty assessments that would otherwise require a tedious, location-by-location checking of assumptions regarding underlying probability distributions (Maia et al., 2006).

Here, we discuss the rationale, advantages and limitations of both, parametric and non-parametric approaches. We illustrate the use of distribution-free methods by assessing the uncertainty of percentiles and POEs estimates for 3-monthly rainfall series from selected locations in Australia and the Southeast of South America.

Data & Methods

To demonstrate the application of distribution-free methods we quantified uncertainties of 3-monthly rainfall percentile and POE estimates from historical climate records at selected case study sites (64 stations from Australia and 60 stations from South America, from which we selected 4 and 2 stations for an in-depth analysis). The aim of this study was to demonstrate the utility of distribution-free methods across a wide range of location-dependent distributional properties (eg. different means and medians as well as various degrees of internal variability). The case studies were specifically chosen to represent these differences.

Here we only assessed uncertainties of 'unconditional climatologies', ie. uncertainties associated with POEs derived from all available, historical rainfall records at given locations, regardless of climatological conditions such as ENSO. For statistical, ENSO-based forecasts, such unconditional climatologies can be partitioned further into conditional climatologies based on either SOI or ENSO classes. Such 'analogue' distributions serve as probabilistic forecasts that demonstrate past ENSO impacts when similar conditions to the current situation prevailed. This is a convenient and robust way to generate a forecast distribution (Stone, 1996; Podestá et al., 2002). An extension of the methods proposed here will eventually include the ability to provide uncertainty assessments for POEs arising from a number of classes.For each location we deliberately selected 3-monthly rainfall periods when the class effect was either low or non-existent. These selections were based on p-values > 0.2 when applying the multisamle Kruskal-Wallis test (KW).

We used descriptive statistics (sample size, median, variance and standard error) to characterise and rank rainfall distributions. This provided an objective basis to explore relationships between location and dispersion parameters.

Here we present case studies for a) two stations in South America with different sample sizes (time series length) and similar rainfall variances, and b) four stations in Australia with identical sample sizes but vastly different variances. For all case studies we calculated confidence limits for unconditional POEs and respective 5, 50 and 95th percentiles based on distribution-free methods (see later).

In order to provide empirical evidence about the postulated violation of normality assumptions, we evaluated Normal goodness-of-fit (GOF) assessment using three GOF tests (Kolmogorov-Smirnov, Anderson- Darling and Cramer-von-Mises tests; Stephens, 1974).

Normal-based as well as distribution-free methods used for calculating two-sided confidence limits (CLs) for percentiles of the *Y* distribution are described in Hahn and Meeker, 1991. Both methods are available in "The Capability Procedure" (PROC CAPABILITY) of the Statistical Analysis System (SAS® version 7 and latter releases).

Distribution-free CLs are based on order statistics while Normal-based CLs are derived from the sample standard deviation (*s*) and a factor (*t*) related to the non-central t-distribution (SAS 2004; Hahn and Meeker 1991). Let $Y_{(1)}, Y_{(2)}, \ldots, Y_{(j)} \ldots, Y_{(n)}$ be the time series of the prognostic variable Y, rearranged in increasing order of magnitude. $Y_{(j)}$ is referred to as the *j-th* order statistic. The lower ($Y_{(l)}$) and upper limits ($Y_{(u)}$) of the distribution-free confidence interval for the 100.p-th percentile are chosen so that $Y_{(l)}$ and $Y_{(u)}$ are as close as possible to? $Y_{([n+1].p)}$, while satisfying the coverage probability requirement (γ=100(1-α)). In some cases is necessary to relax the symmetry requirement in order to get a pre-specified minimum coverage probability (γ).

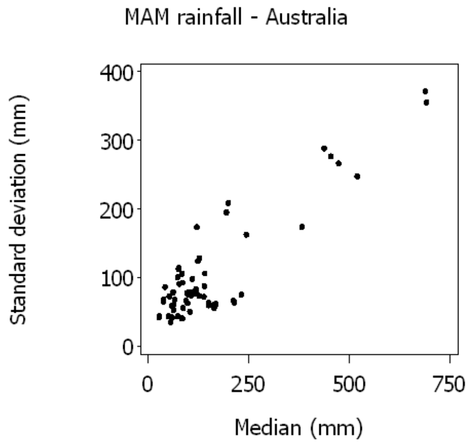
Calculation of CLs for percentiles is requested by the options CIPCTLNORMAL (Normal-based CLs) and CIPCTLDF (distribution-free CLs) of the PROC CAPABILITY (SAS® 2004). Using estimates of lower (LCL) and upper (UCL) confidence limits we calculated range (UCL-LCL) and relative range (range expressed as percent of the respective point estimates) of percentile confidence intervals (CI). Influence of variance and sample size on ranges of percentile CIs were illustrated using data from selected stations.

Confidence limits for the probability of exceeding a particular threshold *c* (*Pr[Y>c]*) are based on Binomial distributions with parameters *n* and p_c, where *n* is the series length and p_c= *Pr[Y>c].* CLs for the Binomial parameter *Pr[Y>c])* can be calculated using either Normal approximations or distribution free methods, which rely on iterative processes. Calculation of confidence limits for *Pr[Y>c]* are requested by the option CIPROBEX of the PROC CAPABILITY (SAS 2004).

Results & Discussion

Exploratory analysis

Periods selected were March-May (MAM) for Australia (KW p-value>0.20 for 75% of stations) and January-March (JFM) for South America (KW p-value>0.20 for 70% of stations). Rainfall time series from Australia had all the same length (n=103 years) while in South America series length ranged form 57 to 87 years. Graphical analysis of rainfall data for those selected periods showed that variance tends to increase with median (Figure 1). That tendency helps to explain patterns of observed CL relative ranges for percentiles (Table 2, Figure 2).



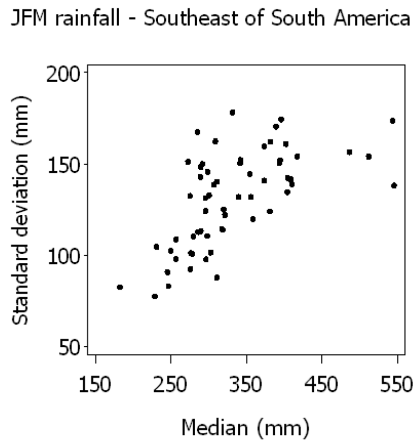


Figure 1. Relationship between 3-monthly rainfall median and respective standard deviation in Australia (64 stations) and the Southeast of South America (60 stations).

Normal goodness-of-fit assessments indicated that normality assumptions for most case study sites were inappropriate. P-values for all GOF tests applied did not exceeded 0.01 for Australia and 0.10 for South America, constituting high evidence against normality assumption. Those results indicate the need for distribution-free approaches as proposed in this paper.

Stations from Australia (n=103) were selected for different MAM rainfall variances in order to show influence of the data dispersion on estimates uncertainties: minimum variance (Nihill), intermediate variances (Walgett and Bundaberg) and maximum variance (Cairns). For evaluating influence of sample size on CL ranges, we used data from stations in South America with similar variances but different series length: Castro (n=57) and Concórdia (n=87) (Table2).

Distribution-free confidence limits for rainfall percentiles and POEs

Distribution-free percentile estimates and respective confidence limits for the selected locations (stations) and periods are shown in Table 2. Notice that Normal based CLs use a factor (*t*) related to the non-central t distribution that requires normality assumption while distribution-free CLs, based on order statistics do not require assumptions about the distribution family (e.g. Normal, Lognomal, Gamma). Therefore, distribution-free tool are particularly useful for spatial uncertainty assessments using data from a dense net of stations.

As expected, for stations with same record length, CI absolute ranges increased with variance (selected stations from Australia, Figure 2a). For stations with similar variance, CI range decreased with sample size (Castro and Concórdia, Brazil, Figure 2b). As large absolute CI ranges for medians are frequently associated with high variances/high medians, when CI absolute ranges are expressed as percent of respective median estimates, relative ranges tend to become similar, even amongst locations with different data dispersion (Table 2, Figure 2).

Graphical displays of POEs and respective CLs are a simple and useful way for summarizing information uncertainties arising from probabilistic forecasts (Figure 3 and 4).

Vertical and horizontal lines can be used to highlight CL ranges at some specified percentile. Segments defined by interception between horizontal lines placed at a specific POE level (e.g. 0.50), the POE and respective confidence bands corresponds to the POE CL. Similarly, a vertical line placed at the correspondent percentile (e.g. POE=0.50 ⇒ percentile=median) defines the respective percentile CL.

Table 2. Distribution-free 3-monthly rainfall percentile estimates and respective confidence limits (CLs) for sample stations in Australia and the Southeast of South America.

Station	Period	N	STD (mm)	Percentile (%)	Estimate (mm)	Lower CL (LCL,mm)	Upper CL (UCL,mm)	CI range[a] (mm)	Relative Range[b] (%)
Nhill (Australia)	MAM	103	41	5	34	26	40	14	41.18
				50	85	74	90	16	18.82
				95	161	146	252	106	65.84
Walgett (Australia)	MAM	103	76	5	11	0	21	21	190.91
				50	100	77	116	39	39.00
				95	264	221	328	107	40.53
Bundaberg (Australia)	MAM	103	163	5	88	53	96	43	48.86
				50	244	211	312	101	41.39
				95	595	470	907	437	73.44
Cairns (Australia)	MAM	103	371	5	325	201	378	177	54.46
				50	689	631	828	197	28.59
				95	1377	1277	2281	1004	72.91
Castro (Brazil)	JFM	57	154	5	274	177	354	177	64.60
				50	512	437	553	116	22.65
				95	733	700	949	249	33.97
Concórdia (Brazil)	JFM	87	160	5	117	64	214	150	128.20
				50	374	316	400	84	22.46
				95	680	607	903	296	43.53

[a] CL range = UCL-LCL;

[b] CL relative range= (UCL-LCL)/Percentile Estimate

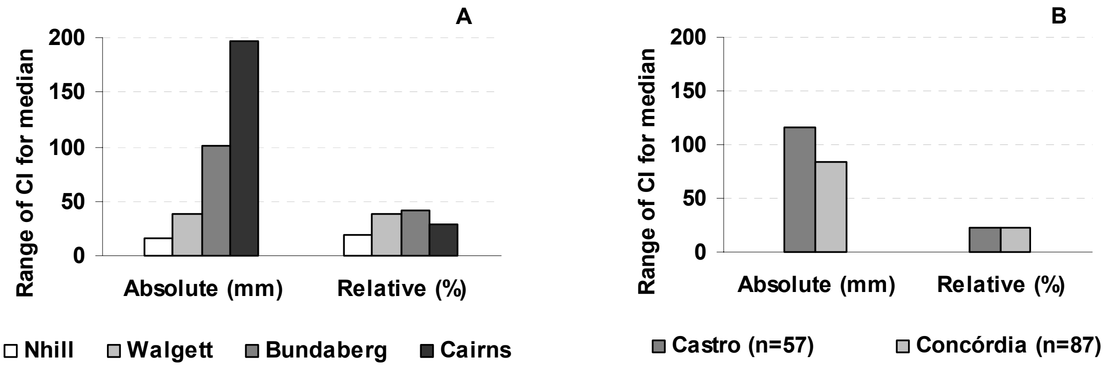


Figure 2. Influence of rainfall variability and series length (n) on the ranges of confidence intervals for median: (A) selected stations from Australia ranked accordingly MAM rainfall standard deviation (mm) and (B) stations with maximum and minimum series length in the Southeast of South America (JFM rainfall).

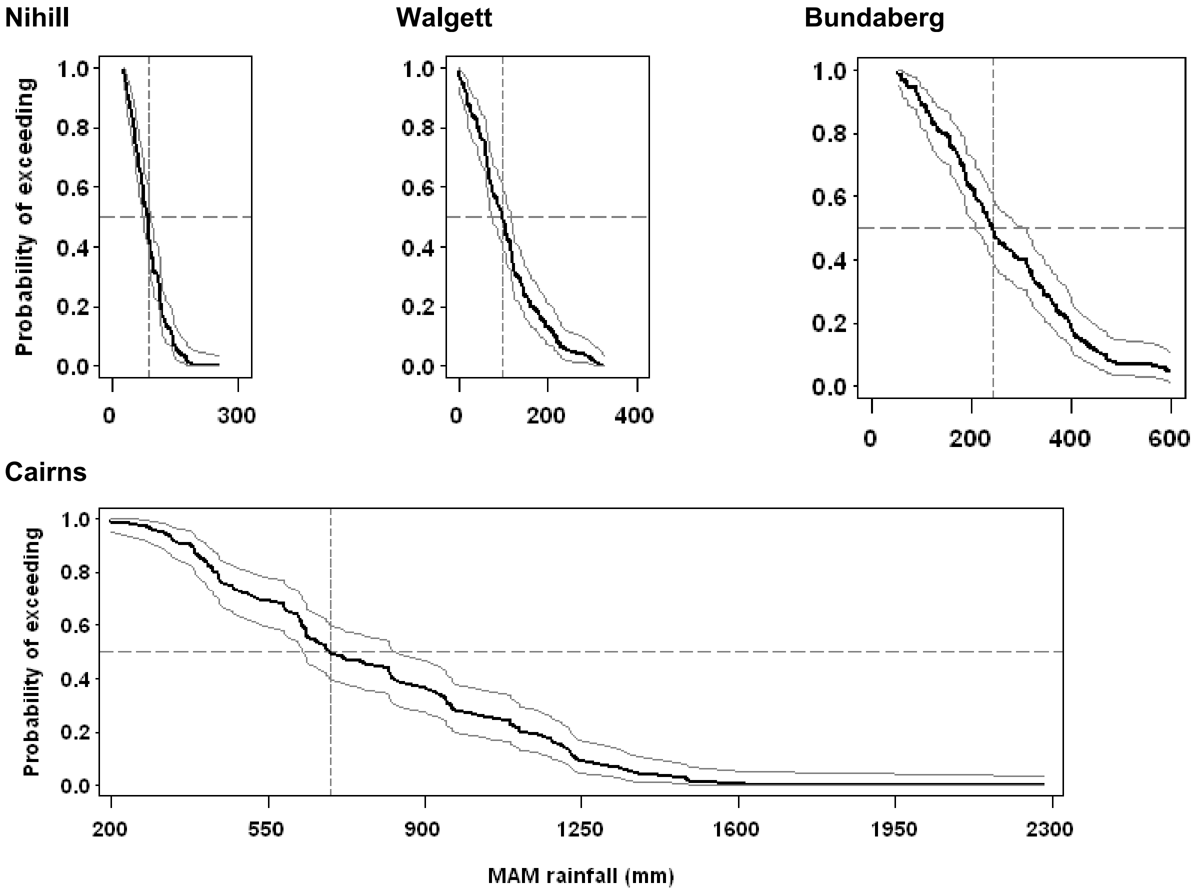


Figure 3. MAM rainfall probability of exceeding functions and respective 95% distribution-free confidence limits at four stations over Australia. Stations were selected accordingly to magnitude of MAM rainfall variances: minimum variance (Nihill), intermediate variances (Walgett and Bundaberg)) and maximum variance (Cairns). Horizontal and vertical dashed lines correspond to median and probability of exceeding the median respectively.

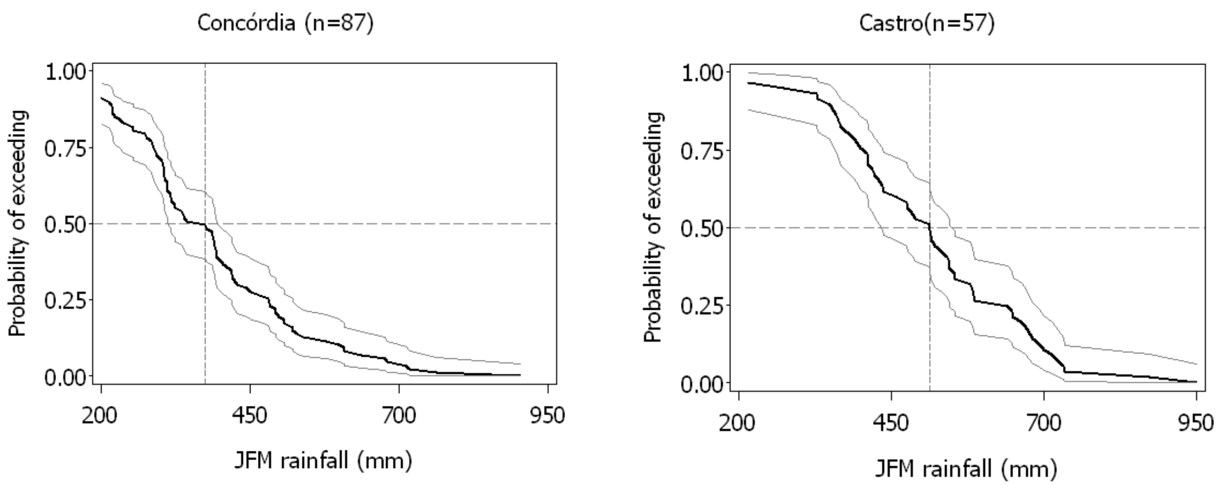


Figure 4. JFM rainfall probability of exceeding functions and respective 95% distribution-free confidence limits at Castro (n=57) and Curitiba (n=87), Brazil. Rainfall series from those stations have similar variances, but different lengths. Horizontal and vertical dashed lines correspond to median and probability of exceeding media respectively.

Concluding remarks

Parametric methods based on non-Normal distributions (e.g. Gamma, Tweedie) are also available for assessing uncertainties associated with CDF estimation. However, such methods become time-consuming for spatial uncertainty assessments which would require location-by-location checking of parametric assumptions, indicating a role for distribution-free approaches.

However, distribution-free methods for assessing uncertainties of percentiles and POE estimates are not widely used in climate science, although such procedures are readily available via statistical software packages (e.g. The Capability Procedure of SAS System),. Divulgation on such tools and their subsequent adoption by the science community are key issues that need to be addressed so that uncertainty assessments will become routine steps in the process of probabilistic climate analysis.

Confidence intervals express the level of knowledge (or ignorance) about forecasts arising from probabilistic systems. Meinke et al. (2006) stressed the importance of quantifying uncertainties for decisions makers so that the best possible decisions can be made in the face of inevitable uncertainty. This is particularly important when only short climate time series are available.

References

Hahn, G. J. and W. Q. Meeker, 1991: *Statistical Intervals: A Guide for Practitioners*. New York: John Wiley & Sons, Inc.

Jolliffe, I.T., 2004: Estimation of uncertainty in verification measures. *Proc. International Verification Methods Workshop*, Montreal, Canada. [Available at http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/Workshop2004/home.html].

Maia, A.H.N., Meinke, H., Lennox, S. and R. C. Stone, 2006. Inferential, non-parametric statistics to assess quality of probabilistic forecast systems. *Monthly Weather Review*, in review.

Meinke, H., Howden, M. and Nelson, R., 2006. Integrated assessments of climate variability and change for Australian agriculture – connecting the islands of knowledge. *3rd Biennial meeting of the International Environmental Modelling and Software Society*, 9-12 July 2006, Burlington, Vermont, USA, in review.

Podestá, G., Letson, D., Messina, C., Rocye, F., Ferreyra, A., Jones, J., Llovet, I., Hansen, J., Grondona, M. and O'Brien, J., 2002, Use of ENSO-related climate information in agricultural decision making in Argentina: a pilot experience. *Agricultural Systems*, 74, 371-392.

Potts, J.M., C. K. Folland, I. T. Jolliffe, and D. Sexton, 1996: Revised "LEPS" scores for assessing climate model simulations and long-range forecasts. *Journal of Climate*, 9, 34-53.

SAS Institute Inc., 2004: *SAS/QC® 9.1 Users Guide*. Cary, NC: SAS Institute Inc.

Stephens, M. A., 1974: EDF Statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69, 730-737.

Stone, R. C., 1992: *SOI phase relationship with rainfall probabilities over Australia*. Unpublished PhD Thesis, University of Queensland, Brisbane, Australia.